Rough surface contact analysis by means of the Finite Element Method and of a new reduced model

Vladislav Yastrebov, Julian Durand, Henry Proudhon, Georges Cailletaud

Rough surface contact analysis by means of the Finite Element Method and of a new reduced model

Vladislav A. Yastrebov, Julian Durand, Henry Proudhon, Georges Cailletaud *

Centre des matériaux, Mines ParisTech, CNRS UMR 7633, BP 87, 91003 Evry cedex, France

* Corresponding author.
E-mail addresses: vladislav.yastrebov@mines-paristech.fr (V.A. Yastrebov)
henry.proudhon@mines-paristech.fr (H. Proudhon)

A B S T R A C T

This article presents two approaches of a normal frictionless mechanical contact between an elastoplastic material and a rigid plane: a full scale finite element analysis (FEA) and a reduced model. Both of them use a representative surface element (RSE) of an experimentally measured surface roughness. The full scale FEA is performed with the Finite Element code Zset using its parallel solver. It provides the reference for the reduced model. The ingredients of the reduced model are a series of responses that are calibrated by means of FEA on a single asperity and phenomenological rules to account for asperity–asperity interaction. The reduced model is able to predict the load–displacement curve, the real contact area and the free volume between the contacting pair during the compression of a rough surface against a rigid plane. The CPU time is a few seconds for the reduced model, instead of a few days for the full FEA.

Keywords:
Roughness
Normal mechanical contact
Real contact area
Free volume
Finite Element Method
Reduced contact algorithm

1. Introduction

It is now widely accepted that the contact between two surfaces is in fact a one to one contact between many asperities, depending on the roughness of the contact pair. This represents a strong deviation from the perfect contact assumed in the Hertz theory, the real contact area being usually significantly smaller than the apparent contact area. This has important consequences for a lot of applications as thermal and electrical conductivity, friction, wear processes or sealing. The detailed study of this problem needs to develop a quantitative description of the real surface profile, and techniques to elucidate the way contact between individual asperities will develop during the loading. They should involve a good description of the local material behaviour and of the load redistribution.

Many models describing surface roughness have followed the pioneering work of Abbot and Firestone [5]. They attempt to characterise roughness by a series of indicators, such as the arithmetic average of vertical deviations R_a and the mean line m, the root mean squared R_s or the standard deviation σ. However, the surface cannot be fully described using only a surface profile in the vertical direction. Bhushan thus proposed to introduce spacing parameters [6]. These parameters can be the number of local maxima per unit length (N_p), the number of surface mean line crossing per unit length (N_0), or the mean peak spacing (A_R). Most of these works are based on the roughness measurement along a line (1D measurements).

The investigation of the contact itself classically follows two types of approach, either stochastic or deterministic. One of the first models has been proposed by Greenwood and Williamson [1], who assumed that the asperity summits are spherical with a constant radius, asperities deform elastically and their height follow a Gaussian distribution. Statistical models have had a considerable impact on contact analysis and have been considered by many authors, e.g. [7,2,8,9]. The main advantage

Table 1
Norem chemical composition.

Elements	C	Mn	Cr	Si	Ni	Mo	N_2	Fe	Co	B
Composition	1.23%	4.3%	25.4%	3.38%	4.0%	2.0%	0.17%	Bal	0.01%	0.001%

of these approaches is that they are able to provide a quick and rather good estimation of the global response — the contact area is almost proportional to the applied normal load, which corresponds to an experimentally observed fact. Nevertheless, these models do not take into account the real geometry of the surface and the interactions between asperities. Deterministic approaches were then developed to introduce a more precise geometric description. Representative models of asperities using mathematical functions were built to describe the behaviour for elastic, purely plastic [10] or elastoplastic deformations [11], and to include interactions between asperities [3].

The Finite elements method has been used to solve the contact problem for artificial fractal surfaces [4]. Starting from roughness measurements, synthesised fractal surfaces were also used in the studies of C. Vallet et al. ([12] and [13]) where they used a numerical procedure [14] to solve the contact problem. This is an interesting research direction, since, after the seminal work due to Mandelbrot [15], many studies have been made to demonstrate the self affine fractal character of surfaces [16–20]. One of the main result that is now admitted is the linearity between the real contact area and the applied load. This remarkable property is verified for any material behaviour and any roughness.

The article is organised as follows. The next section aims at introducing the material behaviour and the description of the geometry of a real surface, a technique to determine a representative surface element is proposed. Height, width and spatial distribution of asperities are the key parameters that characterise surface topography and can be used for real contact area analysis. However to analyse the free volume between a rough surface and a rigid plane, valleys and holes distributions have to be taken into account. The finite element analysis of an isolated asperity and of a regular network of identical asperities is presented in Section 3. From those results, a reduced model for rough surface contact will be further deduced. Section 4 presents the results of a large parallel finite element analysis of a representative surface element in contact with a rigid plane. This analysis is considered as a reference result used for validation of a simplified approach, proposed in Section 5. A general discussion follows in Section 6, with comments on the global displacement–load response, the real contact area and the free volume left between the asperities. The simplified finite element algorithm for analysis of contact with a rigid plane is presented in Appendix A.

2. Contact problem and rough surface analysis

2.1. Description of the contact problem

The motivation of the study comes from the problem of valves used in nuclear power plants where, due to extremely high pressure of fluids, the sealing function is ensured via steel/steel contact. Because of the roughness of the steel surfaces in contact, the tightness is not perfect and micro-leakage can be observed. The quantitative study of the leaks can be made by a simple phenomenological approach, looking for an heuristic relation between the roughness characteristics and the leakage. Nevertheless, a true predictive model can only be developed by characterising the free space between the surfaces, and by determining how fluids behave in real environment. That is the reason why, in addition to the classical results concerning the real contact area evolution with applied normal load, the free volume left between the surface is also studied. The numerical simulations will represent a contact between a steel specimen in Norem (Table 1) and a rigid flat surface. Norem is a hardfacing material developed by Electric Power Research Institute.

Classical elastoplastic constitutive equations are introduced to model the material. The behaviour is time independent, the only hardening variable is isotropic as described by the following equation:

$$\sigma_Y = R_0 + Q\left(1 - e^{-bp}\right) \tag{1}$$

The material parameters allow the model to successfully represent the tensile loading behaviour observed experimentally (Fig. 1). The same parameters are used for all the computation through the whole paper.

All Finite Element Analyses have been performed with implicit parallel Finite Element code Zset [21], using the framework of updated Lagrangian formulation. A comparison between small and large deformation formulations has shown that both force–displacement and contact area–force curves are strongly influenced by these assumptions. The contact area computed with the small deformation assumption is about 35% higher than with large deformation for the same force.

Contact is considered to be frictionless. A preliminary study has been carried out to estimate the influence of friction on contact characteristics. Whatever the asperity shapes, the force/displacement and volume/force curves are left unchanged for frictionless case and for Coulomb's friction coefficient 0.15, the last corresponds to non-lubricated friction between steel and sapphire. The principal difference is that in average the contact area is 10% smaller if Coulomb's friction is taken into account for a given force. All 3D simulations have been performed with a simplified contact algorithm that provides fast and robust sequential and parallel resolutions of the contact problem (see Appendix A).

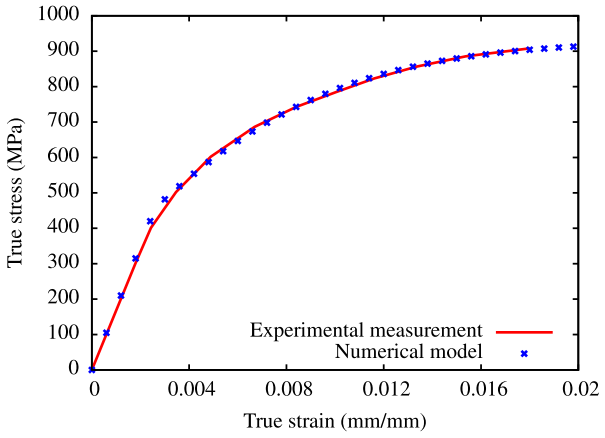

Fig. 1. Material behaviour, experimental measurement and numerical model ($E = 175$ GPa, $\nu = 0.3$, $R_0 = 442.7$ MPa, $Q = 493.5$ MPa and $b = 242.2$).

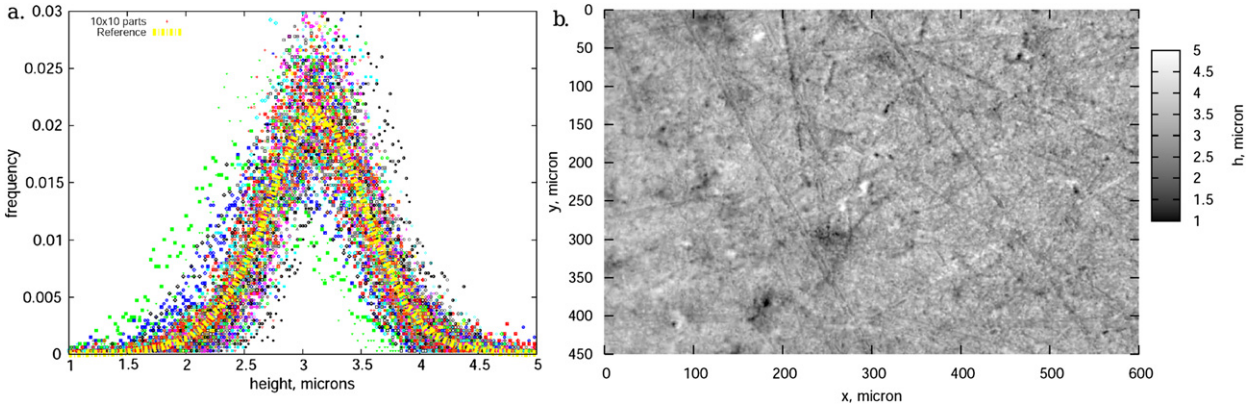

Fig. 2. (a) Peak distribution observed on the full rough surface (dashed line) and on the reduced area (points); (b) roughness map of the studied machined rough surface.

