

Chapter 4

Computational Methods for Contact Problems with Roughness



Marco Paggi, Alberto Bemporad and José Reinoso

Abstract This chapter provides a self-consistent introduction to computational methods for the solution of contact problems between bodies separated by rough interfaces. Both frictional and frictionless contact problems are examined. The mathematical formulation of the boundary element method is presented first, with details on the possible algorithmic implementation strategies and their computational efficiency. In the second part of the chapter, the fundamentals of the finite element method for the solution of contact problems are presented, along with an overview on the different strategies available in the literature to accurately discretize the multiscale features of roughness. A synopsis of the major advantages and disadvantages provided by the computational methods based on the boundary element method or the finite element method concludes the chapter, illustrating also perspective research directions.

Introduction

Contact mechanics between rough surfaces is a very active area of theoretical and applied research in physics and engineering (Vakis et al. 2018). Due to roughness, when two bodies separated by rough boundaries are brought into contact, they exchange forces through the so-called contact spots, which correspond to the tips of the asperities, i.e., the local maxima of the surfaces. As a consequence, the real area of contact is usually a small percentage of the nominal one, which would be attained only if the surfaces were perfectly flat. The evolution of the contact domain, which includes all the contact spots, the size of the real area of contact, and the normal contact stiffness, do depend on the applied normal load level, see Borri-Brunetto et al.

M. Paggi (✉) · A. Bemporad
IMT School for Advanced Studies Lucca, Piazza San Francesco 19, 55100 Lucca, Italy
e-mail: marco.paggi@imtlucca.it

J. Reinoso
School of Engineering, University of Seville, Camino de los Descubrimientos s/n,
41092 Seville, Spain

© CISM International Centre for Mechanical Sciences 2020
M. Paggi and D. Hills (eds.), *Modeling and Simulation of Tribological Problems in Technology*, CISM International Centre for Mechanical Sciences 593,
https://doi.org/10.1007/978-3-030-20377-1_4

(1999), Ciavarella et al. (2000, 2004, 2008a), Campaña et al. (2001), Barber (2003), Nosonovsky and Bhushan (2005), Persson (2006), Hyun and Robbins (2007), Carbone and Bottiglione (2008), Paggi and Ciavarella (2010), Paggi and Barber (2011), Paggi et al. (2014), and Yastrebov et al. (2015) for a selection of studies. Similarly, when a shearing load is applied, there is a progressive transition of the contact spots from a full stick condition, with perfect adhesion and no relative displacement between the bodies, to full slip, when sliding takes place. Such a transition is ruled by the Coulomb friction law at the asperity level, while the emerging quantities, such as for instance the total shearing load versus the size of the contact area in stick or slip conditions, are the result of a collective response emerging from the complex local interactions (Carpinteri and Paggi 2005, 2009; Paggi et al. 2014).

In this context, semi-analytical micromechanical contact theories relying on the statistical distribution of the elevation of the asperities and their radii of curvature have been proposed and widely explored in the engineering community (see McCool 1986; Zavarise et al. 2004a for comprehensive review articles), following the pioneering approach by Greenwood and Williamson (1966) and extending it to more complex statistical distributions of elevations and curvatures (Ciavarella et al. 2006; Greenwood 2006; Paggi and Ciavarella 2010), considering also elastic interactions between asperities (Ciavarella et al. 2008b) that were not included in the original pioneering formulations. Since the 1990s, research focused on the multiscale features of roughness, exploiting the use of fractal geometry for the understanding of its role on the contact behavior (Majumdar and Bhushan 1990; Borri-Brunetto et al. 1999; Carpinteri and Paggi 2005; Persson et al. 2005).

More recently, it has been found that neither the random process theory, which is the theoretical framework for the derivation of micromechanical contact theories, nor the fractal description of roughness is able to reproduce the complex morphology of surfaces (Greenwood and Wu 2001), as recently proved for natural or engineering surfaces with functionalized textures (Borri and Paggi 2015, 2016). Therefore, the predictions of semi-analytical contact models based on random process theory or fractal assumptions should be checked with care and led to a wide range of comparisons and validation studies (Mueser et al. 2017). On the other hand, experimental investigations are challenging to be performed and involve approximations too (Woo and Thomas 1980). For example, very often the contact quantities can only be estimated by indirect measurements of thermal or electric resistances of compressed rough joints (Sridhar and Yovanovich 1994), or they are mostly concerned with the measurement of the real area of contact under special conditions allowing for its inspection (O'Callaghan and Probert 1970; Hendriks and Visscher 1995).

Therefore, due to the general considerations above, numerical methods able to deal with realistic surface topologies without making approximations and assumptions on their shape, and with any constitutive response of the continuum and of the interface, are very important to predict the contact response and infer general conclusions on the observed trends.

In the linear elastic regime, if the multiscale character of roughness covering a wide range of wavelengths is one of the most prominent research topics, then the use of the boundary element method (BEM) has been historically preferred over the finite

element method (FEM) (Andersson 1981; Man 1994). This is essentially due to the fact that only the surface must be discretized in the boundary element method, and not the surrounding continuum, as required by the finite element method. Moreover, it is not necessary to adopt surface interpolation techniques, like Bezier curves, to discretize the interface (see, e.g., the approach in Wriggers and Reinelt (2009)) and make it amenable for the application of contact search algorithms. This avoids an undesired smoothing of the fine-scale geometrical features of roughness.

In the application of the boundary element method, the core of the procedure is based on the knowledge of the so-called Green functions that relate the displacement of a generic point of the half-plane to the action of a concentrated force on the surface caused by contact interactions. An integral convolution of the effects of all the contact tractions provides the deformed contact configuration. After introducing a discretization of the half-plane consisting of a grid of boundary elements, the problem of point-force singularity is solved numerically by using the closed-form solution for a patch load acting on a finite-size boundary element (Johnson 1985, Chaps. 3, 4). The contact problem is then set in terms of equalities and inequalities stemming from the unilateral contact constraints and it can be solved by constrained optimization algorithms, see Polonsky and Keer (1999), Bemporad and Paggi (2015). The basic version of the boundary element method can be also extended to solve contact problems with friction (Li and Berger 2003; Pohrt and Li 2014) and between viscoelastic materials, see Carbone and Putignano (2013) and the references therein given.

However, standard boundary element formulations are based on the fundamental assumptions of linear elasticity and homogeneity of the materials, and their extension to inhomogeneities (Leroux et al. 2010), finite-size geometries (Putignano et al. 2015), or interface constitutive nonlinearities such as adhesion (Rey et al. 2017; Popov et al. 2017; Li et al. 2018) are sometimes possible but are not so straightforward. For these problems, the finite element method would be conceived as a more versatile computational approach to pursuit in order to overcome all the major limitations of the boundary element method. The finite element method can in fact take into account any material or interface constitutive nonlinearity, and it can easily treat finite-size geometries of practical interest in industrial applications. Moreover, it is prone to be extended for the solution of nonlinear multi-field problems involved in heat transfer or in reaction–diffusion systems (Zavarise et al. 2019; Sapora and Paggi 2014; Lenarda et al. 2018), for which the boundary element method has not been applied so far.

In spite of the different appealing aspects of the finite element method over the boundary element method, this approach has been limited to few remarkable studies concerning contact problems with roughness, especially in relation to elastoplasticity (Pei et al. 2005; Hyun et al. 2004). The motivation is primarily due to the need for discretizing the bulk and also the rough interface, which is not an easy task from the mathematical standpoint and it also gives a rise in computation costs. As shown in Wriggers and Reinelt (2009), Bezier interpolation techniques can be employed to regularize rough interfaces to become amenable for contact search algorithms. Nevertheless, smoothing should be applied with care in order to avoid artificial

filtering of finer surface features relevant to the physics of contact. To overcome such limitations, a recent approach which does not explicitly discretize roughness, but it embeds its analytical form in a nominally flat interface finite element has proved to be very efficient in solving contact problems with roughness in the finite element method (Paggi and Reinoso 2018), significantly simplifying the issue of roughness discretization.

In this chapter, an overview of computational methods for solving contact problems with roughness is proposed. Section “[The Boundary Element Method](#)” focuses on the fundamentals of the boundary element method, first in relation to the frictionless normal contact problem. Based on the results in Bemporad and Paggi (2015), special attention is given to the review of the computational challenges of the method, which regard two main aspects: (i) efficiently solve the system of linear equations; (ii) impose the satisfaction of the unilateral contact constraints (contact inequalities). Regarding the first issue, iterative methods like the conjugate gradient algorithm or the Gauss–Seidel method (Francis 1983; Borri-Brunetto et al. 1999, 2001), or the capabilities of multigrid or multilevel methods (Raous 1999; Polonsky and Keer 1999), or even the solution of the linear system of equations in the Fourier space (Nogi and Kato 1997; Polonsky and Keer 2000a, b; Batrouni et al. 2002; Scaraggi et al. 2013; Prodanov et al. 2014; Vollebregt 2014) are possible strategies.

Regarding the imposition of the contact inequalities, on the other hand, a greedy approach where the boundary elements bearing tensile loads are iteratively excluded can be exploited (Kubo et al. 1981; Borri-Brunetto et al. 1999, 2001; Karpenko and Akay 2001; Batrouni et al. 2002), although it has been demonstrated by Bemporad and Paggi (2015) that it often fails for very compact contact domains. The constrained conjugate gradient method proposed by Polonsky and Keer (1999) and based on the theory in Hestenes (1980, Chaps. 2, 3) to solve the linear system of equations and rigorously impose the satisfaction of the contact constraints is also discussed, along with its developments (Polonsky and Keer 2000a, b). Finally, other optimization algorithms based on the solution of the corresponding quadratic program, such as nonnegative least squares (NNLS) and the alternative direction method of multipliers (ADMM), proposed in Bemporad and Paggi (2015), are detailed. A careful comparison of the available methods in terms of computation cost is also provided, along with other more advanced acceleration strategies.

The presentation moves then to the contact problem with friction, which is significantly complicated by the coupling between the normal and the tangential deformations, and by the Coulomb friction law to be locally satisfied in the tangential direction.

Section “[The Finite Element Method](#)” presents the basis of the finite element method, starting from the variational formulation of the problem, including the strong form, the governing equations for the continuum and the interface, and also the weak form. Next, finite element procedures based on the explicit discretization of roughness are briefly summarized in relation to seminal work published in the literature. Finally, the zero-thickness interface finite element with embedded profile for joint roughness (MPJR interface finite element, recently proposed in Paggi and Reinoso (2018)) is presented, along with its implementation details in the research finite element

analysis program FEAP (Zienkiewicz and Taylor 2000). A comparison between the different finite element discretization strategies for the solution of a benchmark Hertzian contact problem concludes the section and it shows the great potential of the MPJR interface finite element for future research.

The last section provides an overall summary of the chapter with a synopsis reporting the major advantages and disadvantages of the computational methods herein reviewed, along with an overview of perspective research directions in the field of contact mechanics between rough surfaces.

The Boundary Element Method

The boundary element method (BEM) is an efficient technique to solve the contact problem between two linear elastic bodies, say Ω_1 and Ω_2 , with rough boundaries Γ_1 and Γ_2 , respectively. The first step for the application of the method is the knowledge of the topographies Γ_i ($i = 1, 2$), which are nowadays acquired using confocal profilometers or atomic force microscopes. They are stored in a matrix containing the x_i , y_i , and z_i coordinates with respect to an arbitrary datum. Such surfaces are nonconforming, i.e., they do not match when in contact. Hence, the shape of the contact area changes with the applied load and the contact area is an unknown of the problem, which is a source of complexity for the solution procedure.

Due to linear elasticity of the continuum, the actual contact problem can be simplified by reducing it to the solution of a fictitious contact problem between a rigid microscopically rough surface with a composite topography and a flat linear elastic half-space with composite elastic moduli E and G dependent on the Young's moduli E_i and the Poisson's ratios ν_i of the two materials ($i = 1, 2$). The composite topography is simply obtained by summing up the elevations of the two rough surfaces Γ_i ($i = 1, 2$). The composite elastic moduli, on the other hand, are given by the following formulae:

$$E = \left(\frac{1 - \nu_1^2}{E_1} + \frac{1 - \nu_2^2}{E_2} \right)^{-1}, \quad (4.1a)$$

$$G = \left(\frac{2 - \nu_1}{4G_1} + \frac{2 - \nu_2}{4G_2} \right)^{-1}, \quad (4.1b)$$

where $G_i = E_i/[2(1 + \nu_i)]$. The composite Poisson ratio ν is related to E and G via $\nu = E/(2G) - 1$.

In this simplified setting, the next step for the application of the method is the knowledge of the so-called Green functions, which relate the displacements of any point belonging to the deformable half-space to the applied tractions on its surface. This allows formulating the problem by involving only contact tractions and surface displacements, getting rid of the surrounding continua. While the simplest expressions for the Green functions are those for a cylindrical punch on a homogenous

and isotropic half-space, other forms for more complex material configurations can be of interest in engineering applications. For instance, multilayered half-spaces, deformable elastic bodies with finite size (Putignano et al. 2015), or half-spaces with spatial inclusions or voids (Leroux et al. 2010) have been investigated in the literature. Quite recently, adhesion effects in tension have been included in BEM for frictionless normal contact problems, see Rey et al. (2017), Popov et al. (2017), Li et al. (2018).

For more complex heterogeneous material compositions due to the random presence of inclusions or voids of arbitrary shape, or functionally graded compositions, the Green functions can be solely determined in numerical form (Paggi and Zavarise 2011). For that, the finite element method (FEM) can be used to preliminarily extract the Green functions. This is done by discretizing the bulk of the deformable body and its internal microstructure. Then, unit tractions are applied at any point over the half-plane boundary, and the induced surface displacements are computed. Therefore, point-by-point, the Green functions can be numerically reconstructed. In the sequel, we shall restrict the attention to linear elastic homogeneous and isotropic half-spaces, for the sake of simplicity. More complex material compositions can be dealt with by modifying the expression for the Green functions.

It is important to remark here that the knowledge of the relation between tractions and surface displacements is a key point for the simplification of the complexity of the contact problem, since it allows avoiding the discretization of the bulk. This is indeed a significant gain over the finite element method in terms of computation costs, especially for problems with rough boundaries where the discretization of the rough interface has to be very fine to capture its essential multiscale features. On the other hand, BEM is rigorously exact only for linear elastic contact problems, since the principle of superposition is applied to convolute the effect of any distribution of surface tractions applied over the half-plane. For elastoplastic contact problems with isotropic or kinematic hardening laws, Chen et al. (2008) generalized the BEM formulation by considering modified discrete expressions for the Green functions and the residual displacements caused by plastic deformation. For other relevant publications on this topic, the reader is referred to Kogut and Etsion (2002), Chang et al. (1987), Nelias et al. (2006). For other problems involving material or geometrical (finite elasticity) nonlinearities of the continuum, or for multi-field problems, the finite element method is indeed preferable over the boundary element method and it would be the ideal framework for their investigation.

The Frictionless Normal Contact Problem

Back to the classical framework of BEM for frictionless normal contact problems in linear elasticity, the normal displacements $u(\mathbf{x})$ at any point of the half-plane identified by the position vector \mathbf{x} are related to the contact tractions $p(\mathbf{y})$ at other points as follows (Johnson 1985; Barber 2018):

$$u(\mathbf{x}) = \int_S H(\mathbf{x}, \mathbf{y}) p(\mathbf{y}) d\mathbf{y}, \tag{4.2}$$

where $H(\mathbf{x}, \mathbf{y})$ represents the displacement at a point \mathbf{x} due to a surface contact pressure p acting at \mathbf{y} , and S is the elastic half-plane. For homogeneous, isotropic, linear elastic materials, the influence coefficients (Green function) are given by Boussinesq:

$$H(\mathbf{x}, \mathbf{y}) = \frac{1 - \nu^2}{\pi E} \frac{1}{\|\mathbf{x} - \mathbf{y}\|}, \tag{4.3}$$

where $\|\cdot\|$ denotes the standard Euclidean norm. The total contact force P is the integral of the contact pressure field:

$$P = \int_S p(\mathbf{x}) dA. \tag{4.4}$$

By referring to Fig. 4.1, in the following we define for each surface point $\mathbf{x} \in S$ its elevation $\xi(\mathbf{x})$, measured with respect to a reference frame, and set $\xi_{\max} \triangleq \max_{\mathbf{x} \in S} \xi(\mathbf{x})$ the maximum elevation. The indentation of the half-plane at the points in contact is denoted by \bar{u} , whereas a generic displacement along the surface is u .

For a given far-field displacement $\Delta \geq 0$ in the direction perpendicular to the undeformed half-plane, the problem is to find the solution of the normal contact $u(\mathbf{x})$, $p(\mathbf{x})$ satisfying (4.2) and the unilateral contact (linear complementarity) conditions

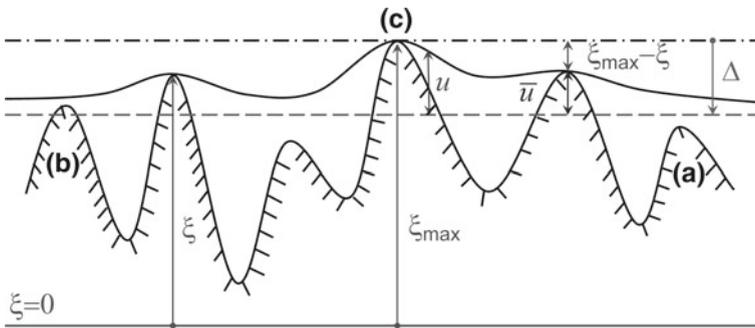


Fig. 4.1 Sketch of the contact problem between a rigid rough surface and an elastic half-plane. Its deformed configuration corresponding to the imposed far-field displacement Δ is depicted with a solid line. The dashed line corresponding to a rigid body motion of the half-plane identifies the heights to be included in the initial trial contact domain. Once the contact problem is solved, one may have (i) heights certainly not in contact from the beginning, type (a); (ii) heights losing contact due to elastic interactions, type (b); (iii) heights in contact, type (c). Adapted from Bemporad and Paggi (2015)

$$u(\mathbf{x}) - \bar{u}(\mathbf{x}, \Delta) \geq 0, \quad (4.5a)$$

$$p(\mathbf{x}) \geq 0, \quad (4.5b)$$

$$(u(\mathbf{x}) - \bar{u}(\mathbf{x}, \Delta))p(\mathbf{x}) = 0, \quad (4.5c)$$

for all points $\mathbf{x} \in S$, where contact tractions are considered positive when compressive.

