Finite-element analysis of contact between elastic self-affine surfaces

S. Hyun, L. Pei, J.-F. Molinari, and M. O. Robbins

Department of Physics and Astronomy and Department of Mechanical Engineering, Johns Hopkins University,
Baltimore, Maryland 21218, USA
(Received 22 March 2004; published 31 August 2004)

Finite-element methods are used to study nonadhesive, frictionless contact between elastic solids with self-affine surfaces. We find that the total contact area rises linearly with the load at small loads. The mean pressure in the contact regions is independent of load and proportional to the root-mean-square slope of the surface. The constant of proportionality is nearly independent of the Poisson ratio and roughness exponent and lies between previous analytic predictions. The contact morphology is also analyzed. Connected contact regions have a fractal area and perimeter. The probability of finding a cluster of area a_c drops as $a_c^{-\tau}$ where τ increases with a decrease in roughness exponent. The distribution of pressures shows an exponential tail that is also found in many jammed systems. These results are contrasted to simpler models and experiments.

PACS number(s): 81.40.Pq, 46.55.+d

I. INTRODUCTION

The forces of friction and adhesion between two surfaces are determined by the interactions between atoms at their interface. Most surfaces are rough enough that atoms are only close enough to interact strongly in areas where peaks or asperities on opposing surfaces overlap. Experiments [1–4] and theory [5–12] show that this real area of contact A is often much smaller than the projected area A_0 of the surfaces. They have also correlated [1–3] the increase in friction with normal load W to a corresponding increase in A.

Given the importance of A it is not surprising that there have been many theoretical studies of the factors that determine it in different limits. Some have examined the plastic limit where the local pressure is large enough to flatten asperities [1]. The mean pressure in the contacts is W/A and in the simplest model this has a constant value that is proportional to the hardness. The resulting linear relation between W and A is often given as an explanation for the linear rise of friction with the load [1,13].

The behavior of elastic contacts is more complicated. In the Hertzian limit, where friction and adhesion are ignored, the contact area between a sphere and a flat rises only as $W^{2/3}$. However, Greenwood and Williamson's pioneering work [14] showed that a nearly linear relation between W and A was obtained if one considered a large number of asperities of different height. While this calculation assumed spherical asperities with a uniform radius, a linear relation was also obtained when it was extended by Bush *et al.* [15] to include both a distribution of radii and aspherical asperities. These calculations consider explicit probability distributions for peaks and sum the Hertzian contact areas calculated for each peak without including correlations between peaks. Persson [8] has recently presented a different approach, motivated by the fact that many surfaces have roughness on all length scales that can be described by self-affine scaling [16–18]. His calculation considers the scaling of stress and contact area with the length scale to which surface features are resolved. Nevertheless, his final result for A only differs from the earlier work of Bush *et al.* [15] by a constant factor of $\pi/2$.

All of the above theories treat correlations between contacting regions approximately. Models based on asperities ignore spatial correlations in asperity heights [14,15]. Persson's calculation includes height correlations and becomes exact in the limit of complete contact [9]. However, when $A < A_0$, correlations between local pressures and contacts are only included in an average way. As a result, screening of small bumps on the side of larger bumps may not be included completely. Such correlations could change A, and also the distribution of pressure along the surface. The range over which A is proportional to the load is also difficult to determine from analytical theories. Indeed, the work of Bush *et al.* [15] seems to suggest that the linear region is confined to infinitesimally small loads for a self-affine surface.

A recent numerical study by Borri-Brunetto *et al.* [19] calculated A and the spatial distribution of contact areas using a method that is restricted to zero Poisson ratio. They found a substantial range where A rose linearly with the load, but did not test analytical results for the slope specifically. As in Persson's work [8], their focus was instead on the change in results with increasing resolution of the surface roughness. Batrouni *et al.* [10] have considered scaling of the contact area with the load using a similar method. They found a slight deviation from linearity, but ruled out larger deviations predicted by previous scaling arguments [12].

In this paper, we present numerical calculations of contact area and pressure distributions for a wide range of Poisson ratios, loads, system sizes, self-affine scaling exponents and roughness amplitudes. We first show that A has a well-defined thermodynamic limit, i.e., that for fixed small scale roughness, the fraction of area in contact at a given average normal pressure is independent of the system size. The ratio of A over the load is shown to scale as the inverse of the root-mean-square (rms) surface slope, with a coefficient that lies between the results of Bush *et al.* [15] and Persson [8]. The dependence of A/W on the roughness exponent and Poisson ratio is also obtained. This allows A to be predicted for any elastic rough surface.