2.2. Rough surface analysis

The construction of the rough surface that will be used in the numerical simulations is based on eight samples (see Fig. 2). Each of them has a size of 610 µm × 460 µm and 736 × 480 pixels. The precision of the measurement normal to the surface is 0.02 µm. Despite machining, the peak distribution does not present a specific shape, as mentioned in [1], but a simple Gauss distribution.

Quantitatively, the average roughness R_a and the root-mean-square σ of peak height are

$$R_a = \frac{1}{A}\int_A |z - z_m|\,\mathrm{d}A = 0.356\mu, \qquad \sigma = \sqrt{\frac{1}{A}\int_A (z - z_m)^2\,\mathrm{d}A} = 0.459\mu, \tag{2}$$

where A is the surface area, $z = z(x, y)$ the height, $z_m = 3.11\mu$ the mean plane. A detailed analysis of the surface shows that the distribution of global peak heights also follows a normal distribution, with $z_m = 3.55\mu$, $R_a = 0.391\mu$, $\sigma = 0.589\mu$ (see Fig. 3). However the distribution of the local asperity heights and of the asperity radii follow a non-symmetric Gamma distribution. This can be due to the fact that, in agreement with the fractal nature of roughness, small asperities should be locally orthogonal to the surface: such a topography is hardly simulated, this is why, in the next realisations, all the asperities are assumed to be orthogonal to the mean plane. Another reason explaining the non-Gaussian character of the peak distribution is likely related to the detection procedure. To detect asperities on the rough surface, all neighbouring data points have been analysed: peaks correspond to points for which all neighbouring points are lower. The radius of the asperity foundation is computed by averaging distances in eight directions, assuming an axisymmetric shape for the asperities.

A surface investigated at a small scale exhibits very high roughness, each asperity or valley being eventually represented by only one measurement point. On the other hand, as shown below, the geometrical discretisation of each peak must be fine enough to get reasonable result in the FEA (at least 10 × 10 elements per asperity, see also [22]). A compromise can

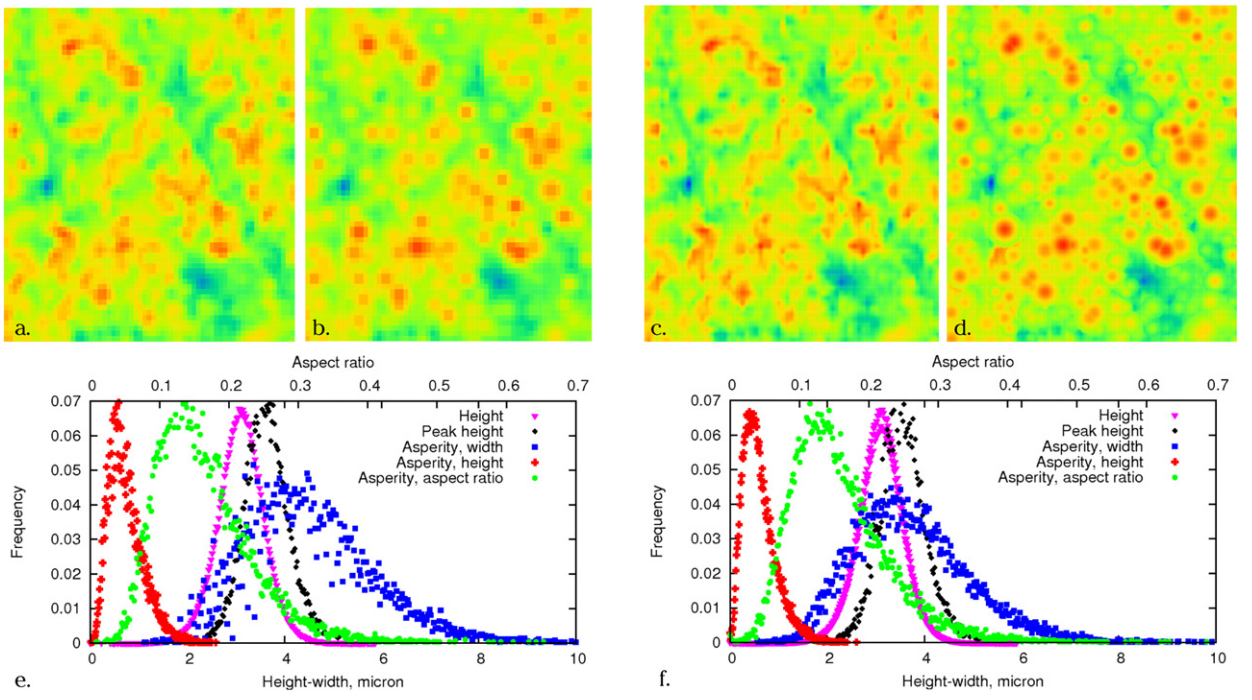

Fig. 3. Construction of a representative rough surface element. Top, from left to right: (a) original part of rough surface; (b) sinusoidal asperities plotted on the rough surface; (c) the same part smoothed with bi-cubic Bezier surfaces; (d) corresponding sinusoidal asperities. Below: (e) surface data distributions for full rough; (f) full smoothed surface.

been achieved by enriching the rough surface using a bi-cubic Bezier smoothing step (with the parameter $\alpha = 0.5$) [23]. This treatment preserves all the measured points and refines the surface between them. The number of segments on the smoothed surface is chosen to be 25 times larger than on original surface. The beneficial effect of this approach is demonstrated in Fig. 3, that shows a comparison between the original surface and the surface after Bezier smoothing. The size of the presented sample is only one hundredth of the full surface. The corresponding distribution of asperities approximated by sinusoidal shapes is projected on the rough surface. The distributions of the critical parameters (height, peak height, asperity height, width and aspect ratio) are represented for the full surface and the smoothed one. More than 16 000 asperities have been detected on the original surface. This number reaches 22 900 for the smoothed surface. The representative surface element smoothed with Bezier curves consists of 230 asperities. The smoothing procedure is found to slightly change the distributions: mean values of asperity width, height and aspect ratio are smaller for smoothed surface.

2.3. Representative surface element

For the following applications, the Representative Surface Element (RSE) is determined in terms of roughness representativity. The RSE is then chosen as the smallest part of the full measured surface, for which the height distribution follows the height distribution of the full surface. The peaks height characterisation is not sufficient to characterise the surface. The spatial distribution in the contact plane needs also to be respected. We found that the first point is the most important, and this is why a particular attention was taken to define it. Concerning the repartition in the contact plane, we only tried to include the same number of peaks in the final surface than in the measured surface.

The reduced surface is built from one of the eight measured surfaces. It is divided into 16, 25, 100 and 144 domains. For each domain, the distribution analysis has been performed. For the case of 100 domains, 23 of them have height distributions that are close enough to the height distribution of the initial surface (see Fig. 2(a)). On the other hand, further division of the surface leads to height distributions which are far from the reference one. This is why the 100 subdomains case is chosen, and one among the 23 subdomains is selected for the finite element analyses. Its dimensions are 54 µm $\times$ 63 µm (slightly larger than one hundredth of the full surface).

Remark 1. The size of the selected surface is comparable to the grain size. This would justify to consider crystal plasticity instead of a classical von Mises criterion to describe the material (see e.g. [24]). It would be expected to find other types of local behaviour, and a different percolation scheme [25].

The geometry used for the finite element model is obtained by considering a Representative Volume Element (RVE) representing the bulk material under the surface. One face of this RVE is nothing but the RSE. The third dimension (normal

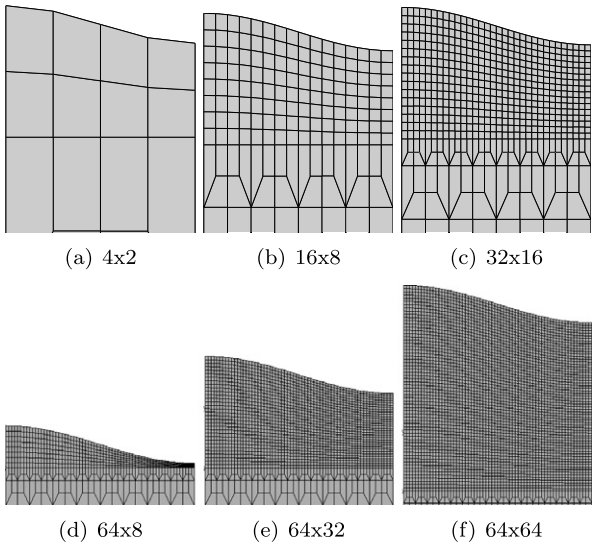

Fig. 4. Different meshed to describe the asperity geometry, (a), (b), (c) varying the element size and (d), (e), (f) varying the number of layers under the surface.

to this surface) must be large enough to represent correctly the bulk material response and to avoid side effects from the boundary conditions. The chosen size is $h = 72$ µm.

3. FEA of the contact on a single asperity

Describing the compression of a single asperity is the starting point to build a reduced model representing the contact process with a surface featuring thousand asperities such as the one presented in Section 2.3. This problem is now investigated trying to address the relevant mesh size, the influence of the aspect ratio and of the actual shape of the asperity. The single asperity response such as the load–displacement, contact area and free volume evolutions will further nurture the simplified model presented in Section 5.

3.1. Mesh influence

The relevance of the contact area predicted by FEA is connected to the number of nodes in contact and thus strongly depends on the finite element mesh. Several computations were performed on axisymmetric meshes with respectively 2×1, 4×2, 8×4, 16×8, 32×16 and 64×32 elements in the region close to asperity surface (see Figs. 4(a)–(c)). The influence of the number of element layers used under the asperity surface has also been studied. The number of layers has been modified from 2 to 64 (Figs. 4(d)–(f)).

For each case, the ratio between real and apparent contact area (A/A_0) and the ratio between the real and the initial free volume (V/V_0) were plotted as a function of the applied load. They are not shown for the sake of brevity, but the study is summarised in Fig. 5, that presents the load–displacement response. One can observe that the difference between the various meshes is more sensitive for small displacements/small forces, when the excessive rigidity plays a major role. On the other hand, the number of layers has little influence on the global response. The force is overestimated by a few percents. Accordingly, the contact area is underestimated, and the free volume is overestimated. A reasonable convergence is reached with 32×16 elements. In the following, all computations will use a 32×16 configuration (32 elements to describe the asperity shape and 16 layers under the asperity).