Introducing the quantity $w(\mathbf{x}, \Delta) = u(\mathbf{x}) - \bar{u}(\mathbf{x}, \Delta)$, Eq. (4.5) can be rewritten as

$$w(\mathbf{x}, \Delta) \geq 0, \quad (4.6a)$$

$$p(\mathbf{x}) \geq 0, \quad (4.6b)$$

$$w(\mathbf{x}, \Delta)p(\mathbf{x}) = 0. \quad (4.6c)$$

The above contact problem is an infinite-dimensional linear complementarity problem. A finite-dimensional approximate solution can be sought by discretizing the surface as a square grid of spacing δ consisting of $N \times N$ average heights. Let S_{ij} be the cell of area δ^2 indexed by $i, j \in I_N$, with $I_N \triangleq \{1, \dots, N\} \times \{1, \dots, N\}$. Let $\mathbf{x}_{i,j} \triangleq \frac{1}{S_{ij}} \int_{\mathbf{x} \in S_{ij}} \mathbf{x} dA$, $\xi_{i,j} \triangleq \frac{1}{S_{ij}} \int_{\mathbf{x} \in S_{ij}} \xi(\mathbf{x}) dA$, $p_{i,j} \triangleq \int_{\mathbf{x} \in S_{ij}} p(\mathbf{x}) dA$, and $u_{i,j} \triangleq \frac{1}{S_{ij}} \int_{\mathbf{x} \in S_{ij}} u(\mathbf{x}) dA$ be, respectively, the barycentric coordinate, average height, resultant of the contact tractions, and the corresponding average displacement on the surface element S_{ij} . Consider the following discretized version of (4.2)

$$u_{i,j} = \sum_{k=1}^N \sum_{l=1}^N H_{i-k,j-l} p_{k,l} \quad (4.7)$$

for all $(i, j), (k, l) \in I_N$, $p_{k,l} \geq 0$, where the term $H_{i-k,j-l}$ is the Green function in (4.3) averaged over the elementary area δ^2 . For instance, Borri-Brunetto et al. (1999) used the following approximation related to a uniform pressure acting on a rounded punch of radius $\delta/2$:

$$H_{i-k,j-l} = \begin{cases} \frac{2}{E\pi\delta}, & \text{if } i = k \text{ and } j = l \\ \frac{\delta}{E\pi\delta} \arcsin \frac{\delta}{2\|\mathbf{x}_{i,j} - \mathbf{x}_{k,l}\|}, & \text{if } i \neq k, j \neq l \end{cases} \quad (4.8)$$

but other formulae for a square punch can also be taken as in Pohrt and Li (2014).

Let $\bar{I}_C \triangleq \{(i, j) \in I_N : \xi_{i,j} < \xi_{\max} - \Delta\}$ be the set of indices corresponding to elements S_{ij} that are certainly not in contact (cf. Fig. 4.1), and hence

$$p_{k,l} = 0, \forall (k, l) \in \bar{I}_C. \quad (4.9)$$

Let $m = \#\bar{I}_C$ be the number of elements of \bar{I}_C and $n = \#I_C$ the number of elements belonging to the initial trial contact domain, $I_C \triangleq I_N \setminus \bar{I}_C$. The set I_C is only a superset of the set I_C^* of actual contact points, since the corrections to the displacements induced by elastic interactions may induce lack of contact in some elements (i, j) , i.e., $u_{i,j} > \bar{u}_{i,j}$, where $\bar{u}_{i,j} \triangleq \Delta - \xi_{\max} + \xi_{i,j}$ is the value of the compenetration of the height corresponding to the element (i, j) in the half-plane (see Fig. 4.1).

For a generic $(i, j) \in I_C$ corresponding to an element of the surface which is potentially in contact with the elastic half-plane, we denote by

$$w_{i,j} \triangleq u_{i,j} - \bar{u}_{i,j} \geq 0 \quad (4.10)$$

the corresponding elastic correction to the displacement. Clearly, it must hold that

$$w_{i,j} p_{i,j} = 0, \quad \forall (i, j) \in I_C \quad (4.11)$$

since $w_{i,j} > 0$ implies no contact between the surfaces and therefore vanishing contact pressures, while $p_{i,j} > 0$ implies contact, $u_{i,j} = \bar{u}_{i,j}$, or equivalently $w_{i,j} = 0$.

By taking into account that $p_{k,l} = 0$ for all $(k, l) \in \bar{I}_C$, Eq. (4.7) can be recast as the following condition:

$$w_{i,j} + \bar{u}_{i,j} = \sum_{(k,l) \in I_C} H_{i-k,j-l} p_{k,l}, \quad \forall (i, j) \in I_C, \quad (4.12)$$

which is limited to the nodes belonging to the initial trial contact domain I_C , whose number of elements is in general significantly smaller than those of I_N . The relations (4.9)–(4.12) can be recast in matrix form as the following Linear Complementarity Problem (LCP) (Cottle et al. 1992):

$$\mathbf{w} = \mathbf{H}\mathbf{p} - \bar{\mathbf{u}} \quad (4.13a)$$

$$\mathbf{w} \geq \mathbf{0}, \quad \mathbf{p} \geq \mathbf{0}, \quad \mathbf{w}^T \mathbf{p} = 0, \quad (4.13b)$$

where $\mathbf{w} \in \mathbb{R}^n$ is the vector of unknown elastic corrections $w_{i,j}$, $(i, j) \in \bar{I}_C$, \mathbf{w}^T denotes its transpose, $\mathbf{p} \in \mathbb{R}^n$ is the vector of unknown boundary element contact forces $p_{i,j}$, $(i, j) \in I_C$, $\bar{\mathbf{u}} \in \mathbb{R}^n$ is the vector of compenetrations $\bar{u}_{i,j}$, $(i, j) \in I_C$, and $\mathbf{H} = \mathbf{H}^T$ is the matrix obtained by collecting the compliance coefficients $H_{i-k,j-l}$, for $(i, j), (k, l) \in I_C$. Due to the properties of linear elasticity (Johnson 1985, p. 144), we have that

$$\mathbf{H} = \mathbf{H}^T > 0, \quad (4.14)$$

that is, \mathbf{H} is a symmetric positive definite matrix (with the additional property deriving from (4.8) of having all its entries positive). After solving (4.13), the vector $\mathbf{u} \in \mathbb{R}^n$ of normal displacements $u_{i,j}$, $(i, j) \in I_C$, is simply retrieved as $\mathbf{u} = \bar{\mathbf{u}} + \mathbf{w}$.

By the positive definiteness property (4.14) of \mathbf{H} , we inherit immediately the following important property (Cottle et al. 1992, Theorem 3.3.7): the discretized

version (4.6), (4.9)–(4.12) of the contact problem admits a unique solution \mathbf{p} , \mathbf{u} , for all $\Delta \geq 0$.

The LCP problem (4.13) corresponds to the Karush–Kuhn–Tucker (KKT) conditions for optimality of the following convex quadratic program (QP):

$$\min_p \frac{1}{2} \mathbf{p}^T \mathbf{H} \mathbf{p} - \bar{\mathbf{u}}^T \mathbf{p} \quad (4.15a)$$

$$\text{s.t. } \mathbf{p} \geq \mathbf{0} \quad (4.15b)$$

in that the solution \mathbf{p} of (4.15) and its corresponding optimal dual solution \mathbf{w} solve (4.13), and vice versa.

The QP problem is consistent with former considerations by Kalker and van Randen (1972) and also summarized in Johnson (Johnson 1985, pp. 151–152). In fact, the contact pressures solving the unilateral contact problem can be obtained by minimizing the total complementary energy W of the linear elastic system, subject to the constraint $p(\mathbf{x}) \geq 0$, $\forall \mathbf{x} \in \mathcal{S}$. For a continuous system, the total complementary energy is

$$W = U - \int_{\mathcal{S}} p(\mathbf{x}) \bar{u}(\mathbf{x}, \Delta) \, d\mathbf{x}, \quad (4.16)$$

where U is the internal complementary energy of the deformed half-plane in contact. For linear elastic materials, we have

$$U = \frac{1}{2} \int_{\mathcal{S}} p(\mathbf{x}) u(\mathbf{x}) \, d\mathbf{x}. \quad (4.17)$$

Although such an energy-based approach can be used to derive finite element formulations, it is also possible to remain within the boundary element method and introduce a surface discretization as before. By invoking the averaged Green functions in (4.8), the discretized version of W , say \tilde{W} , reads

$$\tilde{W} = \frac{1}{2} \sum_{(i,j) \in I_C} \sum_{(k,l) \in I_C} H_{i-k,j-l} p_{k,l} p_{i,j} - \sum_{(i,j) \in I_C} p_{i,j} \bar{u}_{i,j}, \quad (4.18)$$

which represents a quadratic function of \mathbf{p} to be minimized, under the constraints $p_{i,j} \geq 0$, $\forall (i, j) \in I_C$, as in (4.15). Since it is unlikely that the contact area is known a priori, the active set of nodes in contact results only after solving problem (4.13) or equivalently (4.15).

A large variety of solvers for LCP and QP problems were developed in the last 60 years (Beale 1955; Fletcher 1971; Goldfarb and Idnani 1983; Cottle et al. 1992; Schmid and Biegler 1994; Patrinos and Bemporad 2014), and is still an active area of research in the optimization and control communities. Historically, in the mechanics community, Kalker and van Randen (1972) proposed the simplex method, although

it was found to be practical only for relatively small N . More recent contributions adopt algorithms to solve the unconstrained linear system of equations and then correct the solution by eliminating the boundary elements bearing tensile tractions (Francis 1983; Borri-Brunetto et al. 1999, 2001), or use a constrained version of the conjugate gradient (CG) algorithm (Polonsky and Keer 1999). These methods are simply initialized by considering arbitrary nonnegative entries in \mathbf{p} , without taking advantage of the monotonic increase (or decrease) of pressures by increasing (or decreasing) the far-field displacement, an important property guaranteed by rigorous elasticity theorems (Barber 1974). The history of pressures can be saved during a contact simulation and it is easy to access and use and it can be beneficial to save computation time, as proved by Bemporad and Paggi (2015).

Optimization algorithms. Since now on, we use the subscript i to denote the i th component of a vector or the i th row of a matrix, the subscript \mathcal{I} to denote the subvector obtained by collecting all the components $i \in \mathcal{I}$ of a vector (or all the rows i of a matrix), and the double subscript $\mathcal{I}, \mathcal{I}_1$ to denote the submatrix obtained by collecting the i th row and j th column, for all $i \in \mathcal{I}, j \in \mathcal{I}_1$.

In the sequel, following the content in Bemporad and Paggi (2015), a brief overview of algorithms to solve the constrained contact problem is provided, starting first with those for the solution of the LCP, namely, the greedy method and the constrained conjugate gradient algorithm. Next, optimization algorithms for the solution of the corresponding QP are discussed, such as the nonnegative least squares and the alternative direction method of multipliers. Finally, a comparison of the algorithms above in terms of computation performance is presented. Further acceleration strategies and advanced methods are also discussed.

A greedy method corresponds to solve problem (4.15) by iteratively solving the unconstrained linear system of equations $\mathbf{w} = \mathbf{H}\mathbf{p} - \bar{\mathbf{u}} = \mathbf{0}$ with respect to \mathbf{p} and increasingly zeroing negative elements of \mathbf{p} until the condition $\mathbf{p} \geq \mathbf{0}$ is satisfied. By construction we obtain $\mathbf{w}^T \mathbf{p} = 0$. The method is described in Algorithm 1, in which a standard conjugate gradient (CG) is employed to solve the unconstrained linear system of equations. Steps 2.1–2.4 can be replaced by any other algorithm for solving the linear system of equations, like the Gauss–Seidel iterative scheme as in Borri-Brunetto et al. (1999, 2001), the MATLAB’s `mldivide` solver, or even the FFT algorithm as in Karpenko and Akay (2001), Batrouni et al. (2002), Vollebregt (2014).

Assuming that the prescribed initial \mathbf{p} and \mathcal{I} are such that $p_j = 0$ for all $j \in \{1, \dots, n\} \setminus \mathcal{I}$, and K_{\max} is sufficiently large, the output of the greedy algorithm leads to a contact pressure vector \mathbf{p}^* and a normal displacement vector \mathbf{u}^* satisfying $\mathbf{u}^* = \mathbf{H}\mathbf{p}^*$, $\mathbf{p}^* \geq \mathbf{0}$, $(\mathbf{u}^* - \bar{\mathbf{u}})^T \mathbf{p}^* = 0$. In fact, condition $\mathbf{p}^* \geq \mathbf{0}$ is guaranteed by the condition in Step 2 up to ϵ precision. By letting $\mathbf{w}^* \triangleq \mathbf{u}^* - \bar{\mathbf{u}}$, at termination of the algorithm we have $\mathbf{w}_{\mathcal{I}}^* = \mathbf{H}_{\mathcal{I}, \mathcal{I}} \mathbf{p}_{\mathcal{I}}^* - \bar{\mathbf{u}}_{\mathcal{I}} = \mathbf{0}$ because of the solution of the CG method (Step 2.4), or equivalently $\mathbf{u}_{\mathcal{I}}^* = \bar{\mathbf{u}}_{\mathcal{I}}$ (cf. Step 4). By setting $\mathbf{u}_{\mathcal{I}}^* \triangleq \mathbf{H}_{\mathcal{I}, \mathcal{I}} \mathbf{p}_{\mathcal{I}}^*$ in Step 4, and recalling that $\mathbf{p}_{\mathcal{I}}^* = \mathbf{0}$, we have

Input: Matrix $\mathbf{H} = \mathbf{H}^T > 0$, vector $\bar{\mathbf{u}}$; initial guess \mathbf{p} and initial active set $\mathcal{I} \subseteq \{1, \dots, n\}$ such that $\mathbf{p}_{\{1, \dots, n\} \setminus \mathcal{I}} = \mathbf{0}$; maximum number K_{\max} of iterations, tolerance $\epsilon > 0$.

-
1. $i \leftarrow 0$; $\bar{\mathcal{I}} \leftarrow \{1, \dots, n\} \setminus \mathcal{I}$;
 2. **while** ($i \leq K_{\max}$ **and** $\min(\mathbf{p}) < -\epsilon$) **or** $i = 0$ **do**:
 - (2.1) $\mathbf{w}_{\mathcal{I}} \leftarrow \mathbf{H}_{\mathcal{I}, \mathcal{I}} \mathbf{p}_{\mathcal{I}} - \bar{\mathbf{u}}_{\mathcal{I}}$;
 - (2.2) $n_w \leftarrow \|\mathbf{w}_{\mathcal{I}}\|_2$;
 - (2.3) $\mathbf{b}_{\mathcal{I}} \leftarrow -\mathbf{w}_{\mathcal{I}}$
 - (2.4) **while** $n_w > \epsilon$ **and** $i \leq K_{\max}$ **do**:
 - (2.4.1) $\mathbf{s}_{\mathcal{I}} \leftarrow \mathbf{H}_{\mathcal{I}, \mathcal{I}} \mathbf{b}_{\mathcal{I}}$;
 - (2.4.2) $\mathbf{p}_{\mathcal{I}} \leftarrow \mathbf{p}_{\mathcal{I}} - \frac{\mathbf{w}_{\mathcal{I}}^T \mathbf{b}_{\mathcal{I}}}{\mathbf{b}_{\mathcal{I}}^T \mathbf{s}_{\mathcal{I}}} \mathbf{b}_{\mathcal{I}}$;
 - (2.4.3) $\bar{\mathbf{w}}_{\mathcal{I}} \leftarrow \mathbf{H}_{\mathcal{I}, \mathcal{I}} \mathbf{p}_{\mathcal{I}} - \bar{\mathbf{u}}_{\mathcal{I}}$;
 - (2.4.4) $\mathbf{b}_{\mathcal{I}} \leftarrow -\bar{\mathbf{w}}_{\mathcal{I}} + \frac{\bar{\mathbf{w}}_{\mathcal{I}}^T \mathbf{s}_{\mathcal{I}}}{\mathbf{b}_{\mathcal{I}}^T \mathbf{s}_{\mathcal{I}}} \mathbf{b}_{\mathcal{I}}$;
 - (2.4.5) $\mathbf{w}_{\mathcal{I}} \leftarrow \bar{\mathbf{w}}_{\mathcal{I}}$;
 - (2.4.6) $n_w \leftarrow \|\mathbf{w}_{\mathcal{I}}\|_2$;
 - (2.4.7) $i \leftarrow i + 1$;
 - (2.5) **for** $j \in \mathcal{I}$ **do**:
 - (2.5.1) **if** $\mathbf{p}_j < -\epsilon$ **then** $\mathbf{p}_j \leftarrow 0$; $\mathcal{I} \leftarrow \mathcal{I} \setminus \{j\}$; $\bar{\mathcal{I}} \leftarrow \bar{\mathcal{I}} \cup \{j\}$;
 3. $\mathbf{p}^* \leftarrow \mathbf{p}$;
 4. $\mathbf{u}_{\mathcal{I}}^* = \bar{\mathbf{u}}_{\mathcal{I}}$, $\mathbf{u}_{\bar{\mathcal{I}}}^* \leftarrow \mathbf{H}_{\bar{\mathcal{I}}, \mathcal{I}} \mathbf{p}_{\mathcal{I}}^*$;
 5. **end**.
-

Output: Contact force vector \mathbf{p}^* and normal displacement vector \mathbf{u}^* .

Algorithm 1: Greedy method with Conjugate Gradient (greedy CG)

$$\begin{bmatrix} \mathbf{w}_{\mathcal{I}}^* \\ \mathbf{w}_{\bar{\mathcal{I}}}^* \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{H}_{\bar{\mathcal{I}}, \mathcal{I}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{p}_{\mathcal{I}}^* \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ -\bar{\mathbf{u}}_{\bar{\mathcal{I}}} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{\mathcal{I}, \mathcal{I}} & \mathbf{H}_{\mathcal{I}, \bar{\mathcal{I}}} \\ \mathbf{H}_{\bar{\mathcal{I}}, \mathcal{I}} & \mathbf{H}_{\bar{\mathcal{I}}, \bar{\mathcal{I}}} \end{bmatrix} \begin{bmatrix} \mathbf{p}_{\mathcal{I}}^* \\ \mathbf{p}_{\bar{\mathcal{I}}}^* \end{bmatrix} + \begin{bmatrix} -\bar{\mathbf{u}}_{\mathcal{I}} \\ -\bar{\mathbf{u}}_{\bar{\mathcal{I}}} \end{bmatrix}$$

and hence $\mathbf{u}^* = \mathbf{w}^* + \bar{\mathbf{u}} = \mathbf{H} \mathbf{p}^*$. The complementarity condition $(\mathbf{u}^* - \bar{\mathbf{u}})^T \mathbf{p}^* = (\mathbf{w}^*)^T \mathbf{p}^* = 0$ follows by construction, as Step 2.4 zeroes all the components of \mathbf{w}_j^* that correspond to nonnegative \mathbf{p}_j^* , $\forall j \in \mathcal{I}$, and zeroes all the components \mathbf{p}_j^* that correspond to possible nonzero components \mathbf{w}_j^* , $\forall j \in \bar{\mathcal{I}}$.

However, Bemporad and Paggi (2015) demonstrated that there is no formal proof that the condition $\mathbf{w}_{\bar{\mathcal{I}}}^* \geq \mathbf{0}$ (i.e., that $\mathbf{u}^* \geq \bar{\mathbf{u}}$) is satisfied after the algorithm terminates. If the algorithm is applied to randomly generated $\bar{\mathbf{u}}$ vectors and \mathbf{H} positive definite matrices with positive coefficients, Bemporad and Paggi (2015) found that in many cases the LCP is not solved exactly. This problem was noticed especially when the contact domain is densely packed, with many boundary elements close to each other and all in contact. A MATLAB routine of the counterexample is available for download at <http://musam.imtlucca.it/counterexample.m>.

Therefore, as a word of caution, the reliability of the greedy method should be carefully checked in case of applications of the boundary element method to contact

problems governed by other forms of \mathbf{H} , as in the case of contact with an anisotropic or an inhomogeneous half-plane.