We next describe the contact morphology and distribution of connected contact areas. The probability of finding a connected region of area a_c falls off as a power law, $P(a_c)$

$\propto a_c^{-\tau}$, where τ depends only on the roughness exponent. For $H < 0.9$, τ is greater than 2, and the mean contact size is always comparable to the resolution of the calculation. As first noted by Greenwood and Williamson [14], the linear rise in A with the load reflects a linear increase in the number of contacts without any increase in their mean size or probability distribution. Our results are contrasted to a common model where contacts form in the regions where undeformed surfaces would overlap. This approximation has been used to interpret optical images of contacts [2,3], and by Greenwood and Wu as a method of estimating the statistics of asperity sizes [6]. We show that it gives much too large a total contact area, and a qualitatively different distribution of a_c with $\tau < 2$. Optical methods may include regions that are merely close to touching as part of the contact. We find that this can explain the observed discrepancy between experiment and calculation.

The distribution of contact pressures p is also studied. We find that results for all system sizes, roughness amplitudes, and roughness exponents collapse onto a universal curve when the pressure is normalized by its mean value. The mean value, W/A, can be obtained from the roughness amplitude as described above. The distribution decreases monotonically with increasing p, and has an exponential tail at large p. In contrast, approximate analytic results for the pressure distribution decay as a Gaussian at large p [8]. Possible connections to exponential stress distributions in jammed systems [20–22] are mentioned.

In Sec. II, we describe the numerical procedures used to generate self-affine surfaces, mesh them, and determine their deformation using an explicit dynamic finite-element method. In Sec. III, we present numerical results from the contact analysis, including the contact area, contact morphology and pressure distribution. Section IV presents conclusions from the current study and discusses avenues for future research.

II. NUMERICAL SIMULATION

A. Geometry

As in previous analytical and numerical calculations, we use a well-known result from contact mechanics to simplify the geometry. If there is no friction or adhesion between two rough surfaces and the surface slope is small, then elastic contact between them can be mapped to contact between a single rough surface and a rigid flat plane [13]. The effective modulus E' that controls the contact area is given by $1/E' = (1 - \nu_1^2)/E_1 + (1 - \nu_2^2)/E_2$ where ν_i and E_i are the Poisson ratios and Young's moduli of the two surfaces. The height h of the new rough surface is given by the difference between the local heights of the original undeformed surfaces. We consider the case where both of the surfaces have the same self-affine scaling properties but are uncorrelated. Then the height h has the same scaling properties, but with a larger amplitude.

Many surfaces have roughness on all length scales that can be described by self-affine fractal scaling. Unlike self-similar fractals, self-affine fractal surfaces exhibit different scaling normal to the interface than along it. We consider
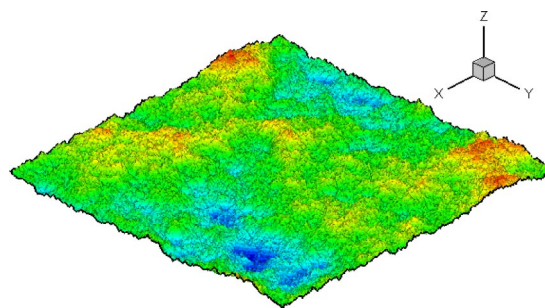

FIG. 1. (Color online) Self-affine fractal surface image (256×256) generated by the successive random midpoint algorithm. Heights are magnified by a factor of 10 to make the roughness visible, and the color varies from dark (blue) to light (red) with increasing height.

geometries with periodic boundary conditions in the $x-y$ plane and specify the surface by the height h along the z axis. For a self-affine surface, variations in the height over a lateral length scale ℓ rise as ℓ^H where $H < 1$ is called the Hurst or roughness exponent. Since $H < 1$, the surface looks smoother at larger length scales. Many researchers also specify the scaling by an effective fractal dimension $d - H$, where d is the spatial dimension.

To generate three-dimensional self-affine fractal surfaces with rms roughness Δ at small scales, we adopted the successive random midpoint algorithm of Voss [18,23]. The x-y plane is divided into a uniform square grid with unit spacing and L nodes along each axis. At the first step, the center of the entire grid is displaced by a height chosen at random from a Gaussian distribution of width $\ell^H \Delta$, where $\ell = L/\sqrt{2}$ is the distance from the center to the corners. This center point then becomes one corner of new squares that are rotated by 45° and have a new corner to center distance ℓ that is smaller by a factor of $\sqrt{2}$. The center of each new square is assigned a height equal to the average of the corner heights plus a random number chosen from a Gaussian of width $\ell^H \Delta$. This process is iterated down to $\ell = 1$, guaranteeing that the variation in height scales with ℓ in the appropriate manner. Figure 1 shows a typical self-affine surface with $H = 0.5$. Note that the height variation is enhanced by a factor of 10 to make it visible in Fig. 1.