3.2. Influence of the boundary conditions

The study which is shown here deals with the influence of the surrounding of the asperity on its global response. This can be represented by various boundary conditions, either in 2D-axisymmetric or in 3D computations. An important parameter is the space around the asperity, that is figured by the relative size of the asperity and of the full mesh. For each calculation, the boundary conditions are imposed at different distances from the asperity summit (see Fig. 6). For all the different meshes, the size of an asperity is fixed ($width = 10$ µm and $height = 1$ µm). A thin mesh (32×16 elements) is used for the asperity summit and a coarser one is define for the volume meshed around the asperity. Each node which belong to the asperity summit (and so to the thinner mesh) is displaced, according to a sinus curve in 2D (see Eq. (3)) and to a sinusoidal surface in 3D (see Eq. (4)), in order to obtain the asperity geometry desired:

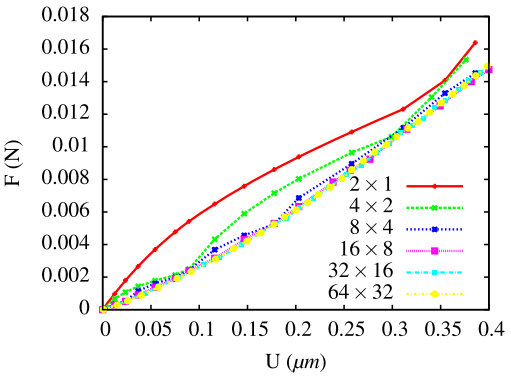

Fig. 5. Effect of the mesh density on the force–displacement response in a single asperity contact.

Fig. 6. Schema representing the different axisymmetric problems.

$$\text{2D axisymmetric computations: } y \times 1.2 + y \times 0.2 \times \cos(\pi \times x) \tag{3}$$

$$\text{3D computations: } y \times 1.2 + y \times \left(0.2 \times \cos\left(2\pi \times \sqrt{\left(x^2 + z^2\right)}\right)\right) \tag{4}$$

For the 3D case, two types of patterns were considered, either square or hexagonal (see Fig. 7). Thanks to symmetry considerations, the mesh was reduced to a quarter in the first case and two neighbouring quarters in the case of an hexagonal network.

The respective behaviours are highlighted through three plots, the relative contact area versus force, the relative free volume versus force, and the force/displacement response (see Fig. 8). The boundary conditions does not affect so much the contact area, as shown in Fig. 8(a), since both “1L” and “4L” cases have the same solution. The difference observed is rather due to the pattern (12% difference between the square and the hexagonal pattern). For the three “compact” meshes, where the surface of the asperity is between 80% and 100% of the whole surface, the difference between the solutions remains lower than 10% for the force–displacement and the volume–force curves. For these three cases, the stress level is high, and plastic deformation occurs. The amount of plasticity is lower for the “4L” case, consequently the force reaches larger values. The free volume is also larger in this latter case, since the geometrical change of the asperity has a small effect on the whole mesh.

3.3. Influence of the asperity shape

Models found in the literature generally use a particular shape for all asperities. For example, Greenwood–Williamson model [1] uses spherical asperities, while Bush et al. model [2] uses ellipsoidal asperities and ZMC model [11] features a parabolic shape. However, an accurate study on the shape influence has never been made. Here, the response of four typical asperity shapes will be compared: sinusoid, paraboloid, ellipsoid and conic (see Fig. 9(a)).

It is worth noting (see Fig. 9(b)) that the contact area evolution as a function of the load is linear and identical for each cases (the difference is less than 2%). The volume variation is slower for massive asperities: less variation with the elliptic shape, then for the parabola. Sinus and conic shape have comparable amounts of matter, but the free volume changes less with the sinusoidal inclusion. This rather intuitive result confirms that inclusions with a tiny top will crush more rapidly than those with a larger summit (Fig. 9(c)).

The influence of the asperity aspect ratio (that is both height/width ratio and width) has also been investigated, based on the study of real rough surface distributions done in Section 2. Several computations were carried out, with ratios of 0.05, 0.075, 0.1 and 0.15. For these computations, the asperity width are taken equal either to 4 µm or to 6 µm, and the sinusoidal shape is retained.

As shown in Figs. 10(b), (c), (e), (f), the contact area and the free volume evolutions weakly depend on the height/width ratio (maximum difference of 10% between the various cases). However, the force/displacement response is ratio dependent, with a non-linear relation between the ratio and the force obtained for a given displacement (Figs. 10(a), (d)).

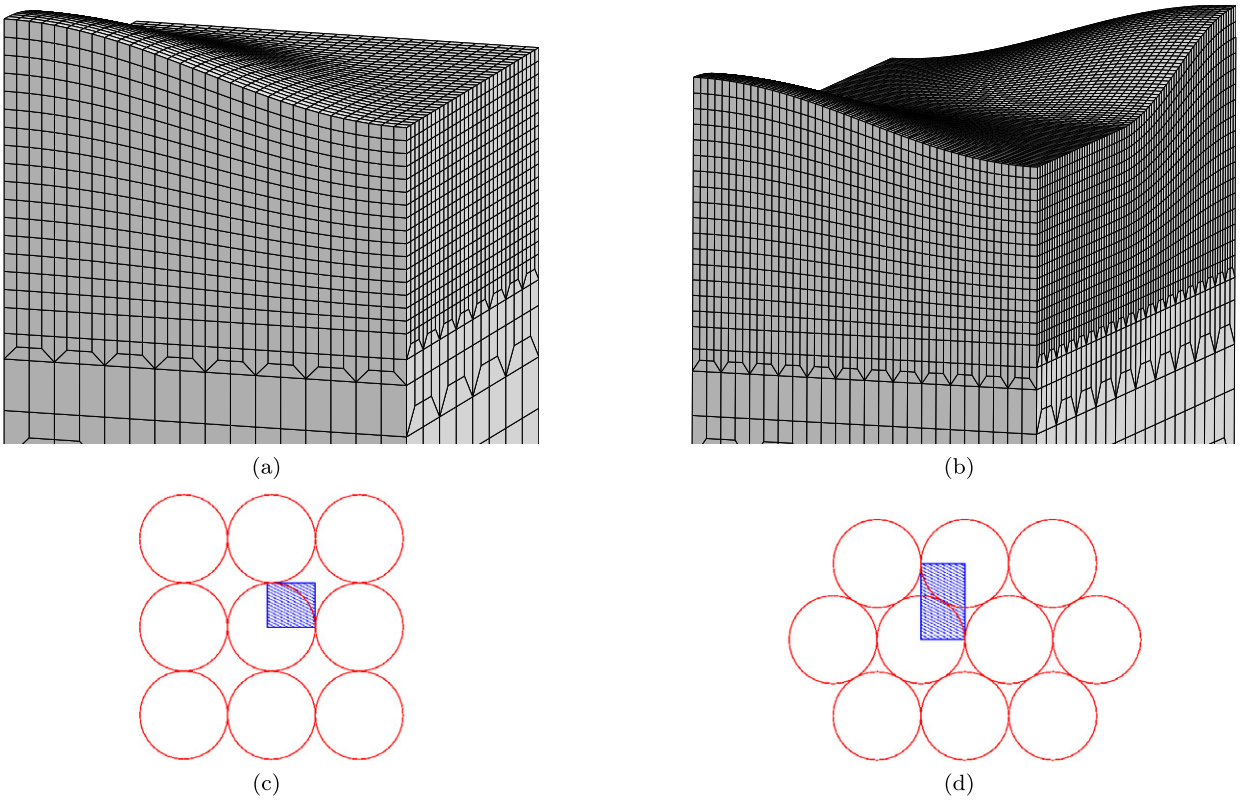

Fig. 7. 3D patterns: (a) 3D mesh used for the case of a square distribution; (b) 3D mesh used for the case of an hexagonal distribution; (c) a view to the square pattern; (d) a view to the hexagonal pattern.

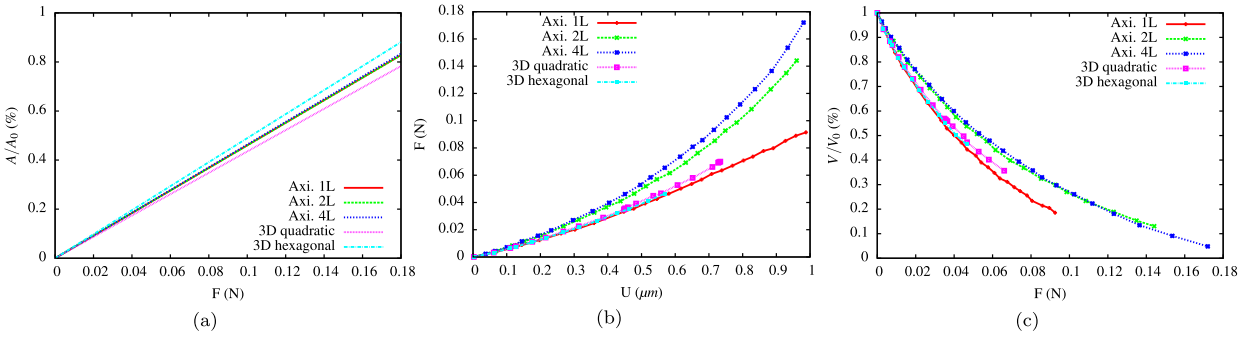

Fig. 8. Influence of boundary conditions: (a) real contact area evolution; (b) force/displacement response; (c) free volume evolution.

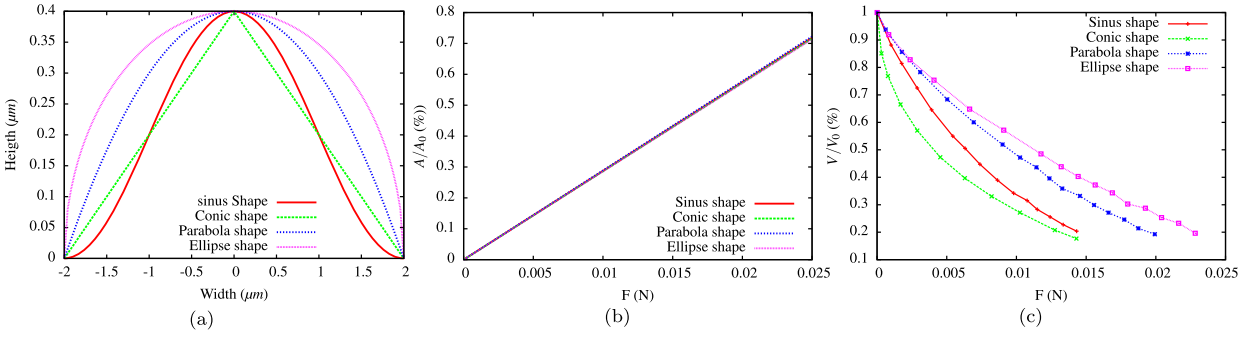

Fig. 9. Influence of the shape: (a) different shapes used; (b) contact area evolution; (c) free volume evolution.