Another drawback of the algorithm is the difficulty to warm start the method with a proper choice of the initial active set \mathcal{I} . Since at Step 2.5.1 the number of elements in the sequence \mathcal{I} is decreased by removing negative enough components \mathbf{p}_j of the current solution vector, i.e., eliminating the points bearing tensile (negative) forces, in a monotonic way (no index j that has been removed from \mathcal{I} can be added back), a safe cold start is to set $\mathcal{I} = \{1, \dots, n\}$ and pick up a vector $\mathbf{p} \geq \mathbf{0}$, usually a vector with arbitrary nonnegative numbers. The history of boundary element contact forces obtained during the solution of a sequence of imposed displacements is not taken into account by the method to accelerate its convergence, although we know that contact forces are monotonically increasing functions of the far-field displacement. In any case, for a complex sequence of loading with an increased or decreased far-field displacement, any warm starting on forces cannot be implemented in the method, since the elimination of contact points is irreversible.

A constrained conjugate gradient (CG) algorithm was proposed by Polonsky and Keer (1999) based on the theory by Hestenes (1980, Chaps. 2, 3), to solve the linear system of equations and rigorously impose the satisfaction of the contact constraints. Algorithm 2 has been applied by Polonsky and Keer (1999) to simulations under load control. However, it can be used also for displacement control. The condition for convergence set by Polonsky and Keer (1999) in terms of relative variation in the local contact forces from an iteration to the next has been recast in terms of the error in the local contact displacements. The two criteria are completely equivalent.

This constrained CG algorithm does not remove the points bearing tensile forces from the active set, as the Greedy algorithms do. Therefore, the size of the linear system of equations is not reduced during the iterations, increasing the computation time for its solution. On the other hand, the method assures the satisfaction of the LCP conditions (4.13) and it was found in Bemporad and Paggi (2015) to converge with a reduced number of iterations as compared to the Greedy CG algorithm. Although not investigated in Polonsky and Keer (1999), it can be warm-started in case of a sequence of loading steps by considering both an initial trial contact domain and a set of contact pressures derived from the previous converged solution. The FFT method can be used to accelerate step (3.8) of Algorithm 2, as in Polonsky and Keer (2000a).

The QP problem with positive definite Hessian matrix having the special form (4.15) and corresponding to the LCP can be effectively solved as a nonnegative least squares problem, as proposed by Bemporad and Paggi (2015).

Thanks to property (4.14), matrix \mathbf{H} admits a Cholesky factorization $\mathbf{H} = \mathbf{C}^T \mathbf{C}$. Hence, we can *theoretically* recast problem (4.15) as the nonnegative least squares (NNLS) problem:

$$\min_p \frac{1}{2} \|\mathbf{C}\mathbf{p} - \mathbf{C}^{-T} \bar{\mathbf{u}}\|_2^2 \quad (4.19a)$$

$$\text{s.t. } \mathbf{p} \geq \mathbf{0}. \quad (4.19b)$$

Input: Matrix $\mathbf{H} = \mathbf{H}^T > 0$, vector $\bar{\mathbf{u}}$, initial guess $\mathbf{p} \geq \mathbf{0}$, initial active set $\mathcal{I} = \{1, \dots, n\}$; maximum number K_{\max} of iterations, tolerance $\epsilon > 0$.

1. $i \leftarrow 0, n_{w,\text{old}} = 1, d = 0, \text{err} = +\infty$;
 2. $\mathbf{w} \leftarrow \mathbf{H}\mathbf{p} - \bar{\mathbf{u}}$;
 3. **while** ($i \leq K_{\max}$ **and** $\text{err} > \epsilon$):
 - (3.1) **if** $i = 0$;
 then $\mathbf{t} \leftarrow \mathbf{w}$;
 else: $\mathbf{t} \leftarrow \mathbf{w} + d \frac{n_w}{n_{w,\text{old}}} \mathbf{t}_{\text{old}}$;
 - (3.2) $\tau = \frac{\mathbf{w}^T \mathbf{t}}{\mathbf{t}^T \mathbf{H} \mathbf{t}}$;
 - (3.3) $\mathbf{p} \leftarrow \mathbf{p} - \tau \mathbf{t}$;
 - (3.4) $\forall j \in \mathcal{I} : p_j \leftarrow \max\{p_j, 0\}$;
 - (3.5) Find $I_{ol} = \{j \in \mathcal{I} : p_j = 0, w_j < 0\}$;
 if $I_{ol} = \emptyset$ **then** $d = 1$ **else** $d = 0$; $p_j \leftarrow p_j - \tau w_j, \forall j \in I_{ol}$;
 - (3.6) $\mathcal{I} \leftarrow \{j : p_j > 0\} \cup I_{ol}$;
 - (3.7) $\mathbf{t}_{\text{old}} \leftarrow \mathbf{t}, n_{w,\text{old}} \leftarrow n_w$;
 - (3.8) $\mathbf{w} \leftarrow \mathbf{H}\mathbf{p} - \bar{\mathbf{u}}$;
 - (3.9) $n_w = \|\mathbf{w}\|_2$;
 - (3.10) $\text{err} \leftarrow |n_w - n_{w,\text{old}}|/n_{w,\text{old}}$;
 - (3.11) $i \leftarrow i + 1$;
 4. $\mathbf{p}^* \leftarrow \mathbf{p}; \mathbf{u}^* = \mathbf{H}\mathbf{p}^*$;
 5. **end.**
-

Output: Contact force vector \mathbf{p}^* and normal displacement vector \mathbf{u}^* .

Algorithm 2: Constrained Conjugate Gradient

A simple and effective active-set method for solving the NNLS problem (4.19) is the one in Lawson and Hanson (1974, p. 161) that was extended by Bemporad and Paggi (2015) in Algorithm 3 to directly solve (4.15) without explicitly computing the Cholesky factor \mathbf{C} and its inverse \mathbf{C}^{-1} , and to handle warm starts. After a finite number of steps, Algorithm 3 converges to the optimal contact force vector \mathbf{p}^* and returns the normal displacement vector \mathbf{u}^* whose components $p_{i,j}$, $u_{i,j}$ satisfy $p_{i,j} \geq 0$, $u_{i,j} \geq \bar{u}_{i,j}$, $(u_{i,j} - \bar{u}_{i,j})p_{i,j} = 0$, and (4.12), $\forall (i, j) \in I_C$.

The method is easy to warm start in case of a loading scenario consisting of an alternating sequence of increasing or decreasing far-field displacements. The contact forces determined for a given imposed displacement are used to initialize vector \mathbf{p} . Due to the monotonicity of the contact solution, this initialization is certainly much closer to the optimal solution \mathbf{p}^* than a zero vector. This usually significantly reduces the iterations of the method to convergence. Such a warm start has a fast implementation requiring a projection of the forces of the points belonging to $I_C^*(\Delta_k)$ to the same points of the trial domain $I_C^*(\Delta_{k+1})$ for a new imposed far-field displacement Δ_{k+1} . For an increasing far-field displacement, i.e., $\Delta_{k+1} > \Delta_k$ the forces in the elements belonging to $I_C^*(\Delta_{k+1}) - I_C^*(\Delta_k)$ are simply initialized equal to zero.

Input: Matrix $\mathbf{H} = \mathbf{H}^T > 0$, vector $\bar{\mathbf{u}}$, initial guess \mathbf{p} ; maximum number K_{\max} of iterations, tolerance $\epsilon > 0$.

1. $\mathcal{I} \leftarrow \{i \in \{1, \dots, n\} : \mathbf{p}_i > \mathbf{0}\}$; $init \leftarrow \text{FALSE}$; $k \leftarrow 0$;
2. **if** $\mathcal{I} = \emptyset$ **then** $init \leftarrow \text{TRUE}$;
3. $\mathbf{w} \leftarrow \mathbf{H}\mathbf{p} - \bar{\mathbf{u}}$;
4. **if** $(\mathbf{w} \geq -\epsilon$ **or** $\mathcal{I} = \{1, \dots, n\})$ **and** $init = \text{TRUE}$) **or** $k \geq K_{\max}$ **then go to** Step 13;
5. **if** $init = \text{TRUE}$ **then** $i \leftarrow \arg \min_{i \in \{1, \dots, n\} \setminus \mathcal{I}} \mathbf{w}_i$; $\mathcal{I} \leftarrow \mathcal{I} \cup \{i\}$; **else** $init \leftarrow \text{TRUE}$;
6. $\mathbf{s}_I \leftarrow$ solution of the linear system $\mathbf{H}_I \mathbf{s}_I = \bar{\mathbf{u}}_I$
7. **if** $\mathbf{s}_I \geq -\epsilon$ **then** $\mathbf{p} \leftarrow \mathbf{s}$ **and go to** Step 3;
8. $j \leftarrow \arg \min_{h \in \mathcal{I} : \mathbf{s}_h \leq 0} \left\{ \frac{\mathbf{p}_h}{\mathbf{p}_h - \mathbf{s}_h} \right\}$;
9. $\mathbf{p} \leftarrow \mathbf{p} + \frac{\mathbf{p}_j}{\mathbf{p}_j - \mathbf{s}_j} (\mathbf{s} - \mathbf{p})$;
10. $\mathcal{I}_0 \leftarrow \{h \in \mathcal{I} : \mathbf{p}_h = \mathbf{0}\}$;
11. $\mathcal{I} \leftarrow \mathcal{I} \setminus \mathcal{I}_0$; $k \leftarrow k + 1$;
12. **go to** Step 6;
13. $\mathbf{p}^* \leftarrow \mathbf{p}$;
14. $\mathbf{u}^* \leftarrow \mathbf{w} + \bar{\mathbf{u}}$;
15. **end.**

Output: Contact force vector \mathbf{p}^* and normal displacement vector \mathbf{u}^* satisfying $\mathbf{u}^* = \mathbf{H}\mathbf{p}$, $\mathbf{u}^* \geq \bar{\mathbf{u}}$, $\mathbf{p}^* \geq \mathbf{0}$, $(\mathbf{u}^* - \bar{\mathbf{u}})^T \mathbf{p} = 0$.

Algorithm 3: Non-Negative Least Squares (NNLS)

Note that Step 6 of Algorithm 3 is equivalent to Step 2.4 of Algorithm 3 and it has been performed in Algorithm 1 by using the MATLAB's `mldivide` solver. This step can be accelerated by the use of an approach based on the FFT (for its implementation, see, e.g., Batrouni et al. (2002)). Alternatively, since the set \mathcal{I}_0 changes incrementally during the iterations of the algorithm, more efficient iterative QR (Lawson and Hanson, 1974, Chap. 24) or LDL^T Bemporad (2014) factorization methods can be employed.

An alternative method to solve the QP problem (4.15) is to use an accelerated gradient projection (GP) method for QP (Nesterov 1983; Patrinos and Bemporad 2014). Because of the simple nonnegative constraints in (4.15), rather than going to the dual QP formulation as in Patrinos and Bemporad (2014), the GP problem was formulated directly for the primal QP problem (4.15). Numerical experiments have shown slow convergence of a pure accelerated GP method to solve (4.15). However, the method can be used to *warm start* Algorithm 3, as described in Algorithm 4 proposed in Bemporad and Paggi (2015). If Algorithm 4 is executed ($K > 0$), it returns a vector \mathbf{p} that is immediately used as an input to Algorithm 3, otherwise one can simply set $\mathbf{p} = \mathbf{0}$ (cold start). As shown in the algorithms' comparison in the sequel, GP iterations provide large benefits in warm starting the NNLS solver, therefore allowing taking the best advantages of the two methods: quickly getting in the neighborhood of the optimal solution (GP iterations of Algorithm 4) and getting

solutions up to machine precision after a finite number of iterations (the active-set NNLS Algorithm 3).

Input: Matrix $\mathbf{H} = \mathbf{H}^T > 0$ and its Frobenius norm L , vector $\bar{\mathbf{u}}$, initial guess \mathbf{p} , number K of iterations.

1. $\bar{\mathbf{p}} \leftarrow \mathbf{p}$;
2. **for** $i = 0, \dots, K - 1$ **do**:
 - (2.1) $\beta = \max\{\frac{i-1}{i+2}, 0\}$;
 - (2.2) $\mathbf{s} = \mathbf{p} + \beta(\mathbf{p} - \bar{\mathbf{p}})$;
 - (2.3) $\mathbf{w} = \mathbf{H}\mathbf{s} - \bar{\mathbf{u}}$;
 - (2.4) $\bar{\mathbf{p}} \leftarrow \mathbf{p}$;
 - (2.5) $\mathbf{p} \leftarrow \max\{\mathbf{s} - \frac{1}{L}\mathbf{w}, \mathbf{0}\}$;
3. **end**.

Output: Warm start for contact force vector \mathbf{p} and elastic correction vector \mathbf{w} .

Algorithm 4: Accelerated Gradient Projection (GP) to be used to warm start the NNLS algorithm

The QP problem (4.15) can also be solved by the alternating direction method of multipliers (ADMM), which belongs to the class of augmented Lagrangian methods. The reader is referred to Boyd et al. (2011) for mathematical details, while its application to contact problems was proposed in Bemporad and Paggi (2015). The method treats the QP (4.15) as the following problem:

$$\begin{aligned} \min_{\mathbf{p}, \mathbf{s}} \quad & \frac{1}{2} \mathbf{p}^T \mathbf{H} \mathbf{p} - \bar{\mathbf{u}}^T \mathbf{p} + g(\mathbf{s}) \\ \text{s.t.} \quad & \mathbf{p} = \mathbf{s}, \end{aligned} \quad (4.20)$$

where

$$g(\mathbf{s}) = \begin{cases} 0 & \text{if } \mathbf{s} \geq \mathbf{0} \\ +\infty & \text{if } \mathbf{s} < \mathbf{0} \end{cases}.$$

Then, the augmented Lagrangian function

$$L_\rho(\mathbf{p}, \mathbf{s}, \mathbf{w}) = \frac{1}{2} \mathbf{p}^T \mathbf{H} \mathbf{p} - \bar{\mathbf{u}}^T \mathbf{p} + g(\mathbf{s}) + \mathbf{w}^T (\mathbf{p} - \mathbf{s}) + \frac{\rho}{2} \|\mathbf{p} - \mathbf{s}\|_2^2$$

is considered, where $\rho > 0$ is a parameter of the algorithm. The basic ADMM algorithm consists of the following iterations:

$$\begin{aligned} \mathbf{p}^{k+1} &= \arg \min_{\mathbf{p}} L_\rho(\mathbf{p}, \mathbf{s}^k, \mathbf{w}^k) \\ \mathbf{s}^{k+1} &= \arg \min_{\mathbf{s}} L_\rho(\mathbf{p}^{k+1}, \mathbf{s}, \mathbf{w}^k) \\ \mathbf{w}^{k+1} &= \mathbf{w}^k + \rho(\mathbf{p}^{k+1} - \mathbf{s}^{k+1}). \end{aligned} \quad (4.21)$$

A scaled form with over-relaxation of the ADMM iterations (4.21) is summarized in Algorithm 5. The algorithm is guaranteed to converge asymptotically to the solution \mathbf{p}^* , \mathbf{u}^* of the problem. The over-relaxation parameter $\alpha > 1$ is introduced to improve convergence. Typical values for α suggested in Boyd et al. (2011) are $\alpha \in [1.5, 1.8]$.

A warm start of the algorithm that takes into account the loading history is possible in a way analogous to that described for the NNLS approach. However, as an additional complexity, also an initialization for the dual variable vector \mathbf{w} must be provided, possibly obtained by projecting the solution obtained for a certain Δ_k to that for Δ_{k+1} .

Input: Matrix $\mathbf{H} = \mathbf{H}^T > 0$, vector $\bar{\mathbf{u}}$, initial guesses \mathbf{p} , \mathbf{w} , parameter $\rho > 0$, over-relaxation parameter $\alpha > 1$, maximum number K_{\max} of iterations, tolerance $\epsilon > 0$.

1. $\mathbf{M} \leftarrow (\frac{1}{\rho}\mathbf{H} + \mathbf{I})^{-1}$;
2. $\mathbf{w}_\rho \leftarrow -\frac{1}{\rho}\mathbf{w}$;
3. $\mathbf{s} \leftarrow \mathbf{p}$;
4. $i \leftarrow 0$;
5. **while** ($i \leq K_{\max}$ **and** $\|\mathbf{p} - \mathbf{s}\|_\infty > \epsilon$) **or** $i = 0$ **do**:
 - (5.1) $\mathbf{s} \leftarrow \mathbf{M}(\mathbf{p} - \mathbf{w}_\rho - \frac{1}{\rho}\bar{\mathbf{u}})$;
 - (5.2) $\bar{\mathbf{s}} \leftarrow \alpha\mathbf{s} + (1 - \alpha)\mathbf{p}$;
 - (5.3) $\mathbf{p} \leftarrow \max\{\bar{\mathbf{s}} + \mathbf{w}_\rho, \mathbf{0}\}$;
 - (5.4) $\mathbf{w}_\rho \leftarrow \mathbf{w}_\rho + \bar{\mathbf{s}} - \mathbf{p}$;
 - (5.5) $i \leftarrow i + 1$;
6. $\mathbf{p}^* \leftarrow \mathbf{p}$;
7. $\mathbf{u}^* \leftarrow \bar{\mathbf{u}} - \rho\mathbf{w}_\rho$;
8. **end**.

Output: Contact force vector \mathbf{p}^* and normal displacement vector \mathbf{u}^* satisfying $\mathbf{u}^* = \mathbf{H}\mathbf{p}$, $\mathbf{u}^* \geq \bar{\mathbf{u}}$, $\mathbf{p}^* \geq \mathbf{0}$, $(\mathbf{u}^* - \bar{\mathbf{u}})^T \mathbf{p} = 0$.

Algorithm 5: Alternative Direction Method of Multipliers (ADMM)

Comparison of the algorithms' performance. To assess the computation efficiency and performance in terms of number of iterations required to achieve convergence, the optimization algorithms reviewed in the previous section can be compared in relation to a benchmark frictionless normal contact problem involving a numerically generated fractal rough surface and a half-plane.