It is common to test the scaling properties of self-affine surfaces by calculating the Fourier transform $C(q)$,

$$C(q) = (2\pi)^{-2} \int d^2r \exp[-i\mathbf{q} \cdot \mathbf{r}]C(\mathbf{r}), \qquad (1)$$

of the height–height correlation function,

$$C(\mathbf{r}) = \langle h(\mathbf{r} + \mathbf{r}')h(\mathbf{r}') \rangle, \qquad (2)$$

where $\langle h(\mathbf{r}') \rangle = 0$, $\mathbf{r}$ and $\mathbf{r}'$ are vectors in the x-y plane, and the angular brackets indicate an average over $\mathbf{r}'$. The function $C(q)$ should be isotropic and decay as $q^{-2(1+H)}$. We have generated surfaces with H between 0.3 and 0.9 and verified that the corresponding $C(q)$ has the correct power law scaling. For a given Δ, the values of $C(q)$ at large q are very insensitive to the random seed. However, there are large fluc-
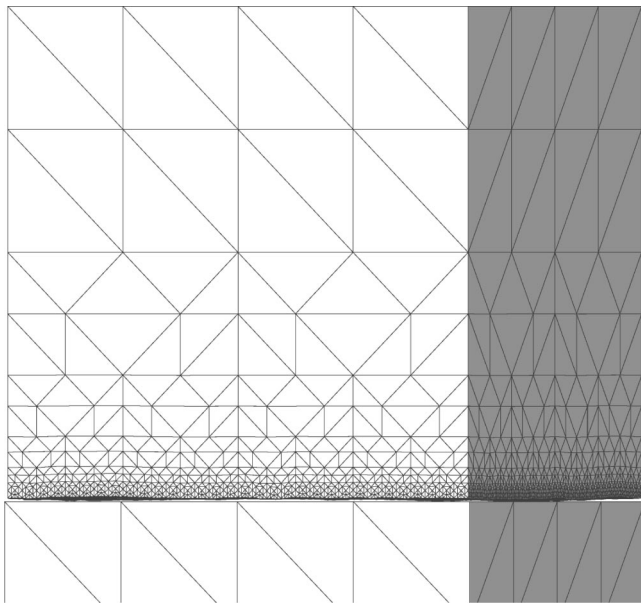
FIG. 2. Geometry of a finite element mesh in an elastic body (top) with a self-affine surface that is pushed down on a flat, rigid substrate. Periodic boundary conditions are applied in the plane of the interface.

tuations at small q where the roughness is dominated by the first few random numbers that are chosen at the largest length scales. These first random numbers also dominate the mean-square height variation:

$$C(r=0) = \langle |h(\mathbf{r}')|^2 \rangle, \qquad (3)$$

and we find large variations in $C(0)$ for surfaces with the same scale roughness Δ. As discussed below, the contact area is determined by Δ and is relatively insensitive to fluctuations in $C(0)$.

B. Mesh generation

As noted above, numerical simulations are done for a rough elastic surface in contact with a perfectly rigid flat surface. A typical finite-element mesh is illustrated in Fig. 2. The mesh is discretized with 10-node tetrahedral elements. These elements contain three integration points and quadratically interpolate the displacement field. A coarse mesh is used for the rigid surface to improve numerical efficiency. A fine mesh for the elastic surface is prepared in two stages.

First, a fine mesh for a flat surface is obtained using a longest edge propagation path refinement scheme, which is ideally suited to obtaining strong mesh gradations as well as preserving a high mesh quality [24]. A cube of side $L = 2^n$ is initially filled with a coarse mesh. Each tetrahedral element at the outer surface is then divided to produce twice as many surface nodes and the mesh is refined. This process is repeated until surface nodes form a uniform square grid of unit spacing. Using this technique, meshes with L up to 512 are created. The mesh for a 512×512 surface grid contains about 911 000 nodes and 568 000 elements.

Next, the desired surface heights $h(x,y)$ are imposed on the contact surface. Moving only the surface nodes produces

badly distorted elements that would at best require impractically small time steps and at worst produce negative Jacobians. Thus all nodes are moved by a fraction of the local height that depends on the initial height z_0 of the node above the bottom of the elastic cube (Fig. 2). The magnitude of the change, Δz, decreases to zero at the top of the cube so that the top surface remains flat. The specific form for the displacement is $\Delta z(x,y,z_0) = h(x,y)(L-z_0)^a$ where $a=6$ usually gives good meshes.