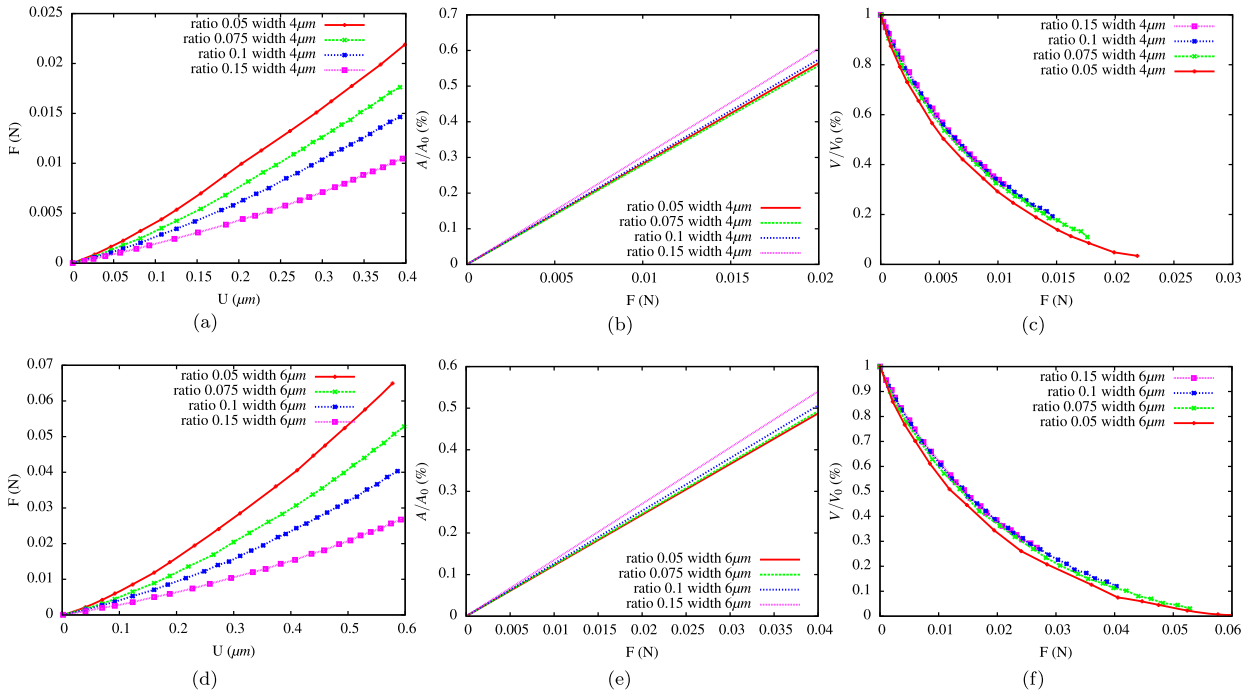

Fig. 10. Influence of height/width ratio: (a), (b), (c) respectively contact area, force/displacement and free volume for a width equal to 4 µm. (d), (e), (f) respectively contact area, force/displacement and free volume for a width equal to 6 µm.

4. Large scale contact: finite element analysis

The study of a single asperity and of a periodic microstructure of identical asperities provides us with a deep understanding of the local deformation process. However the global deformation is much more complicated: asperity shapes are not so simple as it has been considered, the width, height and vertical position follow distribution rules and, due to the fractal nature of rough surfaces, a large asperity is "decorated" by series of smaller ones. Moreover close asperities interact. Simple geometrical and analytical models cannot account for all these aspects. So, full scale finite element analyses are needed to better understand the global deformation of rough surfaces. The loading is assumed to be quasi static, neither dynamic nor viscosity effects have been taken into account.

4.1. Mesh and boundary conditions

The studied contact problem requires a very fine mesh close to the surface to represent complicated geometry of a rough surface and to capture high stress gradients in the zone close to the surface. On the other hand the total height of the mesh has to be sufficiently high in order to represent bulk material deformation and to avoid edge effects. To derive an adequate finite element mesh (fine close to the surface and coarse everywhere else), a special composition of 8-node bricks, 6-node prismatic and 5-node pyramidal elements have been combined in a special arrangement, in order to produce a very fast refinement. One switches from one to 9 elements within one finite element layer, so that refinement is of 9^n elements at n-th layer (Fig. 11(a)). The finite element mesh used for the analysis has 964 000 nodes, and more than 945 000 nodes are situated in 8 layers forming a regularly meshed zone adjacent to the contacting surface (Fig. 11(b)). The rough surface itself consists of 105 625 nodes. Consequently, the average number of nodes used for one asperity is about 460. Moreover, we remarked that asperities which come in contact first have generally large dimensions, and so possess a higher number of nodes. The element size of the regularly meshed surface is 0.167 µm × 0.1944 µm. It also has to be pointed out the fact that only 1.5% of the elements are used to represent the bulk material where stress is more or less homogeneous.

To treat such a large problem in a reasonable time, the use of parallel computation is highly required. Here a classical implementation of the FETI (Finite Element Tearing and Interconnection) method [21] is used together with the sparse DSCPACK solver for the local resolution [26]. Finite element mesh has been divided into 16 equal sub-domains (highlighted by different colours in Fig. 11(b)) in a manner to avoid rigid body motion. Symmetric boundary conditions are applied at all side facets, and a vertical displacement towards the rigid plane is prescribed on the bottom facet. All other displacement components of this facet are blocked, i.e. the problem is symmetric over all lateral facets of the RVE. To analyse the problem a workstation with 8 bi-core processors Intel Xeon X5550 2.67 GHz and 160 Gb of RAM has been used. The problem requires a very large number of computational steps, mainly because of the very large displacements at the interfaces between subdomains, and also due to material and geometrical non-linearity. An updated Lagrangian scheme and a corotational

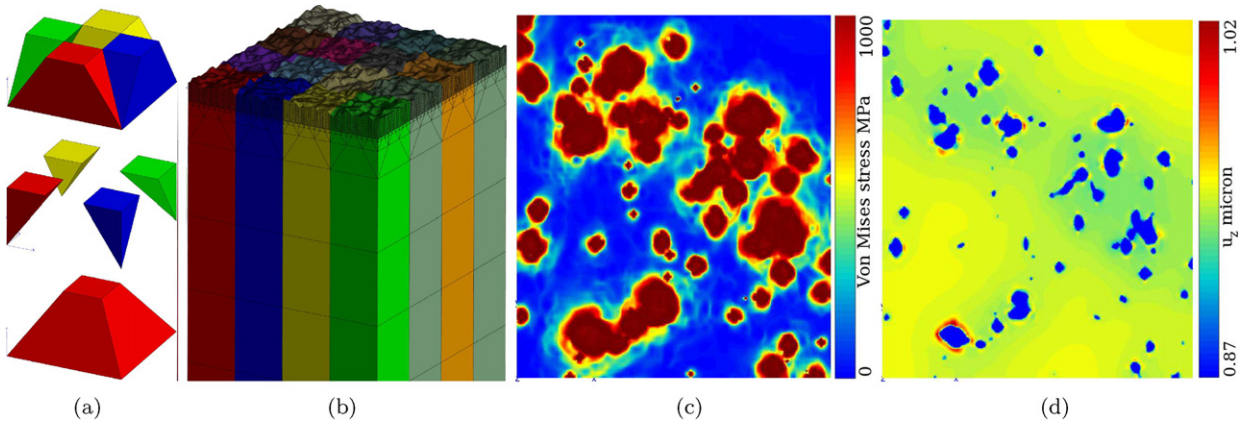

Fig. 11. Full finite element analysis: (a) representation of the transition mesh allowing an efficient refinement from 1 to 9 elements; (b) mesh of the rough surface 54×63 µm^2; (c) distribution of the vertical displacement for a contact area of 6% (blue regions correspond to contact area); (d) distribution of the von Mises stress for a contact area of 6%.

framework are then used to account for finite strain and large displacements. Material is the same which has been used in the study of one asperity. Totally, 220 time steps have been computed, which correspond to a displacement of the sample of 1 micrometer. The convergence for each time step is reached within 3–4 iterations and requires almost 1 hour.

4.2. Results

Fig. 11(c) shows a contour plot of the vertical displacement of the rough surface. The colour scale has been chosen to highlight the contact area and make clear the deformation of the rough surface in a large zone close to asperities. It reveals interactions between asperities which extends at large distances from the local contact zones. It will be shown further how to account for this interaction in the reduced rough surface model. The von Mises stress distribution on the rough surface is presented in Fig. 11(d). It saturates very fast, and the contacting process results in the growth of saturated zones of von Mises stress far beyond the limits of the contacting asperities. For the ultimate displacement of 1 micrometer, the equivalent accumulated plastic strain reaches 200% in several asperities.

Global results such as contact area evolution, free volume changes, etc. will be discussed and compared with a reduced model in Section 6.

5. A new reduced model for the analysis of rough surfaces

Since the percolation analysis is strongly connected with statistics, it requires performing of numerous tests on RSE. The full finite element analyses on an RSE require long CPU times and high performance computers. They should be used essentially as reference computations, and reduced models able to derive global and local characteristics of the contact are highly desirable and of a great practical importance. Such a reduced model is presented in this section. It is based on the study of the contact of one asperity and on the real distribution of asperity shape and position according to the studied rough surface. So the model allows to analyse both global and local characteristics: reaction force, total contact area and total free volume as well as contact area and free volume topology, which is very important for percolation analysis. The main idea is to detect and represent each asperity of the original rough surface by an element (rod) with a force–displacement behaviour derived from the FEA of one asperity with a specific shape, boundary conditions and height/width aspect ratio. The interaction between asperities is ensured by displacing the neighbouring asperities accordingly to the phenomenological law which has been deduced from the study on one asperity and will be presented below.