To this aim, the random midpoint displacement algorithm (Peitgen and Saupe 1988) can be used to generate the synthetic height field of surfaces with multiscale fractal roughness, i.e., with a power spectral density (PSD) function of the height field of power-law type. The surface with a given resolution (pre-fractal) is realized by a successive refinement of an initial coarse representation by adding a sequence of

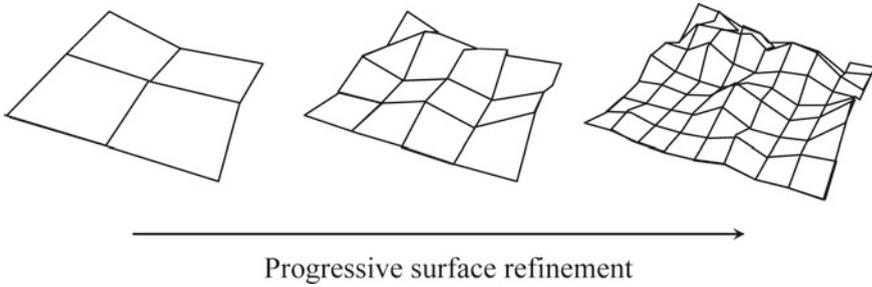


Fig. 4.2 Rough surfaces with multiscale roughness and different resolutions, numerically generated by the random midpoint displacement algorithm. From Bemporad and Paggi (2015)

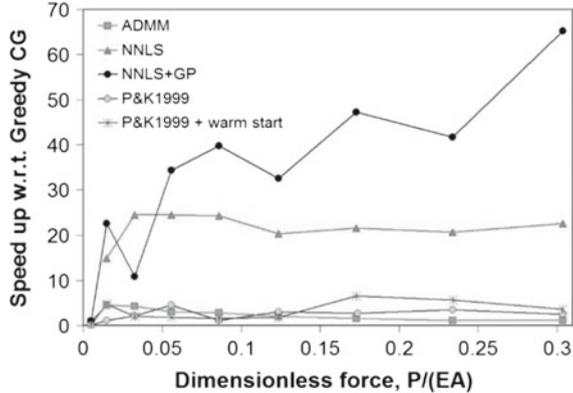
intermediate heights whose elevation is extracted from a Gaussian distribution with a suitable rescaled variance, see a qualitative sketch in Fig. 4.2. Several applications of the method to model rough surfaces for contact mechanics simulations are available in Zavarise et al. (2004a, b), Paggi and Ciavarella (2010).

In particular, let us consider a test problem consisting of a surface with Hurst exponent $H = 0.7$, lateral size $L = 100 \mu\text{m}$, and 512 heights per side. The surface is brought into contact with an elastic half-plane under displacement control. Ten displacement steps are imposed to reach a maximum far-field displacement which is set equal to $(\xi_{\max} - \xi_{\text{ave}})/2$, where ξ_{\max} and ξ_{ave} are the maximum and the average elevations of the rough surface, respectively. All the simulations were carried out with the server 653745-421 Proliant DL585R07 from Hewlett Packard with 128 GB Ram, 4 processors AMD Opteron 6282 SE 2.60 GHz with 16 cores running MATLAB R2014b.

The parameters for the Greedy CG method are the maximum number of iterations $K_{\max} = 1 \times 10^5$ and the convergence tolerance $\epsilon = 1 \times 10^{-8}$. The contact forces are initialized at zero (cold start). The constrained CG method also considers $K_{\max} = 1 \times 10^5$ and the same tolerance $\epsilon = 1 \times 10^{-8}$. Both the original version by Polonsky and Keer (1999) (labeled P&K1999 in Fig. 4.3) and its warm-started variant (labeled P&K1999 + warm start in Fig. 4.3) are considered.

For the NNLS algorithm (Algorithm 3), the warm start strategy based on the projection of contact forces from the solution corresponding to a previous displacement step is adopted. Alternatively, warm starting using gradient projections (denoted as NNLS+GP) is examined, using 100 gradient projections to initialize vector \mathbf{p} . The parameters for the ADMM method are $\alpha = 1.5$, $\rho = 1$, $K_{\max} = 3 \times 10^3$, and $\epsilon = 10^{-8}$. The total number n of optimization variables is varying with the amount of imposed displacement Δ and therefore with the force level. For the highest indentation level of the present test, $n = 35,555$. Warm starting the algorithm is achieved by projecting primal variables as for the NNLS and dual variables \mathbf{w} as well. The

Fig. 4.3 Comparison between the optimization algorithms in terms of speedup of computation time



projection simply consists of assigning the values of $p_{i,j}^*$ and w^* of the boundary elements in contact for the step Δ_k to the same boundary elements belonging to the trial contact domain I_C corresponding to the higher indentation Δ_{k+1} .

Once convergence is achieved for each imposed far-field displacement, the optimization algorithms provide the same normal force P and contact domains, with small roundoff errors due to finite machine precision. The ratio between the CPU time required by each method to achieve convergence and the CPU time employed by the Greedy CG algorithm, which is the slowest, is considered as a measure of speedup. This ratio is plotted vs. the dimensionless normal force $P/(EA)$ in Fig. 4.3, where $A = L^2$ is the nominal contact area. The best performance is achieved by the application of the NNLS method with 100 gradient projections (GP), which is 25 times faster than the original constrained CG method by Polonsky and Keer (1999) and about two orders of magnitude faster than the ADMM and the Greedy CG algorithms, with an increasing efficiency for high loads. The NNLS with warm start is also very well performing, with a stable speedup of about 25 times for any load level.

As outlined in the introduction, the Greedy method can be used in conjunction with other algorithms for solving the unconstrained linear system of equations (Step 2.4) than the CG algorithm. For instance, the CG Step 2.4 in the Greedy algorithm can be replaced with the optimized built-in `mldivide` function of MATLAB, or with the Gauss–Seidel algorithm, as proposed in Borri-Brunetto et al. (1999, 2001).

The MATLAB’s `mldivide` solver (which employs the Cholesky factorization) leads to a reduction of computation time of 30–40% with respect to the CG method, almost regardless of the size of the system n , see Fig. 4.4. Even with this gain in computation speed, the overall performance is still quite far from that of the NNLS Algorithm 3 on the platform used for the tests. Moreover, the MATLAB solver leads to an error of lack of memory for $n > 20,000$, a serious problem for large systems that are not suffered by the CG solver described in Step 2.4 of Algorithm 1. The Gauss–Seidel algorithm does not suffer for the lack of memory but it is about 3 times slower than the CG method.

Fig. 4.4 Ratio between computation times for the Greedy method using different solvers (MATLAB's `mldivide` solver or Gauss–Seidel algorithm) as compared to the conjugate gradient (CG) algorithm, for different sizes n of the contact superset I_C . Adapted from Bemporad and Paggi (2015)

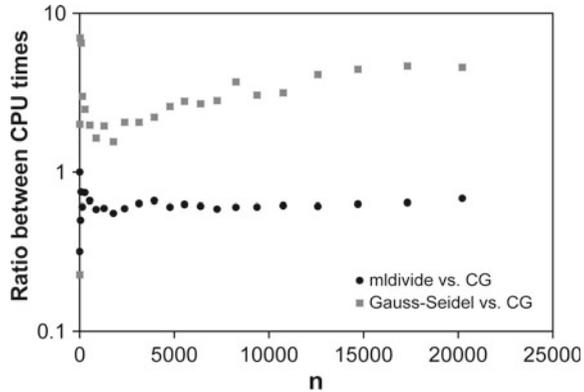
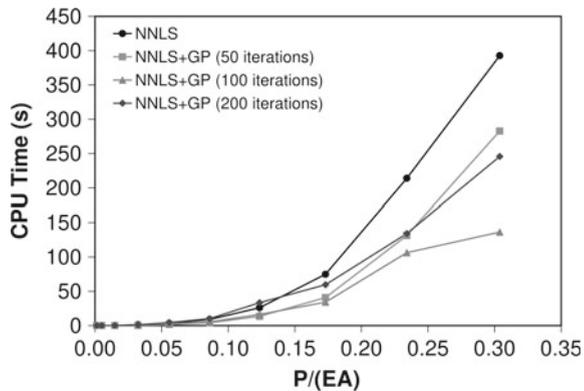


Fig. 4.5 Computation times of the NNLS algorithm depending on the number K of gradient projection (GP) iterations. Adapted from Bemporad and Paggi (2015)



The effect of the number K of GP iterations applied before the NNLS algorithm was also investigated in Bemporad and Paggi (2015). Figure 4.5 shows, for the same test problem whose results are shown in Fig. 4.3, the effect of K on the total computation time. For K from 0 to 100, we observe a reduction in the total computation time due to a decrease in the number of iterations requested by the NNLS algorithm to achieve convergence, thanks to a better initial guess of \mathbf{p} . However, a further increase in K (see, e.g., the curve in Fig. 4.5 corresponding to $K = 200$ iterations) does not correspond to further savings of CPU time. This is due to the fact that the number of NNLS iterations was already reduced to its minimum for $K = 100$ GP iterations, so that the application of further gradient projections is just leading to additional CPU time without further benefit.

Further acceleration strategies and variants of the boundary element method.

A further speedup of computation time, as compared to the NNLS method, can be achieved by improving the criterion for the guess of the initial set I_C of points in contact. The standard criterion based on checking the interpenetration of the surface

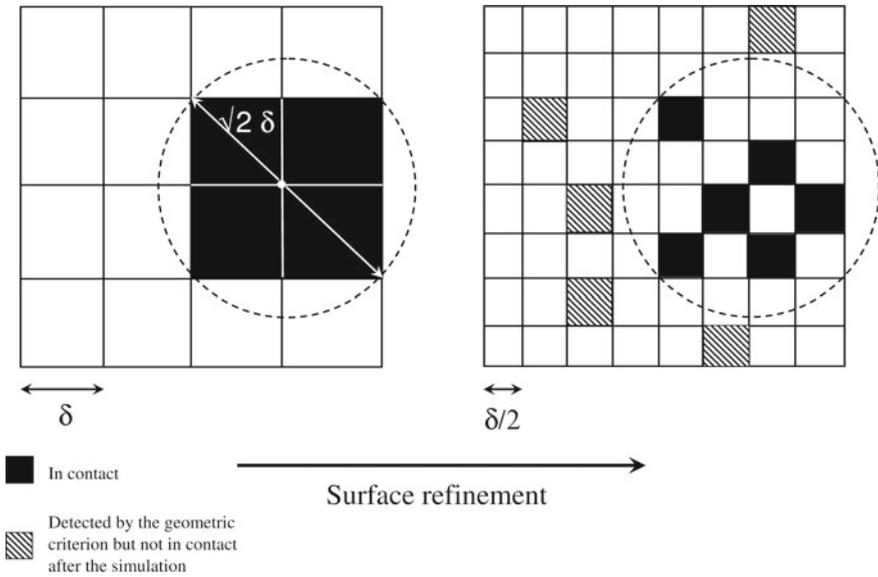


Fig. 4.6 A sketch illustrating the property of lacunarity of the contact domain: the real contact area progressively diminishes by refining the surface, until vanishes in the fractal limit of $\delta \rightarrow 0$. This implies that some boundary elements detected by the rigid body interpenetration criterion (dashed gray elements) can be neglected a priori since they are outside the real contact area corresponding to the coarse scale contact solution. From Bemporad and Paggi (2015)

heights into the half-plane in case of a rigid body motion is the most conservative one. However, at convergence, we know that only a small subset I_C^* of that initial set is actually in contact. Therefore, a better choice of the initial trial contact domain would reduce the size of the system of linear equations with an expected benefit in terms of computation time.

As shown in Borri-Brunetto et al. (1999) via numerical simulations on pre-fractal surfaces with Hurst exponent $H > 0.5$ and different resolutions by refining the surface height field via a recursive application of the random midpoint displacement algorithm, the real contact area of each surface representation decreases by reducing the grid spacing δ , as illustrated in the sketch in Fig. 4.6. In the fractal limit of $\delta \rightarrow 0$, the real contact area vanishes. Therefore, this property of lacunarity implies that the heights that are not in contact for a coarser surface representation are not expected to come into contact by a successive refining of the height field, for the same imposed far-field displacement.

Therefore, a possible better criterion was proposed by Bemporad and Paggi (2015) and was called cascade multiresolution (CMR) algorithm. In the method, the initial trial contact domain is selected by retaining, among all the heights selected by the rigid body interpenetration check, only those located within the areas of influence of

the nodes belonging to the contact domain of a coarser representation of the rough surface for the same imposed displacement Δ . This criterion, inspired by fractal considerations, shares some analogies with multigrid methods, where coarse and fine grids are used to compute the contact solution.

As graphically shown in Fig. 4.6, an area of influence of a given node in contact can be defined by the radius $\sqrt{2}\delta$, where δ is the grid size of the coarser surface representation. Since the criterion is not exact, Bemporad and Paggi (2015) suggested to consider a multiplicative factor h larger than unity for the radius defining the nodal area of influence. It is remarkable to note that this numerical scheme can be applied recursively to a cascade of coarser representations of the same rough surface. As a general trend, computation time is expected to drastically diminish by increasing the number of cascade projections. However, the propagation of errors due to the wrong exclusion of heights that would actually make contact cannot be controlled by the algorithm and it is expected to increase with the number of projections as well. The advantage of the method is represented by the fact that, in addition to saving computation time with respect to that required by the NNLS algorithm to solve just one contact problem for the finest surface, all the contact predictions for the coarser scale representations of the same surface are provided without additional costs, which is a useful result for the multiscale characterization of contact problems. Moreover, the CMR method can be used in conjunction with any of the optimization algorithms reviewed in the previous sections. The algorithm is illustrated in Algorithm 6.

Input: $s = 1, \dots, l$ surface representations with different resolution or grid spacing $\delta(s)$; area of influence parameter $h \geq 1$.

1. **for** $s = 1, \dots, l$ **do**:

(1.1) Determine $I_C(s) = \{(i, j) \in I_N(s) : \xi_{i,j} \geq \xi_{\max}(s) - \Delta\}$;

(1.2) **if** $s = 1$ **then** $I_{C,p}(s) = I_C(s)$

else

$I_{C,p}(s) = \{(i, j) \in I_C(s) : r_{i-k,j-l} = \|\mathbf{x}_{i,j} - \mathbf{x}_{k,l}\| \leq h\delta(s-1)\}, \forall (k, l) \in I_C^*(s-1)$

end

2. Construct \mathbf{H} based on the projected trial contact domain $I_{C,p}(s)$;

3. Apply optimization algorithms (e.g., NNLS) and determine $\mathbf{p}^*, \mathbf{u}^*, I_C^*(s)$;

4. **end**.

Algorithm 6: Cascade multi-resolution (CMR) algorithm

To assess the computational performance of the method, the CMD algorithm was applied in conjunction with the NNLS algorithm to pre-fractal surfaces with different H numerically generated by the RMD method (Bemporad and Paggi 2015). As an example, the lateral size was set equal to 100 μm for all the surfaces and the finest

resolution whose contact response has to be sought corresponded to 256 heights per side. The method requires the storage of the coarser representations of such surfaces that are in any case available by the RMD algorithm during its various steps of random addition.

The cascade of projections was applied starting with a coarser representation of the surfaces with only 16 heights per side and then considering 32, 64, 128, and finally 256 heights per side. A parameter $h = 2$ was used for the definition of the area of influence. The solution of the contact problem for the surface with 16 heights per side was obtained in an exact form, since it is the starting point of the cascade projections, whereas the contact predictions for the finer surface representations can be affected by an error intrinsic in the criterion. The approximate predictions for the surface with 256 heights per side were compared with the reference solution corresponding to the application of the NNLS algorithm with warm start directly to the finest representation of the rough surface.

The computation time of the CMR+NNLS solution is the sum of the CPU time required to solve all the coarser surface representations and it is found to be much less than the CPU time required by the NNLS algorithm to solve just one single surface with the finest resolution, see Fig. 4.7, where we observe a reduction of about 50% in CPU time almost regardless of H . The relative error in the computation of the maximum normal force between the predicted solution and the reference one is a

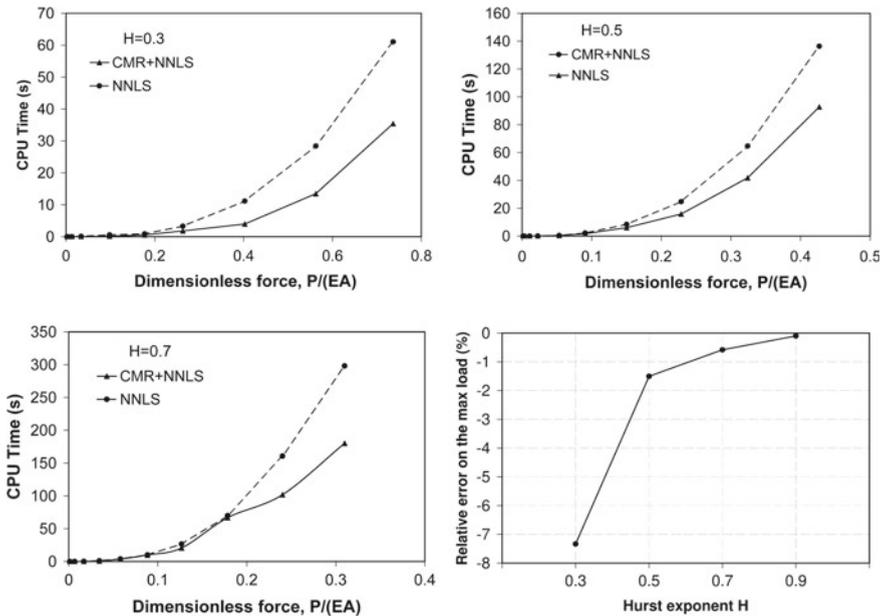
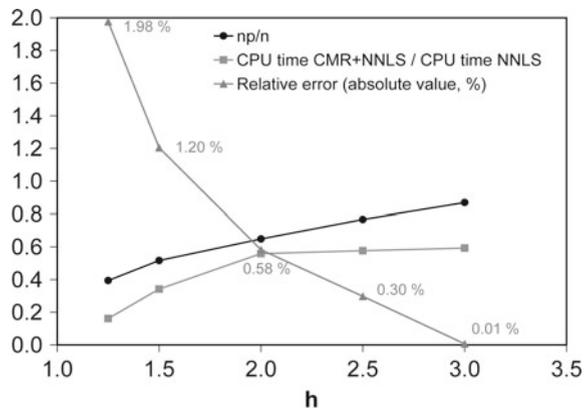


Fig. 4.7 Performance of the CMR+NNLS method applied to numerically generated fractal surfaces with a different Hurst exponent H and $h = 2$. Adapted from Bemporad and Paggi (2015)

Fig. 4.8 Performance of the CMR+NNLS method with respect to NNLS for a numerically generated fractal surface with $H = 0.7$, depending on the parameter h . Adapted from Bemporad and Paggi (2015)



rapidly decreasing function of H , as shown in Fig. 4.7. Considering that real surfaces have often a Hurst exponent $H > 0.5$, the method is very promising.

A synthetic diagram illustrating the effect of the parameter h for the surface with $H = 0.7$ and for a single imposed displacement corresponding to the maximum load is shown in Fig. 4.8. The relative error is rapidly decreasing to values less than 1% by increasing h . The ratio between the number of points expected to be in contact after the application of the CMR projection criterion, n_p , and the number of points that would be included by using the classic rigid body interpenetration check, n , is ranging from 0.4 to 0.8 by increasing h from 1.25 to 3.0. The ratio between CPU times, on the other hand, tends to an asymptotic value of 0.6, which implies a saving of 40% of computation time as compared to the exact solution, with less than 0.01% of relative error.

Among the variants of the boundary element method published in the literature, some concerned with the treatment of the key features of roughness and its evolution during contact. Starting from the fundamental assumption of micromechanical contact theories that only the asperities, i.e., the local maxima of the rough surface, make contact (Greenwood and Williamson 1966; Zavarise et al. 2004a; Paggi and Ciavarella 2010), then one could simplify the boundary element approach by treating only the set of asperities as boundary elements. Each asperity can be modeled as a paraboloid whose geometry is defined by its position (x, y) in the plane, its elevation z above a reference plane, and its mean radius of curvature R , as proposed in Greenwood (2006). For such a Hertzian asperity, the theory of elasticity provides the displacement of the half-plane in the location of the asperity itself, as well as the displacement in any other position, see Nowell and Hills (1989). Such an information can be used to define the Green functions for the application of a generalized boundary element method, where a recursive elimination of asperities supporting tensile forces can be implemented till the final active set of asperities in contact and the

corresponding normal contact forces are identified, using, for instance, the Greedy algorithm. Alternatively, the problem can be formulated by an iterative correction to the asperity deformation as proposed in Ciavarella et al. (2006), avoiding the inversion of the compliance matrix. Moreover, the scheme allows also studying the effect of short- or long-range elastic interaction effects, as carefully investigated in Paggi and Barber (2011).