To find and characterise asperities on the rough surface, a detection technique has been used (see Section 2). For each asperity of the rough surface, this procedure determines the position of the summit (x, y, z), together with asperity height and width. If contacting asperities are situated far from each other, then the interaction between them is negligible, so that the reaction force developing at each asperity can be easily predicted using the study on a single asperity. As an extension of the study shown in Section 3, series of computations have been carried out for different representative asperities, according to statistical data (Fig. 3). They include 6 different widths, $\{2, 3, 4, 5, 6, 7\}$ micrometers and 5 aspect ratios, $\{0.05, 0.075, 0.1, 0.15, 0.25\}$. It is worth mentioning that the distribution of asperity characteristics which first come in contact is quite different from the distributions presented in Fig. 3; for example the mean aspect ratio of asperities in contact at ultimate force of 0.5 N is 0.2–0.22 instead of 0.1. Then, for each detected asperity, the constitutive model has been interpolated from these 15 studies. However when the deformation of an asperity is confined by the deformation of surrounding asperities, the response is quite different. To account for this effect, an interaction between asperities has been taken into account. All these aspects are explained below, and the global response as well as the topography of rough surface in contact are compared with the reference result from the full scale finite element analysis.

The ultimate algorithm of the proposed simplified analysis is the following:

- starting from the initial rough surface data smoothed with bi-cubic Bezier surface, all asperities (position, global height, average local height, radius, local area) are located and retained for the following steps;
- each determined asperity is approximated by an axisymmetric asperity with a sinusoidal shape and an aspect ratio corresponding to a ratio of local height to local width;
- asperity widths are scaled in order to get the total area equal to the reference one;
- a rigid plane is gradually moved against the asperity set;
- if a global height of an asperity appears to be over the rigid plane position so it deforms according to the deduced constitutive law;
- all neighbouring asperities are displaced to follow the deformation of the contacting asperity as well as points of rough surface situated outside contact zones of asperities;
- the total force and contact area are computed as sums over all contacting asperities;
- the free volume is analysed by integration of the gap between rough surface situated outside contact zones of asperities and the rigid plane.

5.1. Constitutive models

The study performed in Section 3 allows to determine constitutive deformation models for asperities in contact. In this model, the motion of the rigid plane is governed by the vertical displacement. The force/displacement model enables us to get out the force associated to this displacement, and also contact area can be deduced for each asperity.

Asperities are supposed to be axisymmetric, the shape of asperity is chosen to follow cosinus. The constitutive laws depend on the asperity width and aspect ratio. The constitutive laws are based on FEA simulation on a single asperity and therefore, represent only the elastoplastic behaviour of the material Norem (see Section 2). Thus, for each pair of width and aspect ratio, two mathematical functions describing the responses $F(U, w, r)$ and $A(F, w, r)$ have been deduced (see Figs. 12(a) and (b)):

$$\text{Force/displacement response:} \quad F(U, r, w) = c_1 U^4 + c_2 U^2 + c_3 U$$

$$\text{Contact area/force response:} \quad A(F, r, w) = A_0 c_4 F$$

For each asperity width, c_i parameters are then approximated by a mathematical function $c_i = c_i(r)$, $i = 1, 4$, depending on the asperity aspect ratio r and involving other parameters c_{ij}, $j = 1, 3$ (see Figs. 12(c), (d), (e) and (f)). These last parameters c_{ij} are also approximated by mathematical functions depending on the asperity width w (see Figs. 12(g), (h) and (i)). Thus, starting with the determination of c_{ij}, we calculate the c_i parameters and, finally, we define the $F(U)$ and $A(F)$ responses for a specific shape of asperity.

For real rough surfaces, the free volume is not determined by asperity deformation but mainly by valleys and scars. That is why no constitutive model $V(F)$ is determined.

5.2. Asperity interaction

When a point of the rough surface comes in contact with a rigid foundation, the adjacent bulk material deforms, the general displacement being in the direction opposite to the contact area both in the contact plane and in orthogonal direction. The size of the deformed region in the contact plane is limited, and points that are far enough retain their original positions. It means that if one asperity comes in contact with a rigid foundation, surrounding asperities move outward the center of contacting asperity as well as outward the rigid plane. It is assumed that the local shape of interacting asperities remains unchanged. An analytical estimation of the surface deformation near a contacting asperity is a quite intricated task in the framework of large deformations and non-linear material. It has been then decided to use a phenomenological interaction law deduced from a set of axisymmetric calculations on a single asperity.

Let us consider a single sinusoidal asperity coming in contact with a rigid plane. The radial displacement in the neighbourhood of the contact zone is of the order of magnitude of the axial displacement, but this radial displacement does not change the contact pressure or the contact area. Its influence on the topology of the contact area and on the free volume is not significant as well, and it will be neglected further. The axial displacement of the surface close to a sinusoidal asperity (width = 4 micrometers, height = 0.4 micrometers) is represented in Fig. 13(a). This displacement can be approximated by the following function

$$u_z = a u_z^0 \sqrt{\frac{2d}{w}} \exp\left(-b \frac{2d}{w}\right) \tag{5}$$

where u_z^0 is the displacement of the asperity summit, d the distance from the center of asperity, w the asperity width and a, b model parameters which might depend on the asperity aspect ratio. In the present study, these coefficients are taken constant.

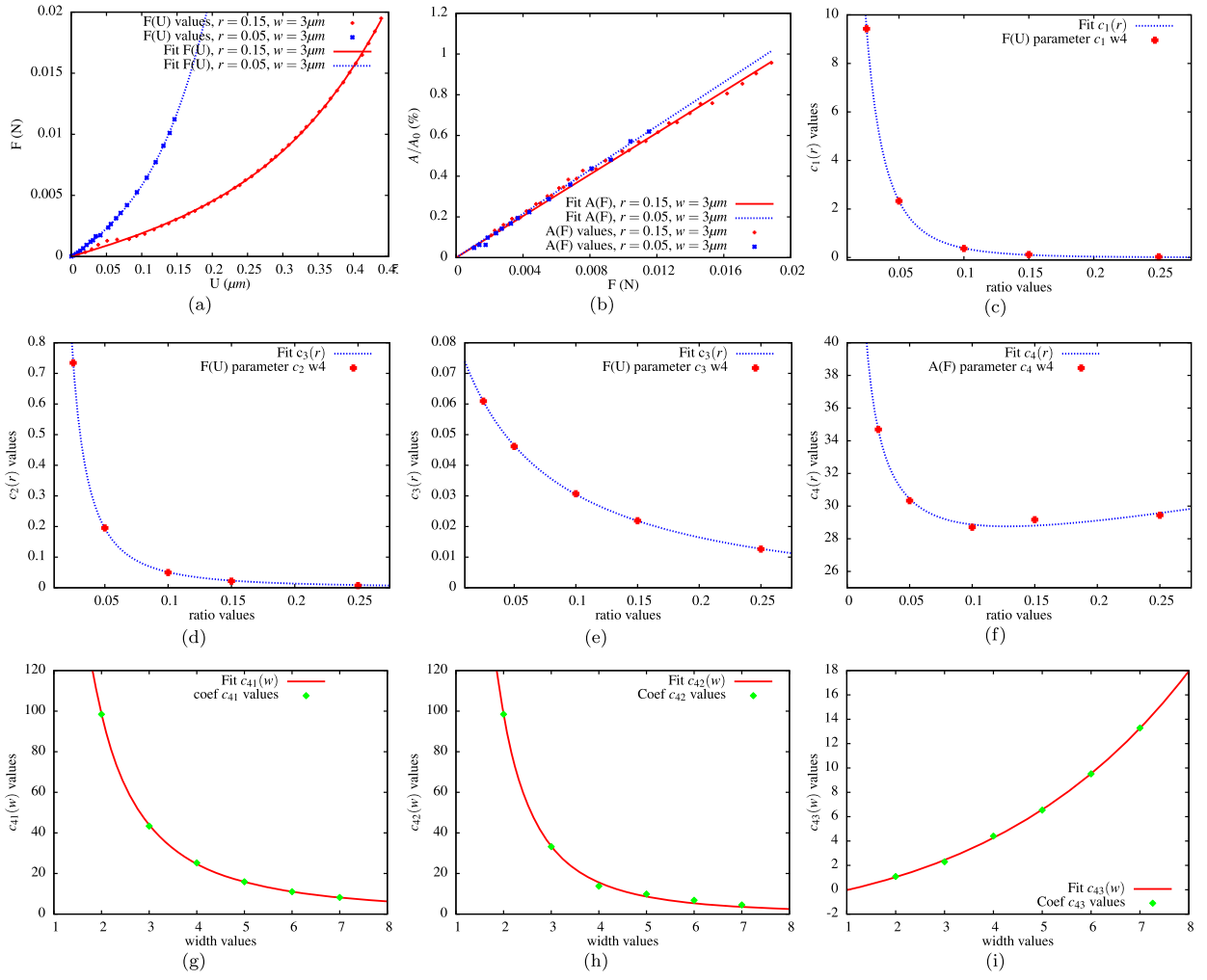

Fig. 12. (a) and (b): representative mathematical functions for $F(U)$ and $A(F)$. (c), (d), (e) and (f): mathematical functions which represent respectively c_1, c_2, c_3 and c_4 in function of asperity ratio r ($w = 4$ µm). (g), (h) and (i): mathematical functions which represent respectively c_{41}, c_{42} and c_{43} in function of asperity width w.

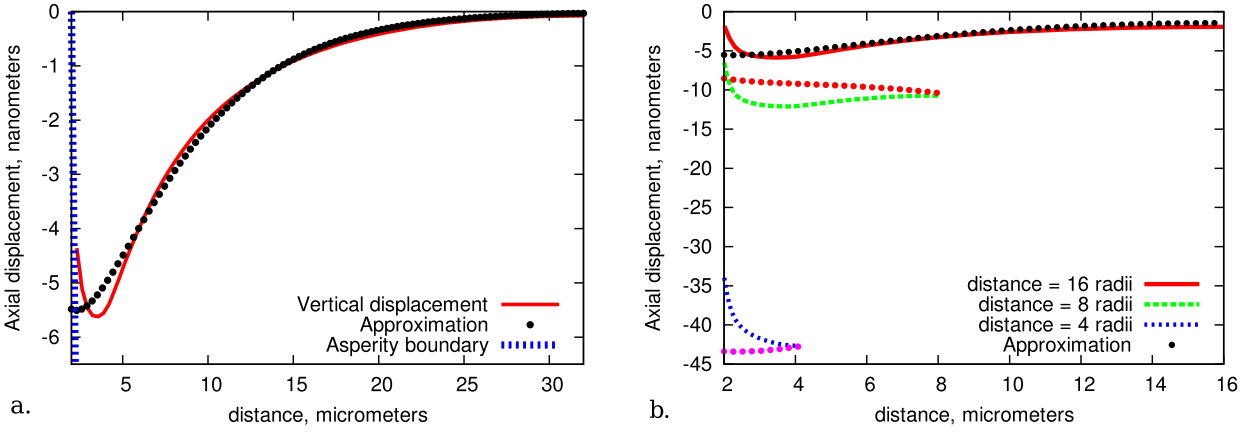

Fig. 13. Illustration of the interaction between contacting asperities: (a) axial displacement close to a single sinusoidal asperity and its approximation; (b) axial displacements for confined asperities and approximations evaluated by assuming displacement additivity.

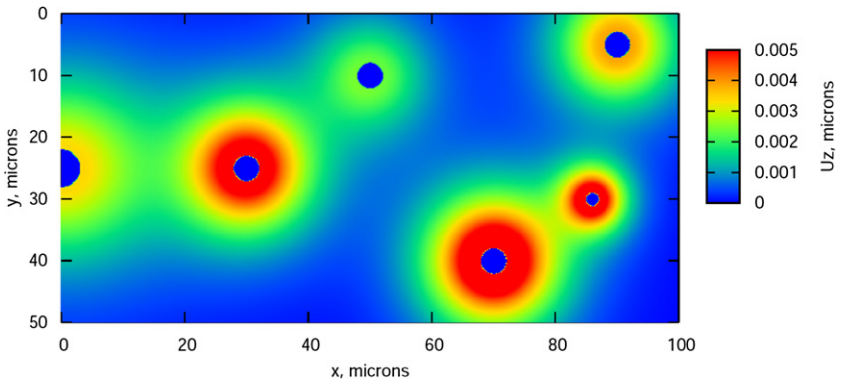

Fig. 14. An example of surface deformation computed with the phenomenological interaction law. The colour scale corresponds to the vertical displacement; blue circles embedded in coloured field designate asperity limits.

As it can be seen, the value of the axial displacement at the surface is rather small but it decreases relatively slowly and approaches zero only at a distance equal to 16 asperity radii. There is then a long range influence of the contacting asperities on the deformed surface profile (see Fig. 11(d)). To account for asperity interactions, an additivity of displacement fields arising from every contacting asperity is assumed. The proposed interaction law has been validated for the same asperity shape on several axisymmetric tests: the lateral boundary fixed in radial direction is placed at 4, 8, 16 micrometers from the center of asperity. Such boundary conditions imply approximately that there is a set of identical asperities distributed over a circle. It can be checked in Fig. 13(b) that the results provided by the upcoming phenomenological rule are in good agreement with the surface deformation computed with by FEA. However the assumption of additivity is very rough and requires a validation on numerous full 3D analyses for different asperity configurations. For the given set of scale and material properties, the parameters in the deformation law are the following: $a = 23.6410^{-3}$; $b = 0.4367$. An example of surface deformation computed by the proposed law for different asperities in contact is presented in Fig. 14.

The constitutive model for one asperity provides us with force–displacement relations $F(U)$. As demonstrated in Fig. 8, there is a strong difference between the deformation process of a single asperity and of a periodic structure of closely packed asperities. If interaction between asperities is not taken into account, then a constitutive deformation model would give the same result for both cases. This is why all the constitutive models (made explicited in Section 5.1) correspond to the case of an isolated asperity to which we add an interaction law. So in the present study, the vertical displacement of the asperity summit U is decomposed into two parts: one due to the interaction between asperities U_i and the second related to local deformation itself U_d. Consequently, the force–displacement relation becomes $F(U_d) = F(U - U_i)$. The force is a monotonically increasing function of the displacement, so that such a procedure results qualitatively in a lower force for multiple asperities (U_i is high) than for an isolated asperity ($U_i = 0$). It remains clear that the assumptions of additivity and independence of interaction remain questionable and have to be further investigated.

6. Results

A qualitative analysis of the contact between a rough RSE and a rigid plane has been studied by means of the same quantities as the analysis of a single asperity. The evolution of the contact force versus the displacement of the rigid foundation is non-linear, with a power 3.4. The maximal force reached in the large FEA is about 0.5 N, which corresponds to a displacement of 1 micrometer of the rigid plane. The RSE's dimensions are 54 µm × 63 µm, so this force corresponds to an apparent pressure of 147 MPa, which is three times smaller than the initial yield strength, however local stress is much higher, since less than 6% of the apparent contact area is in contact for the ultimate load. The evolution of the real contact area with increasing reaction force in large scale analysis is strictly linear, regardless the fact that for a single asperity with high aspect ratio this evolution is slightly concave.

Global results (force/displacement and contact area/force) obtained with the reduced model in comparison with the full FEA are presented in Fig. 15. The distribution of the contact zones predicted by the reduced model and the FEA are presented in Fig. 16 for different displacements. As expected, this plot reveals the importance of accounting for the interaction between asperities in the reduced model. For the maximum displacement, the simplified prediction made without interaction overestimates the force by 40%. In fact, considering the interaction between asperities influences the force–displacement curve almost from the beginning due to an incidental presence of two close asperities which come in contact almost at the same moment. As it can be seen from Fig. 15(a), the interaction reduces significantly the reaction force and makes the reduced model a very powerful tool to analyse rough surface contact. Varying the interaction network between asperities situated close to edges of the RSE allows to model periodic or symmetric boundary conditions (the latter is used in the present simulation). The reduced model presents a good general agreement with the reference curve from FE. The difference between the two curves is negligible for the maximum load. For moderate displacement values, the maximal relative error reaches 23% for the model with interaction. This can be explained by the approximation of the asperities by

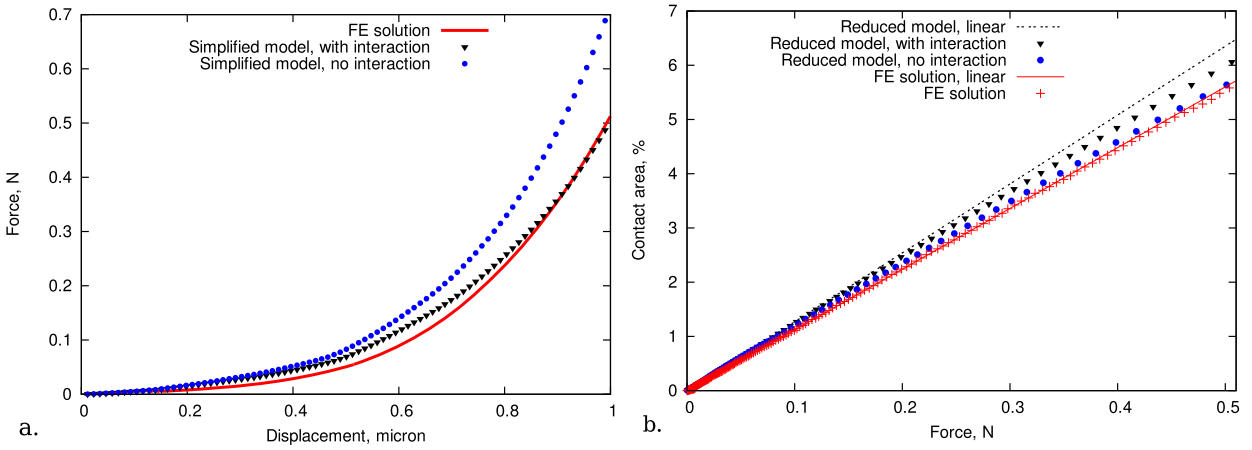

Fig. 15. Comparison between the FEA and the reduced model with and without interaction between asperities: (a) dependence of the contact reaction force on the displacement of the rigid plane; (b) force–real contact area plot.

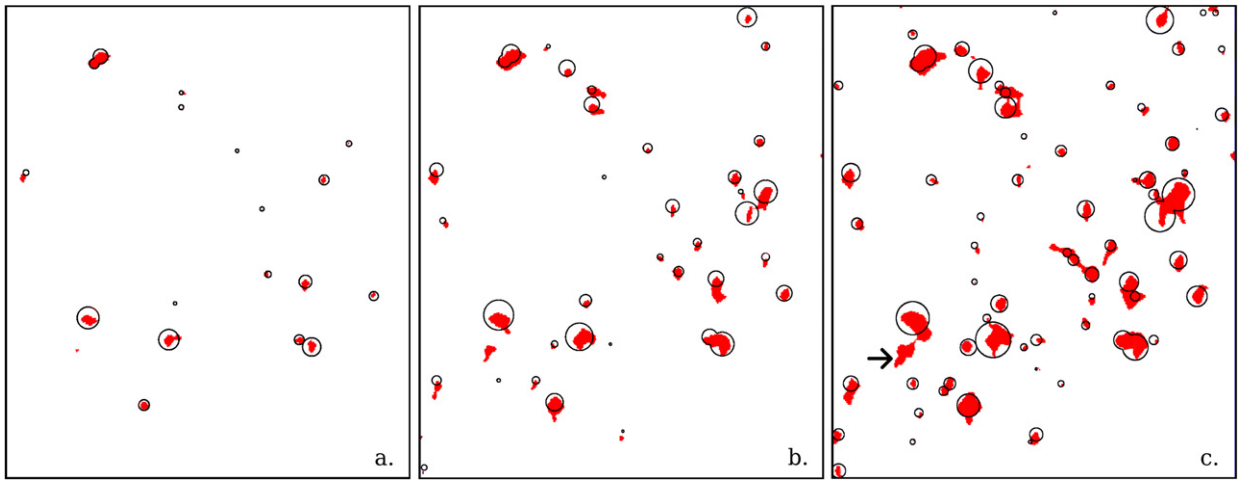

Fig. 16. Comparison of the contact area topology for the reduced model with interaction (black circles) in comparison to the full FEA (red zones) at different loading steps: (a) $u_z = 0.5$ μm, $F = 0.05$ N, contact area 0.56%; (b) $u_z = 0.75$ μm, $F = 0.19$ N, contact area 2.10%; (c) $u_z = 1$ μm, $F = 0.5$ N, contact area 5.58%. The arrow marks the only non-detected contact zone.

a sinusoidal axisymmetric shape, the phenomenological interaction model and the approximated definition of the asperity width and height. The global performance of the reduced model is then very satisfactory, if one thinks that the computing time is now a few seconds instead of a few days.