The above multi-asperity contact problem overcomes the limitations of the original semi-analytical contact theory by Greenwood and Williamson (1966) which did not include elastic interactions in its original form. Moreover, the correction scheme proposed in Ciavarella et al. (2006) inspired also a way to improve the Greenwood and Williamson contact theory by introducing a yet simple but effective mean pressure elastic interaction effect (Ciavarella et al. 2008b). Moreover, the approach presents several advantages over the standard boundary element method, especially in terms of speedup of computation time which is mainly achieved by the fact that the size of the set of potential asperities in contact is significantly much smaller than the total number of boundary elements. On the other hand, drawbacks regard the fact that the method strongly relies on the geometrical parameters of the asperities, which are resolution-dependent as discussed in Majumdar and Bhushan (1990), Zavarise et al. (2004b). Another limitation as compared to a more general boundary element method regards the fact that the contact problem relies on the undeformed asperity geometry. This limitation has been partially overcome by Afferrante et al. (2012), who proposed an update of the asperity detection and their geometrical parameters during contact, to model the phenomenon of merging of asperities by forming bigger ones with completely different geometrical features as compared to what estimated from the undeformed configuration.

The Contact Problem with Friction

In the most general three-dimensional contact problem with friction, the surface displacement vector projected onto an orthogonal Cartesian frame $Oxyz$ has three components, u_x , u_y , and u_z . The component u_z is the component perpendicular to the mean plane of the nominally flat rough surface and it corresponds to the variable u for the frictionless normal contact problem detailed in the previous section. The other components u_x and u_y correspond, on the other hand, to the in-plane surface displacements. The latter are intimately connected to the corresponding surface tractions q_x , q_y , and p through the Green functions.

For instance, a uniform distributed normal traction $p(\mathbf{y})$ acting over a square surface element S of lateral size δ identified by the position vector $\mathbf{y} = (x', y')^T$ leads to the following surface displacements at another point $\mathbf{x} = (x, y)^T$ at a distance $r = \|\mathbf{x} - \mathbf{y}\|$:

$$u_x = -\frac{1-2\nu}{4\pi G} \int_S \frac{x-x'}{r^2} p(\mathbf{y}) d\mathbf{y}, \quad (4.22a)$$

$$u_y = -\frac{1-2\nu}{4\pi G} \int_S \frac{y-y'}{r^2} p(\mathbf{y}) d\mathbf{y}, \quad (4.22b)$$

$$u_z = \frac{1-\nu}{2\pi G} \int_S \frac{p(\mathbf{y})}{r} d\mathbf{y}. \quad (4.22c)$$

Similarly, for a uniform distributed tangential traction $q_x(\mathbf{y})$,

$$u_x = \frac{1}{2\pi G} \int_S \left[\frac{1-\nu}{r} + \nu \frac{(x-x')^2}{r^3} \right] q_x(\mathbf{y}) d\mathbf{y}, \quad (4.23a)$$

$$u_y = \frac{1}{2\pi G} \int_S \nu \frac{(x-x')(y-y')}{r^3} q_x(\mathbf{y}) d\mathbf{y}, \quad (4.23b)$$

$$u_z = \frac{1-2\nu}{4\pi G} \int_S \frac{x-x'}{r^2} q_x(\mathbf{y}) d\mathbf{y}, \quad (4.23c)$$

and for a uniform distributed tangential traction $q_y(\mathbf{y})$:

$$u_x = \frac{1}{2\pi G} \int_S \nu \frac{(x-x')(y-y')}{r^3} q_y(\mathbf{y}) d\mathbf{y}, \quad (4.24a)$$

$$u_y = \frac{1}{2\pi G} \int_S \left[\frac{1-\nu}{r} + \nu \frac{(y-y')^2}{r^3} \right] q_y(\mathbf{y}) d\mathbf{y}, \quad (4.24b)$$

$$u_z = \frac{1-2\nu}{4\pi G} \int_S \frac{y-y'}{r^2} q_y(\mathbf{y}) d\mathbf{y}. \quad (4.24c)$$

Therefore, we recognize that the normal contact problem is in general fully coupled with the tangential one, in the sense that a normal pressure induces not only normal displacements but also not-vanishing in-plane deformation. After introducing a boundary element discretization of the nominally rough surface into $N \times N$ elements as for the frictionless normal contact problem, surface displacements in a point defined by the indices i, j are related to uniform surface tractions acting on a square element defined by the indices k, l via a matrix collecting the Green functions, see Love (1999), Pohrt and Li (2014):

$$\begin{Bmatrix} u_x \\ u_y \\ u_z \end{Bmatrix}_{i,j} = \sum_{k=1}^N \sum_{l=1}^N \begin{bmatrix} H_{xx} & H_{xy} & H_{xz} \\ H_{yx} & H_{yy} & H_{yz} \\ H_{zx} & H_{zy} & H_{zz} \end{bmatrix}_{ijkl} \begin{Bmatrix} q_x \\ q_y \\ p \end{Bmatrix}_{k,l}, \quad (4.25)$$

where $H_{zx} = -H_{xz}$, $H_{zy} = -H_{yz}$, $H_{yx} = H_{xy}$.

The special case $\nu = 1/2$ leads to $H_{xz} = H_{yz} = 0$ and therefore the normal contact problem becomes uncoupled from the tangential one. Another notable case corresponds to $\nu = 0$, where the coefficient H_{xy} vanishes and $H_{xx} = H_{yy} = 2H_{zz}$, leading to coupling between the normal and the tangential contact problems, but uncoupling between the two in-plane directions.

The computation of the surface displacements requires a convolution of the traction effects according to the application of Eq. (4.25), which has a complexity of the order of $O(N^4)$ operations. Although the number of boundary elements $\#I_C$ included in the superset of the possible candidates in contact, I_C , is usually smaller than N^2 , a speedup is certainly required. To this aim, a multilevel multi-integration procedure has been proposed in Lubrecht and Ioannides (1991) reducing the complexity to $O(N^2 \log N)$ by carrying out a summation over a coarser grid and then introducing a correction in the vicinity of the point i, j . The same complexity can be reached by doing the convolution in the Fourier space using a Fast Fourier Transform technique, as proposed in Vollebregt (2014), Pohrt and Li (2014).

Suppose now to apply a monotonically increasing displacement in the normal direction, till a given maximum value which is then held constant. Afterward, a monotonically increasing displacement is applied in a tangential (in-plane) direction. While the normal contact problem is ruled by the unilateral contact condition, in the tangential direction it is customary to postulate the existence of two regimes locally valid for any point in contact: *stick*, when the points of the two surfaces are intimately adhering to each other, and *slip*, when the points experience a relative displacement in the tangential direction. The distinction between such two states is ruled by the Coulomb law of friction, which affirms that any point with $q \leq \mu p$ is in the stick condition, being μ the local coefficient of static friction. Therefore, the solution of the tangential contact problem requires finding the boundary elements belonging to the stick or to the slip state, for a given fixed normal displacement and an imposed tangential one.

Due to coupling, tangential tractions determined from the solution of the tangential contact problem lead to an additional contribution to the normal displacements, which would demand the recursive solution of the normal contact problem to identify the corresponding updated normal contact tractions. Since coupling is in general weak, this feedback effect is often neglected also in the case of $\nu \neq 1/2$, simply setting $H_{xz} = H_{yz} \cong 0$, see, e.g., the implementation in Pohrt and Li (2014). Physically, this approximation implies that the real contact area does not change due to tangential tractions.

Under these assumptions, after solving the frictionless normal contact problem for a given imposed normal displacement Δ , the set I_C^* of boundary elements in contact ($\#I_C^* = N_C^*$) is known. The subsequent application of a tangential displacement Δ_T leads to a partition of the contact set into two parts: a set where stick is observed, $I_{C,st}^*$, and a set where slip is expected, $I_{C,sl}^*$. Hence, in matrix form, the relation between surface displacements and tangential tractions can be formally partitioned as follows:

$$\begin{Bmatrix} \mathbf{u}_{st} \\ \mathbf{u}_{sl} \end{Bmatrix} = \begin{bmatrix} \mathbf{A}_{st,st} & \mathbf{A}_{st,sl} \\ \mathbf{A}_{sl,st} & \mathbf{A}_{sl,sl} \end{bmatrix} \begin{Bmatrix} \mathbf{q}_{st} \\ \mathbf{q}_{sl} \end{Bmatrix}, \quad (4.26)$$

where \mathbf{u}_{st} collects the value of the imposed far-field tangential displacement Δ_T for all the boundary elements in stick condition. On the other hand, by definition, the vector \mathbf{q}_{sl} collects entries which are given by the product between the friction coefficient μ and the local pressure p acting on the boundary element in slip condition, and therefore it is known.

The solution of the problem requires an iterative algorithm to identify $I_{C,st}^*$, $I_{C,sl}^*$, and all the tractions and the surface displacements. To do so, the set $I_{C,st}^*$ is usually initialized equal to I_C^* , i.e., all the boundary elements in contact are supposed to be in stick condition, while $I_{C,sl}^*$ is empty. Then, the tangential tractions of the boundary elements in stick condition associated to the surface deflections can be computed using the first row of Eq. (4.26):

$$\mathbf{A}_{st,st} \mathbf{q}_{st} = \mathbf{u}_{st} - \mathbf{A}_{st,sl} \mathbf{q}_{sl}, \quad (4.27)$$

which requires the inversion of the matrix $\mathbf{A}_{st,st}$ using an inverse fast convolution (a conjugate gradient algorithm, for instance). In the first iteration, $\mathbf{q}_{sl} = \mathbf{0}$ and \mathbf{u}_{st} is a vector of entries all equal to Δ_T .

If all the boundary elements have $q_{st} < \mu p$, then the initial tentative approximation was correct. However, in general, there will be some elements with a tangential traction overcoming the limit value for the stick condition, and therefore they slip and have to be moved to the $I_{C,sl}^*$ set. Correspondingly, the tangential tractions of those boundary elements have to be limited to the maximum admissible value μp . The tangential deflections \mathbf{u}_{st} and \mathbf{u}_{sl} can now be recomputed from the updated distribution of the stick and the slip tractions using again Eq. (4.26). At this point it is still necessary to check if any boundary element in the tentative slip domain presents $u_{sl} \geq \Delta_T$. If this is the case, then such boundary elements should be sticking and therefore they have to be transferred back to the stick domain and u_{sl} has to be limited by Δ_T . Afterward, for the updated partition of stick and slip domains, another iteration is performed, which requires the computation of \mathbf{q}_{st} from Eq. (4.27) and the surface displacements from Eq. (4.26). The iterative procedure stops when the entries in $I_{C,st}^*$ and $I_{C,sl}^*$ do not change anymore, all the boundary elements belonging to $I_{C,st}^*$ are subject to tangential tractions less than μp , and all the boundary elements belonging to $I_{C,sl}^*$ have $u_{sl} < \Delta_T$.

Applications of this algorithm have been made in Paggi et al. (2014) for the identification of the evolution of the stick and slip contact domains from full stick to full slip for fractal rough surfaces subject to a given normal force and an increasing shearing displacement. Results pinpointed that the last boundary elements entering the contact domain are the first slipping, due to the low normal pressures acting on them. Therefore, the contact domain in stick condition is initially coincident with the normal contact domain and it progressively shrinks to zero, when all the boundary

elements slip. From the macroscopical point of view, the shearing force level corresponding to full slip is simply equal to the local friction coefficient multiplied by the applied total normal force, consistently with the Coulomb criterion adopted at the microscopical level.

The Finite Element Method

Variational Formulation

In this section, the variational formulation governing the problem of contact between two bodies across a rough interface is detailed. Since the mathematical formulation leading to the finite element method can easily handle in a consistent manner also adhesive (tensile) interactions at the interface, the most general scenario is herein examined. Therefore, starting from the strong differential form describing the mechanics of the continua and the problem of contact with adhesion along the interface, the corresponding weak form is derived. Afterward, different solution strategies and finite element discretization schemes are discussed, comparing methods based on the explicit discretization of roughness versus a recent method proposed in Paggi and Reinoso (2018) to analytically embed roughness into a special interface finite element.

Governing equations and strong form. Let two deformable bodies occupy the domains $\Omega_i \in \mathbb{R}^2$ ($i = 1, 2$) in the undeformed configuration defined by the reference system $Oxyz$. The two domains are separated by an interface Γ defined by the opposite boundaries Γ_i ($i = 1, 2$) of the two bodies, viz., $\Gamma = \bigcup_{i=1,2} \Gamma_i$, where contact or adhesive interactions take place. The whole boundary of the i th body, $\partial\Omega_i$, is therefore split into three parts: (i) a portion where displacements are imposed, i.e., the Dirichlet boundary $\partial\Omega_i^D$; (ii) a portion where tractions are specified, i.e., the Neumann boundary $\partial\Omega_i^N$; (iii) and the interface $\Gamma_i = \Gamma_i^C \cup \Gamma_i^A$ where specific boundary conditions have to be imposed to model contact on Γ_i^C or adhesion on Γ_i^A , see Fig. 4.9. The partition of Γ_i in Γ_i^C and Γ_i^A is not known a priori, but it is the result of the solution of the elastic problem.

In the most general case, we postulate the existence of a displacement field for each body, $\mathbf{u}_i = (u_i, v_i, w_i)^T$, that maps the transformation from the undeformed configuration to the deformed one, and vice versa. Such functions are thereby assumed to be continuous, invertible, and differentiable functions of the position vector $\mathbf{x} = (x, y, z)^T$ within each body. At the interface Γ^* , on the other hand, the configuration of the system is described by the relative displacement field $\Delta\mathbf{u}$, usually called *gap field* \mathbf{g} across the interface, which is mathematically defined as the projection of the relative displacement $\mathbf{u}_1 - \mathbf{u}_2$ onto the normal and tangential directions of the interface defined by the unit vectors \mathbf{n} , \mathbf{t}_1 and \mathbf{t}_2 , respectively. In components, the vector $\Delta\mathbf{u}$ collects the relative tangential displacements, $\Delta u_{t,1}$, $\Delta u_{t,2}$, and the relative normal displacement, Δu_n , i.e., $\Delta\mathbf{u} = (\Delta u_{t,1}, \Delta u_{t,2}, \Delta u_n)^T$. The total relative

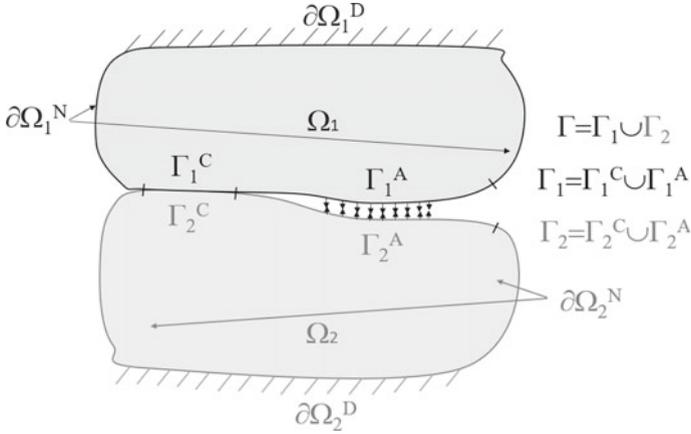


Fig. 4.9 Domains Ω_i ($i = 1, 2$), their Dirichlet ($\partial\Omega_i^D$) and Neumann ($\partial\Omega_i^N$) boundaries, and the interface $\Gamma = \Gamma_1 \cup \Gamma_2$ composed of an adhesive part, Γ_i^A , and a contact part, Γ_i^C . Adapted from Paggi and Reinoso (2018)

displacement in the tangential plane is given by $\Delta \mathbf{u}_t = (\Delta u_{t1}, \Delta u_{t2})^T$. In the sequel, we denote $\mathbf{g}_n = \Delta u_n$ and $\mathbf{g}_t = \Delta \mathbf{u}_t$, with its modulus $g_t = \Delta u_t = \sqrt{\Delta u_{t1}^2 + \Delta u_{t2}^2}$.

Inside each deformable material, the small deformation strain tensor ϵ_i ($i = 1, 2$) is introduced as customary, which is defined as the symmetric part of the displacement gradient: $\epsilon_i = \nabla^s \mathbf{u}_i$. In the sequel, the standard Voigt notation will be used and the strain tensor components will be collected in the vector $\epsilon_i = (\epsilon_{xx}, \epsilon_{yy}, \epsilon_{zz}, \gamma_{xy}, \gamma_{xz}, \gamma_{yz})_i^T$.

In the absence of body forces, the strong (differential) form of equilibrium for each body is provided by the linear momentum equation along with the Dirichlet and the Neumann boundary conditions on $\partial\Omega_i^D$ and $\partial\Omega_i^N$, respectively ($i = 1, 2$), equipped by unilateral contact conditions in the normal direction on Γ_C^* , Coulomb frictional conditions on the stick and slip partitions $\Gamma_{C,st}^*$ and $\Gamma_{C,sl}^*$ of Γ_C^* , and adhesion on Γ_A^* :

$$\nabla \cdot \boldsymbol{\sigma}_i = \mathbf{0} \quad \text{in } \Omega_i, \quad (4.28a)$$

$$\mathbf{u}_i = \bar{\mathbf{u}} \quad \text{on } \partial\Omega_i^D, \quad (4.28b)$$

$$\boldsymbol{\sigma}_i \cdot \mathbf{n} = \mathbf{T} \quad \text{on } \partial\Omega_i^N, \quad (4.28c)$$

$$\mathbf{g}_n = 0, \quad p_n < 0 \quad \text{on } \Gamma_C^*, \quad (4.28d)$$

$$\mathbf{g}_t = \mathbf{0}, \quad \|\mathbf{q}\| < \mu |p_n| \quad \text{on } \Gamma_{C,st}^*, \quad (4.28e)$$

$$\mathbf{q} = -\mu |p_n| \frac{\dot{\mathbf{g}}_t}{\|\dot{\mathbf{g}}_t\|} \quad \text{on } \Gamma_{C,sl}^*, \quad (4.28f)$$

$$\mathbf{g}_n > 0, \quad p_n = p_A > 0 \quad \text{on } \Gamma_A^*, \quad (4.28g)$$

where $\bar{\mathbf{u}}$ denotes the imposed displacement, \mathbf{T} the applied traction vector, $p_A(\mathbf{g}_n)$ is a function of the relative displacement $\Delta \mathbf{u}$, and \mathbf{q} is the shearing traction vector. Therefore, the nonlinearity of the problem stems from the fact that the contact and adhesive portions of the interface Γ^* are known only once the displacement field, solution of the problem, is known. As a consequence, the present problem can be ascribed to the family of the so-called moving boundary value problems and it requires an iterative solution scheme.