As shown in Fig. 15(b), the modelling of the real contact area by the reduced model is better if the interaction is not taken into account, otherwise the contact area is overestimated by about 7%. Nevertheless it remains in a good qualitative and quantitative agreement with the full FEA. The straight line in Fig. 15(b) reveals the slight non-linearity in the change of contact area versus load due to the coalescence of contact zones of neighbouring asperities. If no attention is paid to regions where contact zones of close asperities intersect, i.e. the contact area has been calculated several times in intersecting regions, the global contact area follows the straight line rather accurately. The topology of the contact area obtained with the reduced model is in good agreement with the reference solution. Only one relatively large contact zone has been missed on the RSE (marked with an arrow in Fig. 16). All other contact zones have been determined with a very good precision.

The analysis of the free volume is a more complicated task, because the free volume is mainly determined by valleys and not by asperities. So the main tool to analyse the free volume is a projection of the initial rough surface on a regular mesh and its deformation due to asperities descent and to the interaction between them. In the reduced model, the rough surface is approximated by asperities only where they come in contact with a rigid plane, everywhere else the initial rough surface is projected on the mesh and is deformed due to deduced interaction law. The free volume topologies for load $F = 0.01$ N, $u_z = 0.25$ μm and for the ultimate load $F = 0.5$ N, $u_z = 1$ μm obtained by the FEA are presented in Fig. 17. Also the free volume topology predicted by the reduced model for the ultimate load is presented in this figure. The overall topologies obtained with the reduced model and FEA are in good agreement. Long scars due to machining are explicitly readable. However, in the reduced model, asperities shade somewhere parts of valleys. The global change of the

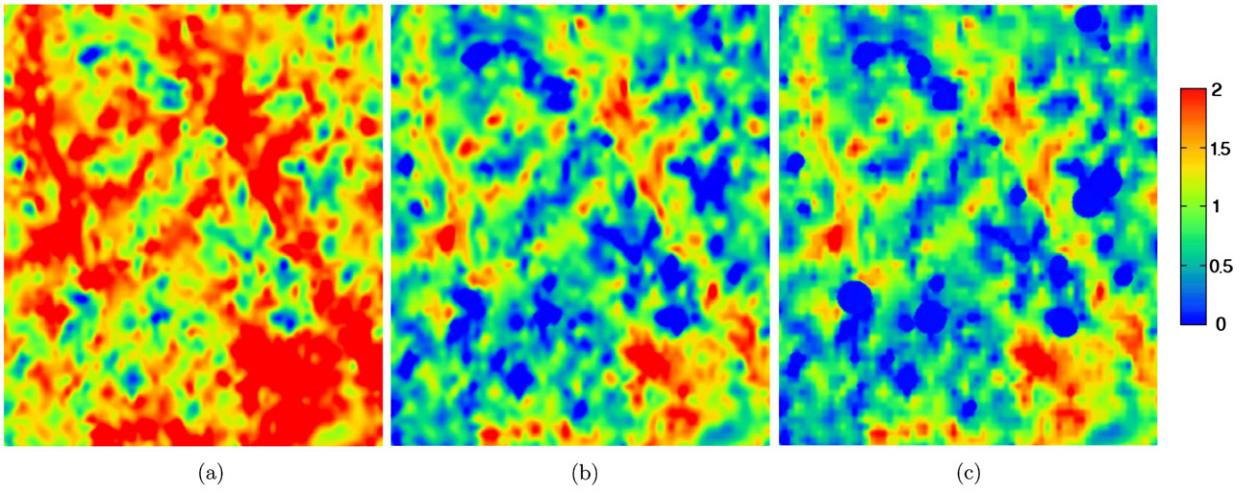

Fig. 17. Comparison of the free volume topology for reduced model with interaction in comparison to the full FEA: (a) FEA for $u_z = 0.25$ µm, $F = 0.011$ N; (b) FEA for the ultimate load $u_z = 1$ µm, $F = 0.5$ N; (c) reduced model for the same ultimate load. Red regions correspond to valleys, blue regions to contact zones.

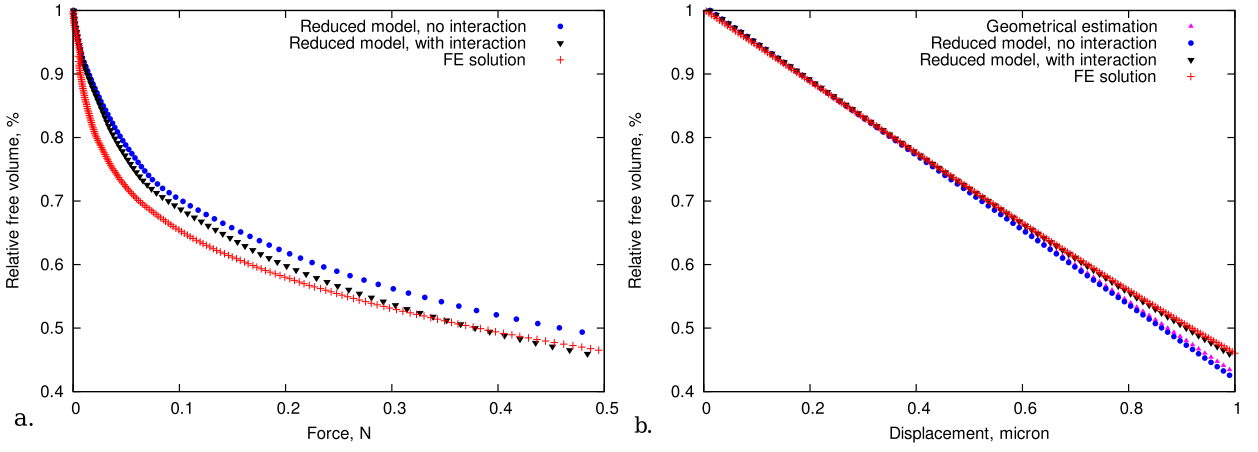

Fig. 18. Full FEA in comparison to reduced model: (a) free volume evolution versus load; (b) free volume evolution versus displacement.

free volume with increasing load is presented in Fig. 18(a) both for finite element analysis and for the reduced model with and without interaction. As expected, it has nothing to do with the evolution of the free volume in case of a single asperity. Qualitatively, two stages can be distinguished: a fast decrease of the free volume for the small load levels, followed, after a smooth transition, by an almost linear evolution, with a much lower slope. The ultimate force of 0.5 N results in a volume reduction of about 54%. The reduced model without interaction slightly overestimates the free volume variation. In fact, this response corresponds to the simple result that would be obtained by means of a pure geometrical analysis, i.e. by cutting the rough surface at the level of the descending rigid plane and integrating the free volume under this plane (Fig. 18(b)). The model with interaction provides a better prediction (less than 2% error with respect to FEA at the ultimate load), demonstrating the interest of this specific study. However for small loads the difference is more pronounced due to a less precise prediction of the force–displacement curve (Fig. 15).

7. Conclusions

Several propositions are made in this article to improve our knowledge concerning the contact of a rough surface and a rigid plane. A representative surface element (RSE) is built by means of an enriching interpolation scheme. This allows us to have in hand a geometrical model, that can be exploited for two purposes. The first one is a full scale finite element analysis (performed on a parallel computer), with more than 100 000 nodes in the contacting surface, that provides a reference solution. The second one is the development of a reduced model which provides a solution to our rough contact problem in a few seconds in very good agreement with full FEA needing a few days. The ingredients of this new reduced model are:

– a series of basic curves obtained by means of elementary finite element computations on a single asperity;
– phenomenological relations to take into account the interaction between neighbouring asperities of the rough surface.

The proposed reduced model predicts very accurately the global responses for the contact between a rough surface and a rigid plane (force/displacement, contact area/force, free volume/force). It provides also the exact contact topologies. Up to now, the agreement has been tested on one RSE only. The work must then be continued, in order to further investigate the assumption of additivity of displacements (those coming from asperity deformation and those coming from the surroundings) that is used in the reduced model.

In the present state, the reduced model is a useful tool for studying very large surfaces that cannot be computed by means of finite element codes, and doing statistics. It should allow to perform quick studies of various roughness scales, and could then be used to check the Greenwood–Williamson model. Only a few roughness scales are classically taken into account due to FE limitations and measurement, but this model will allow to introduce more complex shapes.

The reduced model can be used for the analysis of a soft material in contact with a rigid one. The extension to a real metal/metal contact depends on the validity of Johnson's hypothesis for elastoplastic materials in the framework of large deformations. The next step of this type of study will then be to switch from a rough/smooth contact to a rough/rough contact type. In this case, tangential forces and friction will be of greater importance and should be taken into account. The well-known problem related to geometry accidents (hooks) has also to be considered.

The finite element analysis also demonstrated that even relatively moderate normal load (150 MPa) leads to extremely high local deformations: in the finite element analysis of the considered RSE, accumulated plastic deformation reaches 200%. Normal metal materials are not supposed to bear such high deformations. Consequently a damage or/and a fracture model has to be incorporated in the analysis to get more physical results.

Finally, a simplified contact algorithm (see Appendix A) has been proposed, which allowed to solve fast and efficiently all 3D contact problems presented in the article.

Acknowledgements

This work is supported by the "Chaire EDF-GDF-Suez" (J. Durand). One of the authors (V. Yastrebov) is supported by a CNRS-SNECMA grant #47900 to develop parallel contact computations.