For its solution, the strong form has to be equipped by the constitutive equations for the bulk and for the interface. For the bulk, recalling standard thermodynamic arguments, general (linear or nonlinear) constitutive stress–strain relations can be postulated without any loss of generality for the i th material domain: $\boldsymbol{\sigma}_i := \partial_{\boldsymbol{\epsilon}_i} \Psi(\boldsymbol{\epsilon}_i)$ and $\mathbb{C}_i := \partial_{\boldsymbol{\epsilon}_i \boldsymbol{\epsilon}_i}^2 \Psi(\boldsymbol{\epsilon}_i)$, whereby $\Psi(\boldsymbol{\epsilon}_i)$ is the Helmholtz free-energy function for the body i , whereas its corresponding Cauchy stress tensor and the constitutive operator are, respectively, denoted by $\boldsymbol{\sigma}_i$ and \mathbb{C}_i . The two bodies are in general both deformable, but in the present setting it is also possible to consider one of them as rigid. This condition is of paramount interest for contact mechanics in the presence of two dissimilar linear elastic bodies. In such a case, it is possible to simplify the matter by replacing the bi-material system by a rigid body indenting a linear elastic material having composite elastic parameters, function of the Young's moduli E_i , and Poisson's ratios ν_i ($i = 1, 2$) of the two elastic materials, as previously detailed for the boundary element method, see also Barber (2010, 2018).

Regarding the interface, the constitutive response should be introduced by distinguishing between the normal and the tangential directions. In the normal direction, the contact condition imposes that the displacement field solution leads to a vanishing normal gap $\mathbf{g}_n = 0$ for the points in contact. Correspondingly, contact tractions are negative valued in the finite element method, while it is remarkable to note that the opposite convention was adopted in the boundary element method. For $\mathbf{g}_n > 0$, positive-valued adhesive tractions apply, and they can be, for instance, given by a relation dictated by an adhesion model inspired by the interatomic Lennard-Jones potential:

$$p_A = 24\varepsilon \left[\frac{\kappa^6}{(\mathbf{g}_n + \mathbf{g}_{n,0})^7} - 2 \frac{\kappa^{12}}{(\mathbf{g}_n + \mathbf{g}_{n,0})^{13}} \right], \quad (4.29)$$

where ε and κ are the model parameters and $\mathbf{g}_{n,0}$ is the molecular equilibrium distance. The parameter $\mathbf{g}_{n,0}$ is such that the condition $\mathbf{g}_n = 0$ leads to vanishing adhesive tractions and it correctly captures the transition from adhesion to contact. Other nonlinear adhesive models can be used in the present approach without any loss of generality, e.g., the surface potential derived from the interatomic Lennard-Jones potential, see Yu and Polycarpou (2004). To treat both contact and adhesive tractions in a unified framework, relaxing at the same time the unilateral contact constraint, a generalized penalty approach can be efficiently exploited (Paggi and Reinoso 2018). Physically speaking, a nonlinear spring model is inserted along the interface between the two bodies, where p_n is given by

$$p_n(\mathbf{g}_n) = \begin{cases} K \mathbf{g}_n, & \text{if } \mathbf{g}_n < 0, \\ p_A, & \text{if } \mathbf{g}_n \geq 0. \end{cases} \quad (4.30)$$

This formulation leads to a solution allowing for small compenetration, depending on the value of the penalty stiffness K . Hence, K should be high enough to reduce material penetration between adjacent continua and, at the same time, it should not be too high to cause ill-conditioning of the tangent operator resulting from the computational scheme. Following the pioneering work in Zavarise et al. (2019, 1992), the penalty stiffness K could be related to the normal contact stiffness predicted by semi-analytical micromechanical contact models, giving a physical ground for its estimation.

In the tangential direction, \mathbf{q} is the tangential traction vector which obeys the Coulomb friction law. To simplify its treatment into a computational scheme, a regularized dependency of \mathbf{q} upon $\dot{\mathbf{g}}_t$ is usually put forward, smoothing the sharp transition from the stick to the slip condition, see, e.g., Wriggers (2006, Sect. 5.2.3).

It is remarkable to note here that the treatment of the contact problem with friction in the finite element method is more general than the analogous treatment in the boundary element method, since it allows simulating any loading path in three dimensions. According to Eq. (4.28), the tangential traction vector \mathbf{q} changes sign depending on the velocity of sliding, $\dot{\mathbf{g}}_t$, which has to be computed using a time integration routine, usually based on the implicit Euler scheme. Moreover, the stick and slip portions of \mathbf{q} can be computed using a return mapping algorithm, in analogy with elastoplasticity. For small tangential displacements leading to a situation intermediate from full stick to full slip, before the onset of gross sliding, and for simple monotonic loading paths in one given tangential direction, the formulation can be simplified and rewritten in terms of the total relative displacement \mathbf{g}_t , instead of its time derivative. This leads to a penalty-like formulation similar to that used for Mode II cohesive zone models for fracture, see Paggi et al. (2006), Carpinteri et al. (2008), where tangential tractions are opposing to the relative sliding deformation and are specified as a closed-form equation in terms of \mathbf{g}_t .

Weak form. According to the principle of virtual work, the weak form associated to the strong form Eq. (4.28) with the penalty regularization in the normal and tangential directions reads

$$\begin{aligned} \Pi = & \int_{\Omega_1} \boldsymbol{\sigma}_1(\mathbf{u}_1)^T \boldsymbol{\epsilon}_1(\mathbf{v}_1) d\Omega + \int_{\Omega_2} \boldsymbol{\sigma}_2(\mathbf{u}_2)^T \boldsymbol{\epsilon}_2(\mathbf{v}_2) d\Omega \\ & - \int_{\partial\Omega_1^N} \mathbf{T}^T \mathbf{v}_1 d\partial\Omega - \int_{\partial\Omega_2^N} \mathbf{T}^T \mathbf{v}_2 d\partial\Omega \\ & - \int_{\Gamma_C^*} p_n(\Delta\mathbf{u}) \mathbf{g}_n(\Delta\mathbf{v}) d\Gamma - \int_{\Gamma_C^*} \mathbf{q}(\Delta\mathbf{u}) \mathbf{g}_t(\Delta\mathbf{v}) d\Gamma \\ & - \int_{\Gamma_A^*} p_A(\Delta\mathbf{u}) \mathbf{g}_n(\Delta\mathbf{v}) d\Gamma = 0, \end{aligned} \quad (4.31)$$

where \mathbf{v}_i is the test function (virtual displacement field), and $\mathbf{g}_n(\Delta\mathbf{v})$ and $\mathbf{g}_t(\Delta\mathbf{v})$ are the virtual normal and tangential relative displacements at the interface Γ^* . The test function in the i th body fulfills the condition $\mathbf{v}_i = \mathbf{0}$ on $\partial\Omega_i^D$ and the adhesive-contact condition on Γ^* . The displacement field \mathbf{u}_i solution of the weak form (4.31) is such that it corresponds to the minimum of Π for any choice of the test functions \mathbf{v}_i .

The numerical treatment of the weak form (4.31) within the finite element method requires the introduction of two different types of finite element discretization, one for the bulk, $\Omega_{i,h}$, and another for the interface, Γ_h^* , where the subscript h refers to the respective discretized geometrical feature. For the bulk, standard linear quadrilateral or triangular isoparametric finite elements can be invoked, see classical finite element textbooks (Zienkiewicz and Taylor 2000) for details. For the interface, different strategies can be exploited and they are discussed in the next sections.

Methods Based on the Explicit Discretization of Roughness

Methods based on the explicit discretization of roughness introduce special discretization schemes to model the rough interface topology.

The simplest method was proposed by Hyun et al. (2004) and Pei et al. (2005), who investigated the frictionless normal contact problem between a rough surface and a flat half-plane for elastic (Hyun et al. 2004) or elastoplastic (Pei et al. 2005) continua. In their approach, a three-dimensional mesh for a rough surface was constructed in two stages. First, a flat surface with nodes at each point on the square grid was considered. A local refinement technique was used to achieve a strong mesh gradation with very small elements near the surface and a coarser discretization in the bulk, to reduce the number of finite elements and, therefore, indirectly, the computation cost associated to the solution of the algebraic equations associated to the finite element method. In the second step, all nodes belonging to the surface were displaced to create the desired roughness.

Since only a small fraction of the nodes of the rough surface are in contact after the application of the load, it was convenient to assemble only nonvanishing contact contributions to the weak form. To do that, the conventional master/slave node-to-surface contact search was employed in Hyun et al. (2004), Pei et al. (2005). The contact search computes the value of the normal gap for each node of the rough surface with respect to the opposing master flat surface, considering the projection of the node position in the direction normal to the flat surface, see Wriggers (2006, Sect. 10.1). Depending on the sign of the gap function, nodes experiencing a compenetration are retained in the so-called *active set*. After inserting the finite element discretization, an explicit Newmark time-stepping algorithm was adopted in Hyun et al. (2004), Pei et al. (2005) to further reduce the computation cost. The advantage is that the explicit method leads to a set of uncoupled algebraic equations whose solution can be parallelized. The disadvantage is represented by the fact that an artificial damping has to be introduced to simulate quasi-static contact problems and the algorithm is not unconditionally stable, thus requiring small time steps.

More sophisticated discretization methods available in the literature exploit smooth interpolation schemes based on splines with randomly chosen heights to generate the asperities of the surface (Wriggers and Reinelt 2009), in conjunction with contact search algorithms and implicit Newton–Raphson incremental-iterative solution schemes for the solution of the nonlinear algebraic equations. Such an approach is unconditionally stable, thanks to the computation of the tangent stiffness matrix. However, sparse global stiffness matrices are obtained, which do not allow a straightforward parallelization of the linearized set of algebraic equations as for the explicit approach. Moreover, the introduction of smooth interpolation schemes can be beneficial for modeling wavy surfaces or spheres, as in the case of NURBS used in De Lorenzis and Wriggers (2013) to depict sinusoidal wavy profiles, but it is cumbersome for the description of multiscale roughness features over multiple wavelengths. Due to the complexity of the implicit solution scheme and the explicit discretization of the rough geometry, applications have been confined to the solution of small-scale problems on representative surface elements. This is, for instance, the case of two-scale finite element simulations as proposed in Wriggers and Reinelt (2009), where a reduced fine-scale model of a wavy surface was used for the computation of a microscopically constitutive law to be passed to standard smooth macroscale finite element computations.

Interface Finite Elements with Embedded Profile for Joint Roughness

As an alternative approach to the explicit discretization of roughness, which is computationally demanding and poses severe limitations for the use of the finite element method, a new interface finite element with analytically embedded roughness has been proposed in Paggi and Reinoso (2018).

The method assumes that the boundaries Γ_i in Fig. 4.9 are *nominally flat but microscopically embedding rough profiles*. Γ_1 and Γ_2 can be one the negative of the other, as in the case of an interface originated by fracture, or different from each other, as for two bodies coming into contact, without any restriction.

It is convenient to introduce for the i th rough profile Γ_i its *smoother line* $\bar{h}_i(\xi_i)$ parallel to the average line of the profile and with datum set in correspondence of its deepest valley (see Fig. 4.10). A point along the curve $\bar{h}_i(\xi_i)$ is identified by a value of the curvilinear coordinate $\xi_i = \xi_i(x, y)$, which establishes a one-to-one correspondence with the coordinates of the same point in the global reference system Oxy . It also associates the tangential and normal unit vectors $\mathbf{t}_i(\xi_i)$ and $\mathbf{n}_i(\xi_i)$ to $\bar{h}_i(\xi_i)$, to identify the normal and the tangential directions at any point along the smoothed line $\bar{h}_i(\xi_i)$, with \mathbf{n}_i pointing outward from the domain Ω_i . Due to the assumption that the two non-conformal profiles are microscopically rough but nominally flat, the two smoother lines $\bar{h}_i(\xi_i)$ are parallel to each other and therefore $\mathbf{n}_1(\xi) = -\mathbf{n}_2(\xi)$ and $\mathbf{t}_1(\xi) = -\mathbf{t}_2(\xi)$, $\forall \xi$.

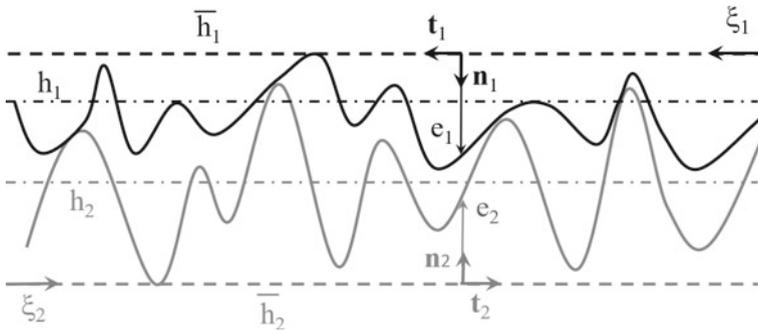


Fig. 4.10 Parametrization of two microscopically rough profiles composing an interface Γ . Adapted from Paggi and Reinoso (2018)

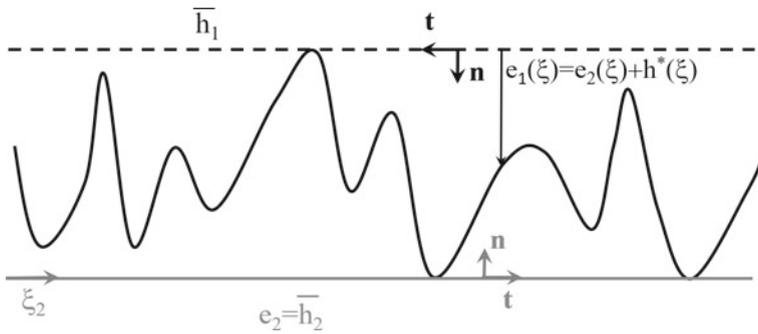


Fig. 4.11 Composite topography of the interface Γ . Adapted from Paggi and Reinoso (2018)

The actual elevation of the rough profile measured from $\bar{h}_i(\xi_i)$ is finally described by the *roughness function* $h_i(\xi_i)$. Therefore, the i th boundary Γ_i is parametrized such that its actual elevation $e_i(\xi_i)$ in the curvilinear setting is given by $e_i(\xi_i) = \bar{h}_i(\xi_i) + h_i(\xi_i)$.

It is in general convenient to exploit the concept of *composite topography* Γ^* of the interface Γ , as also routinely done for the boundary element method. The contact problem between two linear elastic materials with dissimilar rough boundaries is therefore simplified into the contact problem between an infinitely stiff indenter with such a composite topography taken as boundary, and a linear elastic half-plane with composite elastic parameters function of those of the parent elastic bodies. This transformation also allows the study of the contact problem involving a rigid indenter of arbitrary profile (spherical, conical, etc.) and a half-plane.

In the context of the present method, the composite topography is mathematically represented by a flat line, $e_2 = \bar{h}_2(\xi)$, and a profile with elevation $e_1(\xi) = \bar{h}_2(\xi) + h^*(\xi)$, where $h^*(\xi) = \max_{\xi} [h_1(\xi) + h_2(\xi)] - [h_1(\xi) + h_2(\xi)]$ (see Fig. 4.11).

This transformation does not apply for two elastoplastic or viscoelastic materials, while it does apply for any nonlinear interface constitutive model provided that the two materials are linear elastic. Elastoplastic or viscoelastic contact problems can be

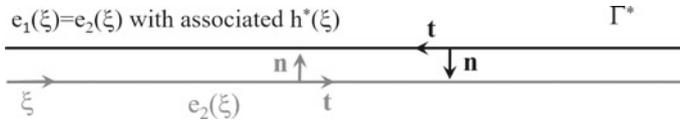


Fig. 4.12 Zero-thickness interface representation Γ^* of the composite topography. Adapted from Paggi and Reinoso (2018)

still simulated using the proposed approach, provide that the attention is restricted to the contact problem between a rigid indenter with rough boundary Γ_1 and a flat half-plane Γ_2 with any prescribed material constitutive relation. In such a case, in fact, the composite topography simply reduces to Γ_1 and the original indenter geometry is kept unchanged.

After this transformation, a *zero-thickness interface model* for Γ^* is introduced and defined by the two initially coincident but distinct (not-joined) flat lines described by the function $e_2(\xi)$, plus the associated function $h^*(\xi)$. This composite topography has also unique tangential and normal unit vectors \mathbf{t} and \mathbf{n} , as previously discussed, see Fig. 4.12.

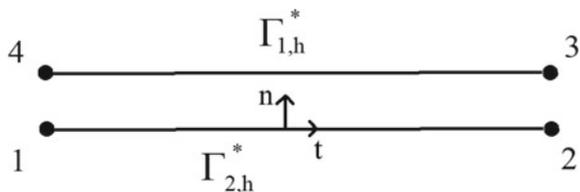
In this framework, the normal gap g_n of the composite topography, which represents the actual physical separation between the composite topography and the smooth curve e_2 after deformation, is given by $g_n = \Delta u_n + h^*$, since material 1 acts as a rigid indenter with a rough profile h^* . Based on the value of g_n , the portion of the interface in contact, $\Gamma^C = \Gamma_1^C \cup \Gamma_2^C$, is identified by the condition $g_n = 0$. On the other hand, the portion subject to adhesion, $\Gamma_A = \Gamma_1^A \cup \Gamma_2^A$, presents a positive-valued normal gap $g_n > 0$. A negative-valued normal gap is not admissible so far by definition, since it would imply compenetration between the bodies.

At the interface, a conforming finite element discretization for the continua can be simply adopted. Consequently, a special interface finite element with embedded profile for joint roughness (MPJR interface finite element) whose kinematics departs from the formulation of interface elements used in nonlinear fracture mechanics for cohesive crack growth (Ortiz and Pandolfi 1999; Paggi and Wriggers 2011, 2012; Reinoso and Paggi 2014; Paggi and Reinoso 2015) has been proposed in Paggi and Reinoso (2018) to be inserted along the interface.

In 2D problems, the interface element is defined by nodes 1 and 2, which belong to $\Gamma_{2,h}^*$, and by nodes 3 and 4, which belong to $\Gamma_{1,h}^*$, see Fig. 4.13.

For frictionless normal contact problems, the contribution of the interface to the weak form is provided by the integral $\int_{\Gamma^*} p(g_n)g_n d\Gamma$ in Eq. (4.31), which can be

Fig. 4.13 Sketch of the interface finite element topology. From Paggi and Reinoso (2018)



computed as the sum of the contributions of the whole interface elements, invoking the property of compactness of isoparametric shape functions:

$$\int_{\Gamma^*} p(\mathbf{g}_n) \mathbf{g}_n d\Gamma \cong \int_{\Gamma_h^*} p(\mathbf{g}_n) \mathbf{g}_n d\Gamma = \mathcal{A}_{e=1}^{n_{el}} \left\{ \int_{\Gamma_e^*} p(\mathbf{g}_n) \mathbf{g}_n d\Gamma \right\}, \quad (4.32)$$

where the subscript e refers to the e th interface element $e = 1, \dots, n_{el}$, and \mathcal{A} symbolically denotes an assembly operator.

The interface integral is herein computed exactly, by using the two-point Newton–Cotes quadrature formula which implies the sampling of the integrand at the nodes 1 and 2 (or, equivalently, at nodes 3 and 4):

$$\int_{\Gamma_e^*} p(\mathbf{g}_n) \mathbf{g}_n d\Gamma = \sum_{j=1,2} p_i(\mathbf{g}_n) \mathbf{g}_{n,i} \det J, \quad (4.33)$$

where $\det J$ is the standard determinant of the Jacobian of the transformation that maps the geometry of the interface element from its global reference frame to the natural reference system.

To evaluate the normal gap \mathbf{g}_n at any point inside the interface element, the nodal displacement vector $\mathbf{d} = (u_1, v_1, \dots, u_4, v_4)^T$ has to be introduced, which collects the displacements u and v of the four interface finite element nodes. The relative displacement $\Delta \mathbf{u}$ for the nodes 1–4 and 2–3 is then computed by applying a matrix operator \mathbf{L} which makes the difference between the displacements of nodes 1 and 4, and between nodes 2 and 3. The relative displacement within the interface finite element is then given by the linear interpolation of the corresponding nodal values, performed by the multiplication with the matrix \mathbf{N} which collects the shape functions at the element level. Finally, the tangential and the normal gaps are determined by the multiplication with the rotation matrix \mathbf{R} defined by the components of the unit vectors \mathbf{t} and \mathbf{n} . In formulae, we have

$$\Delta \mathbf{u} = \mathbf{RNLd}, \quad (4.34)$$

where the operators present the following matrix form:

$$\mathbf{L} = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \end{bmatrix}, \quad (4.35a)$$

$$\mathbf{N} = \begin{bmatrix} N_1 & 0 & N_2 & 0 \\ 0 & N_1 & 0 & N_2 \end{bmatrix}, \quad (4.35b)$$

$$\mathbf{R} = \begin{bmatrix} t_x & t_y \\ n_x & n_y \end{bmatrix}, \quad (4.35c)$$

where n_x , n_y , t_x , and t_y are the components of the unit vectors \mathbf{n} and \mathbf{t} along the x and y directions, respectively.

Once $\Delta \mathbf{u} = (\Delta u_t, \Delta u_n)^T$ is determined, the actual normal gap is given by a correction to Δu_n to account for the embedded profile that models the non-planarity of the rigid indenter Γ^* , i.e., $g_n = \Delta u_n + h^*$. The normal gap is used to compute the normal traction p_n according to Eq. (4.30). Similarly, for further extensions to adhesive-contact problems with friction in the tangential direction, a relationship between the shearing traction p_t and the relative sliding displacement \mathbf{g}_t , or its velocity, should be introduced, in analogy with the normal problem.

Due to the intrinsic nonlinearity, a full Newton–Raphson iterative and incremental scheme was adopted in Paggi and Reinoso (2018) to solve the implicit nonlinear algebraic system of equations resulting from the finite element discretization:

$$\mathbf{K}^{(k)} \Delta \mathbf{d}^{(k)} = -\mathbf{R}^{(k)}, \quad (4.36a)$$

$$\mathbf{d}^{(k+1)} = \mathbf{d}^{(k)} + \Delta \mathbf{d}^{(k)}, \quad (4.36b)$$

where the superscript k denotes the iteration inside the Newton–Raphson loop. The residual vector $\mathbf{R}_e^{(k)}$ and the tangent stiffness matrix $\mathbf{K}_e^{(k)}$ associated with the e th interface finite element, to be assembled to the global residual vector \mathbf{R} and to the global stiffness matrix \mathbf{K} , are

$$\mathbf{R}_e^{(k)} = \int_{\Gamma_e^*} \mathbf{L}^T \mathbf{N}^T \mathbf{R}^T \mathbf{p} \, d\Gamma, \quad (4.37a)$$

$$\mathbf{K}_e^{(k)} = \int_{\Gamma_e^*} \mathbf{L}^T \mathbf{N}^T \mathbf{R}^T \mathbb{C} \mathbf{R} \mathbf{N} \mathbf{L} \, d\Gamma, \quad (4.37b)$$

where $\mathbf{p} = (p_t, p_n)^T = (0, p_n)^T$ for frictionless normal contact problems, and \mathbb{C} is the linearized interface constitutive matrix:

$$\mathbb{C} = \begin{bmatrix} \frac{\partial p_t}{\partial g_t} & \frac{\partial p_t}{\partial g_n} \\ \frac{\partial p_n}{\partial g_t} & \frac{\partial p_n}{\partial g_n} \end{bmatrix}, \quad (4.38)$$

where, again for the frictionless normal contact problem, one needs to specify only $\partial p_n / \partial g_n$ depending on the sign of the normal gap, distinguishing between the penalty relation in compression or the adhesive relation in tension.

In principle, the MPJR interface element contributions to the stiffness matrix could be added only for the elements in contact, using a conventional contact search algorithm. However, in case of adhesive contact with long-range adhesive effects, all the interface elements can contribute to tension or compression and have to be assembled in any case. Due to the simplicity in modeling roughness according to this approach, which is embedded in the computation of the normal gap, the cost of

assembling the whole set of interface finite elements is much less important than for the methods relying on an explicit discretization of roughness. Therefore, contact search algorithms can be skipped, further simplifying the numerical implementation and the robustness of the computational method.

Comparison Between Different Approaches

As a benchmark problem to compare the different approaches presented in Sects. 4.3.2 and 4.3.3, we simulate the bidimensional frictionless normal contact problem without adhesion between a rigid cylinder indenting a half-plane. For comparison purposes, we recall the Hertzian analytical solution, which is available to assess the model accuracy.

The standard procedure for solving this problem within the finite element method requires modeling of the circular cross section of the cylinder and the use of a contact formulation to enforce the unilateral contact constraint along the interface between the cylinder and the half-plane. For that, among the possible numerical strategies, the penalty approach, the Lagrange multiplier method, and the mortar method are among the most popular formulations, see Wriggers (2006). In spite of the simplicity of this nonconforming contact problem, it is well known that all such methods require very fine meshes to resolve the contact area and the contact traction distribution, especially near the edges of the contact strip. This is primarily due to the fact that a C^1 linear finite element interpolation scheme is not sufficiently accurate to describe the circular shape of the cylinder. To overcome this drawback and increase the accuracy in the boundary element method and in the finite element method relying on the explicit discretization of the interface geometry, adaptive mesh refinement was proposed by Oysu (2007), see Fig. 4.14.

Alternatively, the NURBS finite element technique, which adopts shape functions with a very high regularity and smoothness to approximate curvilinear shapes, can be adopted to explicitly discretize the interface geometry, since it has been demonstrated in Dimitri et al. (2014) to provide the best accuracy over other discretization techniques. In spite of that, NURBS still presents problems in capturing the analytical Hertzian contact solution for the frictionless normal contact problem between a sphere and a half-plane, though a very fine mesh was used in Dimitri et al. (2014), see Fig. 4.15.

In the MPJR interface finite element, instead of modeling the geometry of the circular cross section, the non-planarity of the interface is simply embedded in the interface finite element with its exact analytical function. The actual circular shape of the boundary Γ_1 is therefore given by the composite topography of the interface profile: $e_1(x) = \bar{h}_2 + h^*(x)$, where $\bar{h}_2 = x_l$ and $h^*(x) = R - \sqrt{R^2 - x^2}$. This strategy resolves the issues related to the accuracy of finite element interpolation schemes for the bulk and the interface, which can now have low-order linear shape functions. In this context, the geometry of the cylinder of radius R occupying the domain Ω_1 can be simply replaced by a rectangular block of lateral size x_l and thickness $x_l/20$, while

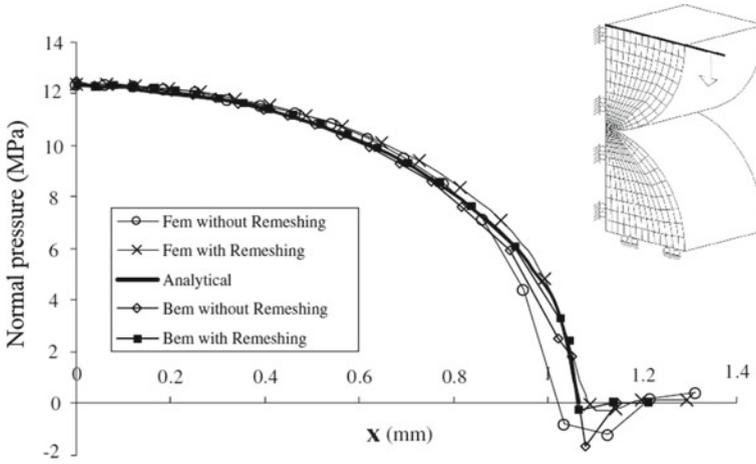
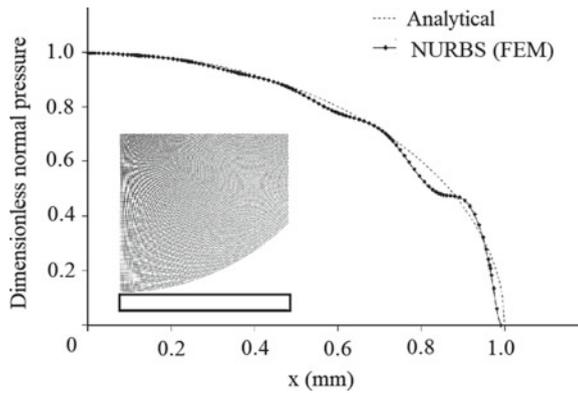


Fig. 4.14 Problems in resolving contact tractions in Hertzian normal contact problems arising from the boundary element method and the finite element method with or without mesh refinement. Adapted from Oysu (2007)

Fig. 4.15 Problems in resolving contact tractions in Hertzian normal contact problems arising from the finite element method with NURBS discretization of the interface geometry. Adapted from Dimitri et al. (2014)



the half-plane occupying the domain Ω_2 can be modeled as a plane strain domain with size x_l , see Fig. 4.16. Since the indenter is rigid, the equivalent model is exact, because there is no error resulting from the deformation associated to the different geometries of body 1. The low-order interpolation scheme used for the interface, on the other hand, is enhanced by the exact analytical representation of the circular indenter profile.

To achieve the condition of a rigid cylinder pressed onto an elastically deformable half-plane, E_1 can be simply set equal to $1000E_2$, where the subscripts 1 and 2 identify the rigid (indenter) and deformable (half-plane) bodies, respectively. Regarding the ratio between the cylinder radius and the lateral size of the half-plane, R/x_l , two cases were examined in Paggi and Reinoso (2018): (i) $R/x_l = 100$, which corresponds

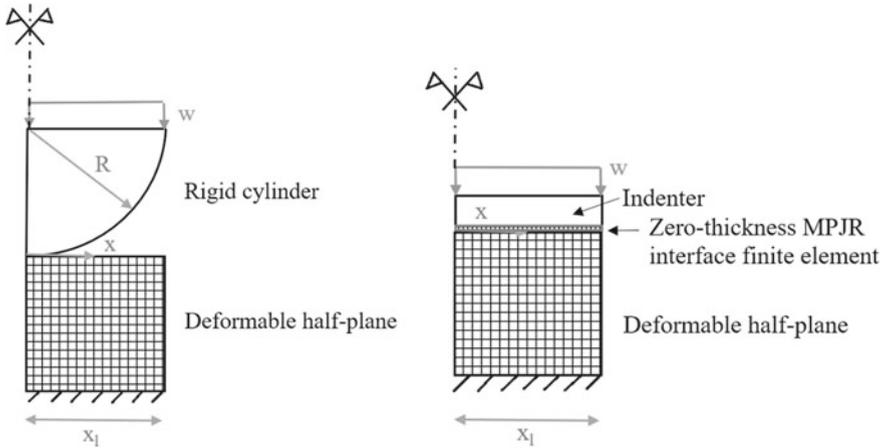


Fig. 4.16 The actual geometry of the Hertz contact problem (left), and its finite element model based on the present variational approach with embedded roughness (right), which incorporates the analytical expression of the curved interface profile into the MPJR interface finite elements instead of explicitly discretizing the interface geometry

to a slightly nonplanar interface; (ii) $R/x_l = 1$, which corresponds to a significant deviation of the interface from the non-planarity. For both cases, uniform meshes for the domains Ω_1 and Ω_2 were used, employing four nodes linear finite elements for the bulk and the proposed MPJR interface elements with embedded roughness for the interface. The whole interface was discretized in the horizontal direction by only $n_{el} = 100$ finite elements, which is much less than what was used in Fig. 4.15 for NURBS, and without adopting any mesh refinement.

Dirichlet boundary conditions are represented in this test by imposed downward vertical displacements w on the topmost side of the domain Ω_1 , monotonically increasing with a pseudo-time variable to simulate the quasi-static normal contact problem; a fully restrained lower side of the domain Ω_2 ; and a symmetry condition on the vertical size of domains Ω_1 and Ω_2 to account for the symmetry in the geometry and in the loading (Fig. 4.16).

Numerical predictions are provided in terms of the dimensionless normal contact pressure, p/E , versus the dimensionless position along the interface, x/R . The contact pressure p is given by $p = -p_n$, and therefore it is positive valued on the portion of the interface Γ_C in contact and it must be zero elsewhere, since adhesion is not considered here. The penalty stiffness K is set $K = 10E_1/x_l$, to model a very stiff interface and avoid material penetration.

Results from the current simulations are shown in Fig. 4.17 for the case $R/x_l = 100$ and $R/x_l = 1$, considering nine increasing values for the imposed far-field displacement w . Analytical Hertzian results, corresponding to the same contact radii, are also superimposed by circles. As can be observed in these graphs, the agreement between the present model predictions and theory is excellent, also for the case

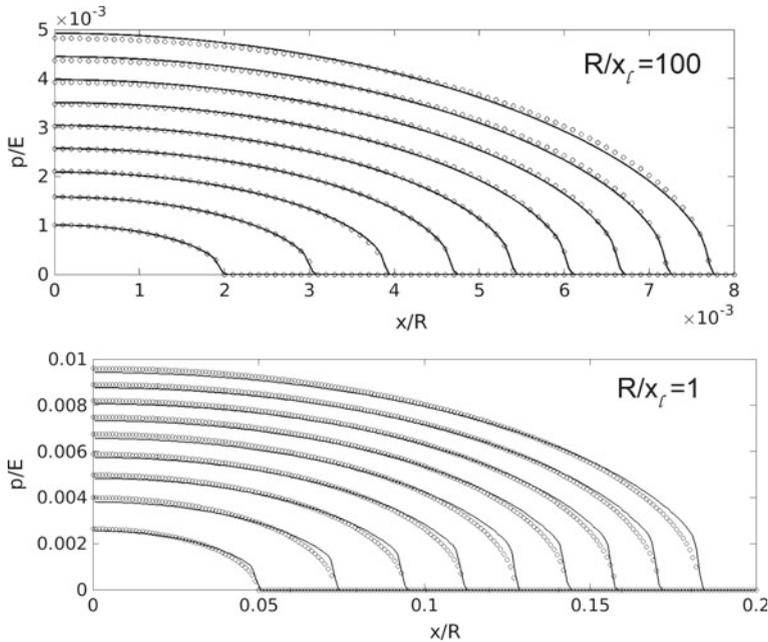


Fig. 4.17 Dimensionless contact pressure along the interface for different imposed far-field displacements and two different values of R/x_l . E , R , and x_l denote, respectively, the composite Young's modulus, the cylinder radius, and the lateral size of the domain, respectively. The analytical Hertzian solution is superimposed with circles. Adapted from Paggi and Reinoso (2018)

$R/x_l = 1$, which is indeed very challenging from the computational point of view due to the significant non-planarity of the interface.

Conclusive Remarks

The solution of the contact problem between bodies separated by rough boundaries is very important in many tribological applications and it is a challenging research topic due to the multiscale features of roughness that span over multiple length scales. So far, especially in relation to the linear elastic frictionless normal contact problem, the boundary element method has been proved to be very efficient from the computational point of view and preferable over the finite element method. It has been used to validate semi-analytical approaches based on asperities or other prominent contact theories, see, e.g., Mueser et al. (2017), inspiring also further developments in the framework of multi-asperity contact models including elastic interaction effects. Moreover, it has been applied to identify valuable trends on the effect of roughness on the emerging contact response, such as the real contact area

Table 4.1 Synopsis of the computational methods reviewed in this chapter

Approach	Roughness discretization	Continua	Interface	Pros	Cons
BEM	Exact representation	Linear elasticity or elastoplasticity	Recently extended to adhesion	Fast and accurate	Nonlinearities and multi-field problems are difficult to be addressed
Multi-asperity BEM	Only asperities	Linear elasticity	Resolution-dependent discretization	Very fast	As BEM, plus the limits of defining the asperities
FEM with explicit discretization of the interface geometry (NURBS, splines, etc.)	Modeling multiscale roughness is problematic	No restrictions	No restrictions	Ideal for nonlinear and multi-field problems	Limits in the discretization of roughness; computationally demanding
FEM-MPJR with embedded roughness	Exact	No restrictions	No restrictions	Ideal for nonlinear and multi-field problems	Faster and simpler than FEM, but still computationally demanding as compared to BEM

versus load relation (Paggi and Ciavarella 2010), the normal contact stiffness versus load dependency (Paggi and Barber 2011), as well as the evolution of the free volume between rough surfaces during the progress of contact (Paggi and He 2015).

If the research on the frictionless normal contact problem between linear elastic materials has seen a significant impulse during the last decades, the contact problem in the presence of friction still presents open issues and aspects deserving investigation. So far, the evolution of the stick and slip contact domains in the case of a monotonically increasing tangential displacement applied by keeping constant the normal load has been investigated using the boundary element method in Paggi et al. (2014). Another important topic regards the effect of roughness on hysteretic energy dissipation caused by cyclic tangential loading paths, see, e.g., Borri-Brunetto et al. (2006), Barber et al. (2011).

Finally, there is a range of contact problems involving nonlinear constitutive relations, nonlinear geometric (finite elasticity) effects, and coupled multi-field problems that have been only marginally challenged so far. For instance, the problem of electromechanical interaction between excitable deformable cells in finite elasticity investigated in Lenarda et al. (2018) is an exemplary problem requiring the exten-

sion of the methods of contact mechanics to soft biological matter. For this class of problems, the constitutive nonlinearities for the interface and the continua suggest passing to the finite element method. In this regard, further advancements in the interface finite element discretization are indeed required for an efficient treatment of contact problems with roughness features covering multiple length scales. The novel approach based on the interface finite element with embedded roughness proposed in Paggi and Reinoso (2018) opens new perspectives in this direction.

As a guideline for new researchers entering the field, and also for experienced researchers willing to explore future directions of research in contact mechanics between rough surfaces, a synopsis of the computational methods reviewed in this chapter is provided in Table 4.1, emphasizing the advantages and the disadvantages of each approach.

Acknowledgements This chapter is derived in part from: an article published in the International Journal of Solids and Structures (Elsevier), available online June 16, 2015, doi: <http://doi.org/10.1016/j.ijsolstr.2015.06.005>; an article published in Mechanics of Advanced Materials and Structures (Taylor & Francis), available online November 5, 2018, doi: <http://doi.org/10.1080/15376494.2018.1525454>.