Appendix A. Simplified contact algorithm

The hypothesis once made by Johnson is frequently used to study contact between rough surfaces: the contact between two solids with rough surfaces can be replaced without loss of generality by a contact of a rigid plane and one deformable solid with effective elastic properties and a roughness profile following the undeformed gap between both surfaces [27]. However this hypothesis seems to be questionable in presence of plasticity, friction or large deformations, so in the present paper material properties and surface roughness remain unchangeable. Consequently, we confine ourself to a contact analysis between a metallic surface (both in case of one asperity and full RSE contact) and a rigid plane, which in reality can be presented by a cleavage plane of diamond or sapphire having 10 and 9 Mohs scale hardnesses respectively.

There are several methods which are widely used in the FEA to treat contact problems: penalty method, Lagrange multiplier method [28], augmented Lagrangian method [29]. A method associated with a discretisation technique (for instance node-to-segment [28], mortar method [30], contact domain method [31], etc.) provides solution for a contact problem (see for details [32]). However all of them have drawbacks. The penalty method gives just approximated solution and allows finite penetration between contacting solids, penetration is proportional to the contact pressure, coefficient of proportionality is inversed penalty parameter. The greater the penalty, the smaller penetration but at the same time the higher condition number of the finite element stiffness matrix and the higher probability of divergence. The Lagrange multiplier method is more precise but on the other hand it increases the number of unknowns, which leads to significantly longer resolution time, especially in the considered case where the number of nodes in contact is very high. There is a method which allows to avoid both drawbacks — the augmented Lagrangian method with Uzawa algorithm, but in general this method takes a lot of iterations to converge. However in case of contact occurring between a deformable solid and a rigid plane, a much more simple and reliable method can be used.

Classically in contact mechanics, non-penetration/non-adhesion condition also called Signorini's or Karush–Kuhn–Tucker condition can be written as

$$\sigma_n g_n = 0, \quad \sigma_n \leqslant 0, \; g_n \geqslant 0 \tag{6}$$

where $\sigma_n = \mathbf{n} \cdot \boldsymbol{\sigma} \cdot \mathbf{n}$ — normal component of stress vector on the contact surface, g_n — normal gap between solids, $\boldsymbol{\sigma}$ — stress tensor, $\mathbf{n}$ — normal to the contact surface. This set of conditions means that if there is a contact $g_n = 0$ then contact pressure is not positive, otherwise if contact pressure is zero $\sigma_n = 0$ then normal gap is not negative, there is no contact. A simplified method used for numerical analysis consists in replacement of Signorini's contact conditions by gradually imposed Dirichlet boundary conditions on the contact surface. "Gradually" means here that boundary conditions are updated during iterations and time steps.

To demonstrate the idea let us consider an example. There is a contact surface Γ_c, all nodes of this surface can come in contact with a plane $z(x, y) = z_0$. The contact condition is the following: $\forall i \in 1, N^n : x_z^i \leqslant z_0$, where i is a node number, N^n – number of nodes at considering surface Γ_c, x_z^i – z component of the i-th node vector in the current frame, i.e. $\mathrm{x}_z^i = \mathrm{X}_z^i + u_z^i$, $\mathbf{x}^i$ is the vector of position in the reference frame and $\mathbf{u^i}$ is the displacement vector of the i-th node. The main idea is the following: if node i penetrates the surface z_0, then the z position of the node is imposed to be $\mathrm{x}_z^i = z_0$. In such a way the condition of non-penetration is fulfilled

$$g_n^i = z(x, y) - \mathrm{x}_z^i = z_0 - \mathrm{x}_z^i \geqslant 0 \tag{7}$$

Additionally, for each iteration, it is necessary to verify the non-adhesion condition $R_z^i \leqslant 0$: the z reaction of each node for which Dirichlet "contact" boundary conditions have been imposed should be opposite to the surface (in the current case negative). If this is not the case, then the imposed boundary conditions at considered node have to be removed. Such verification is repeated at each iteration until convergence.

Remark 2. This method is applicable for linear finite elements for frictionless contact with a rigid surface. Rigid surface can be a plane, a cylinder or a sphere in Cartesian, polar or spherical coordinate systems respectively. However if one make use of so-called Multi-Point Constraints the method can be applied for contact with arbitrarily curved surfaces $f(x, y, z) = 0$. If Neumann boundary conditions are imposed additionally to Dirichlet boundary conditions in a special manner then frictional forces can be taken into account as well.

Contrary to the other methods, the present technique does not require any change of the stiffness matrix, even if frictional contact is considered, i.e. stiffness matrix remains symmetric (if material model provide a symmetric matrix) and therefore any solver can be used. Since one of contacting solids is a rigid plane, there is no need in contact detection. All aforementioned things also leads to a very simple parallelisation of contact problems, i.e. if any parallel method is implemented in the Finite Element code, then no additional efforts should be overtaken to perform parallel FEA of problems including contact. This method complemented with a specific post-processing for displacement, contact force/pressure, contact area and free volume analysis has been successfully applied for all sequential and parallel 3D FE simulations in the current study.

References

[1] J.A. Greenwood, J.B.P. Williamson, Contact of nominally flat surfaces, Proceedings of the Royal Society of London Series A 295 (1966) 300–319.
[2] A.W. Bush, R.D. Gibson, T.R. Thomas, The elastic contact of a rough surface, Wear 153 (1992) 53–64.
[3] Y. Zhao, L. Chang, A model of asperity interactions in elastic–plastic contact of rough surfaces, Journal of Tribology 123 (2001) 857–864.
[4] L. Pei, S. Hyun, J.F. Molinari, M.O. Robbins, Finite element modeling of elasto–plastic contact between rough surfaces, Journal of the Mechanics and Physics of Solids 53 (2005) 2385–2409.
[5] E.J. Abbott, F.A. Firestone, Specifying surface quality – a method based on accurate measurement and comparison, Mechanical Engineering 55 (1933) 569.
[6] B. Bhushan, Surface roughness analysis and measurement techniques, in: Modern Tribology Handbook, vol. 1, CRC Press, London, 2001, pp. 49–119.
[7] D.J. Whitehouse, J.F. Archard, The properties of random surface of significance in their contact, Proceedings of the Royal Society of London Series A 316 (1970) 97–121.
[8] A.W. Bush, R.D. Gibson, G.P. Keogh, Strongly anisotropic rough surfaces, Journal of Lubrification Technology 101 (1979) 15–20.
[9] A.W. Bush, R.D. Gibson, The elastic contact of a rough surface, Wear 35 (1987) 87–111.
[10] W.R. Chang, I. Etsion, D.B. Bogy, An elastic–plastic model for the contact of rough surfaces, Journal of Tribology 109 (1987) 257–263.
[11] Y. Zhao, D.M. Maietta, L. Chang, An asperity microcontact model incorporating the transition from elastic deformations to fully plastic flow, Journal of Tribology 122 (2000) 86–93.
[12] C. Vallet, D. Lasseux, P. Sainsot, H. Zahouani, Real versus synthesized fractal surfaces: Contact mechanics and transport properties, Tribology International 42 (2009) 250–259.
[13] C. Vallet, D. Lasseux, H. Zahouani, P. Sainsot, Sampling effect on contact and transport properties between fractal surfaces, Tribology International 42 (2009) 1132–1145.
[14] P. Sainsot, N. Jacq, D. Nélias, A numerical model for elastoplastic rough contact, CMES 3 (4) (2002) 497–506.
[15] B.B. Mandelbrot, The Fractal Geometry of Nature, Freeman, New York, 1982.
[16] B. Bhushan, Contact mechanics of rough surfaces in tribology: Multiple asperity contact, Tribology Letters 4 (1997) 1–35.
[17] B.N.J. Persson, Theory of rubber friction and contact mechanics, Journal of Chemical Physics 115 (2001) 3840–3861.
[18] A. Majumdar, B. Bhushan, Fractal model of elastic–plastic contact between rough surface, Journal of Tribology 113 (1991) 1–11.
[19] A. Majumdar, B. Bhushan, Elastic–plastic contact model of bifractal surfaces, Wear 35 (1975) 87–111.
[20] C. Vallet, Fuite liquide au travers d'un contact rugueux : application l'étanchéité interne d'appareils de robinetterie. PhD thesis, Ecole Nationale Supérieure d'Arts et Métiers, 2008.
[21] C. Farhat, F.-X. Roux, Implicit parallel processing in structural mechanics, Computational Mechanics Advances 2 (1994) 1–24.
[22] B.N.J. Persson, Contact mechanics for randomly rough surfaces, Surface Science Reports 61 (2006) 201–227.
[23] G.E. Farin, Curves and Surfaces for CAGD: A Practical Guide, 5th edition, Morgan Kaufmann, 2002.
[24] T. Dick, G. Cailletaud, Fretting modelling with a crystal plasticity model of Ti6Al4V, Computational Materials Science 38 (2006) 113–125.
[25] B.N.J. Persson, O. Albohr, U. Tartaglino, A.I. Volokitin, E. Tosatti, On the nature of surface roughness with application to contact mechanics, sealing, rubber friction and adhesion, Journal of Physics: Condensed Matter 17 (2005) R1–R62.
[26] P. Raghavan, DSCPACK: Domain-separator codes for the parallel solution of sparse linear systems, Tech. rep. CSE-02-004, Department of Computer Science and Engineering, The Pennsylvania State University, University Park, PA 16802, 2002.
[27] K.L. Johnson, Contact Mechanics, Cambridge University Press, 1987.
[28] N. Kikuchi, J.T. Oden, Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods, SIAM, Philadelphia, 1988.

[29] P. Alart, A. Curnier, A mixed formulation for frictional contact problems prone to newton like solution method, Computer Methods in Applied Mechanics and Engineering 92 (1991) 353–375.
[30] T.W. McDewitt, T.A. Laursen, A mortar-finite element formulation for frictional contact problems, International Journal for Numerical Methods in Engineering 48 (2000) 1525–1547.
[31] J. Oliver, S. Hartmann, J.C. Cante, R. Weyler, J.A. Hernández, A contact domain method for large deformation frictional contact problems. Part I: Theoretical basis, Computer Methods in Applied Mechanics and Engineering 198 (2009) 2591–2606.
[32] P. Wriggers, Computational Contact Mechanics, 2nd edition, Springer, 2006.