References

- Afferrante, L., Carbone, G., & Demelio, G. (2012). Interacting and coalescing Hertzian asperities: A new multisasperity contact model. *Wear*, 278–279, 28–33.
- Andersson, T. (1981). The boundary element method applied to two-dimensional contact problems with friction. *Boundary Element Methods*, 3, 239–258.
- Barber, J. R. (1974). Determining the contact area in elastic-indentation problems. *Journal of Strain Analysis*, 9, 230–232.
- Barber, J. R. (2003). Bounds on the electrical resistance between contacting elastic rough bodies. *Proceedings of the Royal Society of London, Series A*, 459, 53–66.
- Barber, J. R. (2010). *Elasticity* (3rd ed.). Dordrecht: Springer.
- Barber, J. R. (2018). *Contact mechanics*. Springer International Publishing.
- Barber, J. R., Davies, M., & Hills, D. A. (2011). Frictional elastic contact with periodic loading. *International Journal of Solids and Structures*, 48, 2041–2047.
- Batrouni, G. G., Hansen, A., & Schmittbuhl, J. (2002). Elastic response of rough surfaces in partial contact. *Europhysics Letters*, 60, 724–730.
- Beale, E. M. L. (1955). On minimizing a convex function subject to linear inequalities. *Journal of the Royal Statistical Society, Series B*, 173–184.
- Bemporad, A. (2014). A quadratic programming algorithm based on nonnegative least squares with applications to embedded model predictive control. *IEEE Transactions on Automatic Control*.
- Bemporad, A., & Paggi, M. (2015). Optimization algorithms for the solution of the frictionless normal contact between rough surfaces. *International Journal of Solids and Structures*, 69–70, 94–105.
- Borri, C., & Paggi, M. (2015). Topological characterization of antireflective and hydrophobic rough surfaces: Are random process theory and fractal modeling applicable? *Journal of Physics D: Applied Physics*, 48, 045301.
- Borri, C., & Paggi, M. (2016). Topology simulation and contact mechanics of bifractal rough surfaces. *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, 230, 1345–1358.

- Borri-Brunetto, M., Carpinteri, A., & Chiaia, B. (1999). Scaling phenomena due to fractal contact in concrete and rock fractures. *International Journal of Fracture*, *95*, 221–238.
- Borri-Brunetto, M., Chiaia, B., & Ciavarella, M. (2001). Incipient sliding of rough surfaces in contact: A multiscale numerical analysis. *Computer Methods in Applied Mechanics and Engineering*, *190*, 6053–6073.
- Borri-Brunetto, M., Carpinteri, A., Invernizzi, S., & Paggi, M. (2006). Micro-slip of rough surfaces under cyclic tangential loading. In P. Wriggers & U. Nackenhorst (Eds.), *Analysis and simulation of contact problems*. Lecture notes in applied and computational mechanics (Vol. 27, pp. 191–200). Berlin, Heidelberg: Springer.
- Boyd, S., Parikh, N., Chu, E., Peleato, B., & Eckstein, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, *3*, 1–122.
- Campana, C., Persson, B. N. J., & Mueser, M. H. (2001). Transverse and normal interfacial stiffness of solids with randomly rough surfaces. *Journal of Physics: Condensed Matter*, *23*, 085001.
- Carbone, G., & Bottiglione, F. (2008). Asperity contact theories: Do they predict linearity between contact area and load? *Journal of the Mechanics and Physics of Solids*, *56*, 2555–2572.
- Carbone, G., & Putignano, C. (2013). A novel methodology to predict sliding and rolling friction of viscoelastic materials: Theory and experiments. *Journal of the Mechanics and Physics of Solids*, *61*, 1822–1834.
- Carpinteri, A., & Paggi, M. (2005). Size-scale effects on the friction coefficient. *International Journal of Solids and Structures*, *42*, 2901–2910.
- Carpinteri, A., & Paggi, M. (2009). A fractal interpretation of size-scale effects on strength, friction and fracture energy of faults. *Chaos, Solitons & Fractals*, *39*, 540–546.
- Carpinteri, A., Paggi, M., & Zavarise, G. (2008). The effect of contact on the decohesion of laminated beams with multiple microcracks. *International Journal of Solids and Structures*, *45*, 129–143.
- Chang, W. R., Etsion, I., & Bogy, D. B. (1987). An elastic-plastic model for the contact of rough surfaces. *Journal of Tribology*, *109*, 257–263.
- Chen, W. W., Wang, Q. J., Wang, F., Keer, L. M., & Cao, J. (2008). Three-dimensional repeated elasto-plastic point contacts, rolling, and sliding. *Journal of Applied Mechanics*, *75*, 021021.
- Ciavarella, M., Demelio, G., Barber, J. R., & Jang, Y. H. (2000). Linear elastic contact of the Weierstrass profile. *Proceedings of the Royal Society of London, Series A*, *456*, 387–405.
- Ciavarella, M., Murolo, G., Demelio, G., & Barber, J. R. (2004). Elastic contact stiffness and contact resistance for the Weierstrass profile. *Journal of the Mechanics and Physics of Solids*, *52*, 1247–1265.
- Ciavarella, M., Delfino, V., & Demelio, G. (2006). A “re-vitalized” Greenwood and Williamson model of elastic contact between fractal surfaces. *Journal of the Mechanics and Physics of Solids*, *54*, 2569–2591.
- Ciavarella, M., Dibello, S., & Demelio, G. (2008a). Conductance of rough random profiles. *International Journal of Solids and Structures*, *45*, 879–893.
- Ciavarella, M., Greenwood, J. A., & Paggi, M. (2008b). Inclusion of “interaction” in the Greenwood and Williamson contact theory. *Wear*, *265*, 729–734.
- Cottle, R. W., Pang, J.-S., & Stone, R. E. (1992). *The linear complementarity problem*. Academic Press.
- Dimitri, R., De Lorenzis, L., Scott, M. A., Wriggers, P., Taylor, R. L., & Zavarise, G. (2014). Isogeometric large deformation frictionless contact using T-splines. *Computer Methods in Applied Mechanics and Engineering*, *269*, 394–414.
- Fletcher, R. (1971). A general quadratic programming algorithm. *IMA Journal of Applied Mathematics*, *7*, 76–91.
- Francis, H. A. (1983). The accuracy of plane strain models for the elastic contact of three-dimensional rough surfaces. *Wear*, *85*, 239–256.
- Goldfarb, D., & Idnani, A. (1983). A numerically stable dual method for solving strictly convex quadratic programs. *Mathematical Programming*, *27*, 1–33.
- Greenwood, J. A. (2006). A simplified elliptic model of rough surface contact. *Wear*, *261*, 191–200.

- Greenwood, J. A., & Williamson, J. B. P. (1966). Contact of nominally flat surfaces. *Proceedings of the Royal Society of London, Series A*, 295, 300–319.
- Greenwood, J. A., & Wu, J. J. (2001). Surface roughness and contact: An apology. *Meccanica*, 36, 617–630.
- Hendriks, C. P., & Visscher, M. (1995). Accurate real area of contact measurements on polyurethane. *ASME Journal of Tribology*, 117, 607–611.
- Hestenes, M. R. (1980). *Conjugate direction methods in optimization* (Chaps. 2 and 3). New York: Springer.
- Hyun, S., & Robbins, M. O. (2007). Elastic contact between rough surfaces: Effect of roughness at large and small wavelengths. *Tribology International*, 40, 1413.
- Hyun, S., Pei, L., Molinari, J.-F., & Robbins, M. O. (2004). Finite-element analysis of contact between elastic self-affine surfaces. *Physical Review E*, 70, 026117.
- Johnson, K. L. (1985). *Contact mechanics*. Cambridge, UK: Cambridge University Press.
- Kalker, J. J., & van Randen, Y. A. (1972). A minimum principle for frictionless elastic contact with application to non hertzian problems. *Journal of Engineering Mathematics*, 6, 193–206.
- Karpenko, Y. A., & Akay, A. (2001). A numerical model of friction between rough surfaces. *Tribology International*, 34, 531–545.
- Kogut, L., & Etsion, I. (2002). Elastic-plastic contact analysis of a sphere and a rigid flat. *Journal of Applied Mechanics*, 69, 657–662.
- Kubo, A., Okamoto, T., & Kurokawa, N. (1981). Contact stress between rollers with surface irregularity. *Journal of Tribology*, 116, 492–498.
- Lawson, C., & Hanson, R. (1974). *Solving least squares problems* (Vol. 161, Chap. 24). SIAM.
- Lenarda, P., Gizzi, A., & Paggi, M. (2018). A modeling framework for electro-mechanical interaction between excitable deformable cells. *European Journal of Mechanics—A/Solids*, 72, 374–392.
- Leroux, J., Fulleringer, B., & Nélias, D. (2010). Contact analysis in presence of spherical inhomogeneities within a half-space. *International Journal of Solids and Structures*, 47, 3034–3049.
- Li, J., & Berger, E. J. (2003). A semi-analytical approach to three-dimensional normal contact problems with friction. *Computational Mechanics*, 30, 310–322.
- Li, Q., Argatov, I., & Popov, V. (2018). Onset of detachment in adhesive contact of an elastic half-space and flat-ended punches with non-circular shape: Analytic estimates and comparison with numeric analysis. *Journal of Physics D: Applied Physics*.
- De Lorenzis, L., & Wriggers, P. (2013). Computational homogenization of rubber friction on rough rigid surfaces. *Computational Materials Science*, 77, 264–280.
- Love, H. A. E. (1999). The stress produced in a semi-infinite solid by pressure on part of the boundary. *Philosophical Transactions of the Royal Society of London*, 228, 377–420.
- Lubrecht, A. A., & Ioannides, E. (1991). A fast solution of the dry contact problem and the associated sub-surface stress field, using multilevel techniques. *ASME Journal of Tribology*, 113, 128–133.
- Majumdar, A., & Bhushan, B. (1990). Role of fractal geometry in roughness characterization and contact mechanics of surfaces. *ASME Journal of Tribology*, 112, 205–216.
- Man, K. W. (1994). *Contact mechanics using boundary elements, topics in engineering* (Vol. 22). Boston: Southampton.
- McCool, J. I. (1986). Comparison of models for the contact of rough surfaces. *Wear*, 107, 37–60.
- Mueser, M. H., Dapp, W. B., Bugnicourt, R., Sainsot, P., Lesaffre, N., Lubrecht, T. A., et al. (2017). Meeting the contact-mechanics challenge. *Tribology Letters*, 65, 118.
- Nélias, D., Boucly, V., & Brunet, M. (2006). Elastic-plastic contact between rough surfaces: Proposal for a wear or running-in model. *Journal of Tribology*, 128, 236–244.
- Nesterov, Y. (1983). A method of solving a convex programming problem with convergence rate $O(1/k^2)$. *Soviet Mathematics Doklady*, 27, 372–376.
- Nogi, T., & Kato, T. (1997). Influence of a hard surface layer on the limit of elastic contact—Part I: Analysis using a real surface model. *Journal of Tribology*, 110, 376–493.
- Nosonovsky, M., & Bhushan, B. (2005). Roughness optimization for biomimetic superhydrophobic surfaces. *Microsystem Technologies*, 11, 535–376.

- Nowell, D. D., & Hills, D. A. (1989). Hertzian contact of ground surfaces. *ASME Journal of Tribology*, *111*, 175–179.
- O’Callaghan, P. W., & Probert, S. D. (1970). Real area of contact between a rough surface and a softer optically flat surface. *Journal of Mechanical Engineering Science*, *11*, 259–267.
- Ortiz, M., & Pandolfi, A. (1999). Finite deformation irreversible cohesive elements for three-dimensional crack-propagation analysis. *International Journal for Numerical Methods in Engineering*, *44*, 1267–1282.
- Oysu, C. (2007). Finite element and boundary element contact stress analysis with remeshing technique. *Applied Mathematical Modelling*, *31*, 2744–2753.
- Paggi, M., & Barber, J. R. (2011). Contact conductance of rough surfaces composed of modified rmd patches. *International Journal of Heat and Mass Transfer*, *54*, 4664–4672.
- Paggi, M., & Ciavarella, M. (2010). The coefficient of proportionality κ between real contact area and load, with new asperity models. *Wear*, *268*, 1020–1029.
- Paggi, M., & He, Q.-C. (2015). Evolution of the free volume between rough surfaces in contact. *Wear*, *336–337*, 86–95.
- Paggi, M., & Reinoso, J. (2015). An anisotropic large displacement cohesive zone model for fibrillar and crazing interfaces. *International Journal of Solids and Structures*, *69*, 106–120.
- Paggi, M., & Reinoso, J. (2018). A variational approach with embedded roughness for adhesive contact problems. *Mechanics of Advanced Materials and Structures*. <https://doi.org/10.1080/15376494.2018.1525454>.
- Paggi, M., & Wriggers, P. (2011). A nonlocal cohesive zone model for finite thickness interfaces—Part II: FE implementation and application to polycrystalline materials. *Computational Materials Science*, *50*(5), 1634–1643.
- Paggi, M., & Wriggers, P. (2012). Stiffness and strength of hierarchical polycrystalline materials with imperfect interfaces. *Journal of the Mechanics and Physics of Solids*, *60*(4), 557–572.
- Paggi, M., & Zavarise, G. (2011). Contact mechanics of microscopically rough surfaces with graded elasticity. *European Journal of Mechanics—A/Solids*, *30*, 696–704.
- Paggi, M., Carpinteri, A., & Zavarise, G. (2006). A unified interface constitutive law for the study of fracture and contact problems in heterogeneous materials. In *Analysis and simulation of contact problems. Lecture notes in applied and computational mechanics* (Vol. 27, pp. 297–304). Springer.
- Paggi, M., Pohrt, R., & Popov, V. L. (2014). Partial-slip frictional response of rough surfaces. *Scientific Reports*, *4*, 5178.
- Patrinos, P., & Bemporad, A. (2014). An accelerated dual gradient-projection algorithm for embedded linear model predictive control. *IEEE Transactions and Automatic Control*, *59*, 18–33.
- Pei, L., Hyun, S., Molinari, J. F., & Robbins, M. O. (2005). Finite element modeling of elasto-plastic contact between rough surfaces. *Journal of the Mechanics and Physics of Solids*, *53*, 2385–2409.
- Peitgen, H. O., & Saupe, D. (1988). *The science of fractal images*. New York: Springer.
- Persson, B. N. J. (2006). Contact mechanics for randomly rough surfaces. *Surface Science Reports*, *261*, 201–227.
- Persson, B. N. J., Albohr, O., Tartaglino, U., Volokitin, A. I., & Tosatti, E. (2005). On the nature of surface roughness with application to contact mechanics, sealing, rubber friction and adhesion. *Journal of Physics: Condensed Matter*, *17*, R1.
- Pohrt, R., & Li, Q. (2014). Complete boundary element formulation for normal and tangential contact problems. *Physical Mesomechanics*, *17*, 334–340.
- Polonsky, I. A., & Keer, L. M. (1999). A numerical method for solving rough contact problems based on the multi-level multi-summation and conjugate gradient techniques. *Wear*, *231*, 206–219.
- Polonsky, I. A., & Keer, L. M. (2000a). A fast and accurate method for numerical analysis of elastic layered contacts. *Journal of Tribology*, *122*, 30–35.
- Polonsky, I. A., & Keer, L. M. (2000b). Fast methods for solving rough contact problems: A comparative study. *Journal of Tribology*, *122*, 36–41.
- Popov, V. L., Pohrt, R., & Li, Q. (2017). Strength of adhesive contacts: Influence of contact geometry and material gradients. *Friction*, *5*, 308–325.

- Prodanov, N., Dapp, W. B., & Mueser, M. H. (2014). On the contact area and mean gap of rough, elastic contacts: Dimensional analysis, numerical corrections and reference data. *Tribology Letters*, 53, 433–448.
- Putignano, C., Carbone, G., & Dini, D. (2015). Mechanics of rough contacts in elastic and viscoelastic thin layers. *International Journal of Solids and Structures*, 69–70, 507–517.
- Raous, M. (1999). Quasistatic signorini problem with coulomb friction and coupling to adhesion. *New developments in contact problems* (Vol. 384, pp. 101–178). CISM Series: Springer.
- Reinoso, J., & Paggi, M. (2014). A consistent interface element formulation for geometrical and material nonlinearities. *Computational Mechanics*, 54, 1569–1581.
- Rey, V., Anciaux, G., & Molinari, J.-F. (2017). Normal adhesive contact on rough surfaces: Efficient algorithm for FFT-based BEM resolution. *Computational Mechanics*, 60, 69–81.
- Sapora, A., & Paggi, M. (2014). A coupled cohesive zone model for transient analysis of thermoelastic interface debonding. *Computational Mechanics*, 53, 845–857.
- Scaraggi, M., Putignano, C., & Carbone, G. (2013). Elastic contact of rough surfaces: A simple criterion to make 2D isotropic roughness equivalent to 1D one. *Wear*, 297, 811–817.
- Schmid, C., & Biegler, L. T. (1994). Quadratic programming methods for reduced hessian SQP. *Computers & Chemical Engineering*, 18, 817–832.
- Sridhar, M. R., & Yovanovich, M. M. (1994). Review of elastic and plastic contact conductance models: Comparison with experiments. *Journal of Thermophysics Heat Transfer*, 8, 633–640.
- Vakis, A. I., Yastrebov, V. A., Scheibert, J., Nicola, L., Dini, D., Minfray, C., et al. (2018). Modeling and simulation in tribology across scales: An overview. *Tribology International*, 125, 169–199.
- Vollebregt, E. A. H. (2014). A new solver for the elastic normal contact problem using conjugate gradients, deflation, and an FFT-based preconditioner. *Journal of Computational Physics*, 257, 333–351.
- Woo, K. L., & Thomas, T. R. (1980). Contact of rough surfaces: A review of experimental work. *Wear*, 58, 331–340.
- Wriggers, P. (2006). *Computational contact mechanics*. Berlin Heidelberg: Springer.
- Wriggers, P., & Reinelt, J. (2009). Multi-scale approach for frictional contact of elastomers on rough rigid surfaces. *Computer Methods in Applied Mechanics and Engineering*, 198, 1996–2008.
- Yastrebov, V. A., Anciaux, G., & Molinari, J.-F. (2015). From infinitesimal to full contact between rough surfaces: Evolution of the contact area. *International Journal of Solids and Structures*, 52, 83–102.
- Yu, N., & Polycarpou, A. A. (2004). Adhesive contact based on the Lennard-Jones potential: A correction to the value of the equilibrium distance as used in the potential. *Journal of Colloid and Interface Science*, 278, 428–435.
- Zavarise, G., Wriggers, P., Stein, E., & Schrefler, B. A. (2019). Real contact mechanisms and finite element formulation—A coupled thermomechanical approach. *International Journal for Numerical Methods in Engineering*, 35, 767–785.
- Zavarise, G., Wriggers, P., Stein, E., & Schrefler, B. A. (1992). A numerical model for thermo-mechanical contact based on microscopic interface laws. *Mechanics Research Communications*, 19(3), 173–182.
- Zavarise, G., Borri-Brunetto, M., & Paggi, M. (2004a). On the reliability of microscopical contact models. *Wear*, 257, 229–245.
- Zavarise, G., Borri-Brunetto, M., & Paggi, M. (2004b). On the resolution dependence of micromechanical contact models. *Wear*, 262, 42–54.
- Zienkiewicz, O. C., & Taylor, R. L. (2000). *The finite element method: Solid mechanics* (Vol. 2). Butterworth-Heinemann.