C. Finite-element simulation

The goal is to determine the equilibrium contact geometry at a given load. An implicit approach is too memory intensive for the system sizes of interest. Instead we use an explicit integration algorithm combined with a dynamic relaxation scheme. Three different algorithms were compared to ensure accuracy. In the first, the top surface is given a small velocity and its impact with the bottom surface is followed. In the second, the displacement of the nodes at the top of the elastic cube (Fig. 2) is incremented at a fixed rate or in small discrete steps. In the third, a constant force is applied to each of these nodes and gradually incremented. In the second and third algorithms kinetic energy is removed using the method described below. All three methods give equivalent results for the total area. Unless otherwise noted, the results presented below were obtained with the third algorithm. We confirmed that the mean normal stress was independent of the height to ensure that stress had equilibrated throughout the system.

Within the Lagrangian framework, finite-element discretization of the field equations leads to a discrete system of equations:

$$\mathbf{M}\ddot{\mathbf{x}}_{n+1} + \mathbf{F}^{\text{int}}_{n+1}(\mathbf{x},\dot{\mathbf{x}}) = \mathbf{F}^{\text{ext}}_{n+1}, \qquad (4)$$

where $\mathbf{x}$ is the array of nodal coordinates, $\mathbf{M}$ the mass matrix, $\mathbf{F}^{\text{ext}}_{n+1}$ the external force array, and $\mathbf{F}^{\text{int}}_{n+1}$ the internal force array arising from the current state of stress. The second-order accurate central difference scheme is used to discretize Eq. (4) in time [25,26]. A small time step was used in order to be below the limit of stability [25].

The above equations conserve energy and will not converge to the static equilibrium configuration. Optimum convergence is achieved by removing a fraction of the kinetic energy of each node at regular intervals. The characteristic time for stress equilibration across the elastic cube is given by the time τ_L for sound propagation across its height L. Equilibrium is reached in a few τ_L by scaling all velocities by a factor of 0.9 at intervals of $\tau_L/10$. Other procedures gave equivalent results, but with longer run times.

The internal forces F^{int} are calculated using a linear elastic isotropic constitutive law. All our results are expressed in dimensionless form by normalizing pressures by the effective modulus E'. The Poisson ratio ν was varied from 0 to 0.45. Periodic boundary conditions are imposed at the contact surfaces to eliminate boundary effects.

A contact algorithm is used only to enforce the impenetrability constraint on the two surfaces. Adhesive and frictional forces are not considered in the current work. We adopt a

conventional master/slave approach with a predictor/corrector split in the Newmark time-stepping algorithm [27]. The rough surface of the elastic cube is identified as slave while the rigid surface is master. The predictor part of the Newmark algorithm neglects the contact constraints and therefore consists of an unconstrained step, with the result

$$\mathbf{x}_{n+1}^{\text{pred}} = \mathbf{x}_n + \Delta t \mathbf{v}_n + \frac{\Delta t^2}{2} \mathbf{a}_n, \tag{5}$$

$$\mathbf{v}_{n+1}^{\text{pred}} = \mathbf{v}_n + \frac{\Delta t}{2} \mathbf{a}_n. \tag{6}$$

This predictor solution needs to be corrected in order to comply with the impenetrability constraints. The net result of imposing these constraints is a set of self-equilibrated contact forces that modify the predictor positions and velocities. Since the contact surfaces are presumed to be smooth, normals are well defined and the surfaces can be unambiguously classified as master and slave. The configuration before calculating internal forces is therefore,

$$\mathbf{x}_{n+1}^S = \mathbf{x}_{n+1}^{S,\text{pred}} - \frac{\Delta t^2}{2} \frac{\mathbf{N}_{n+1}^S + \mathbf{F}_{n+1}^S}{M^S}, \tag{7}$$

$$\mathbf{v}_{n+1}^S = \mathbf{v}_{n+1}^{S,\text{pred}} - \Delta t \frac{\mathbf{N}_{n+1}^S + \mathbf{F}_{n+1}^S}{M^S}. \tag{8}$$

Here $M^{()}$ denotes the nodal mass, the superscript $()^S$ designates nodes that belong to the slave surface, and the vectors $\mathbf{N}$ and $\mathbf{F}$ are the normal and frictional forces, respectively. Friction will not be taken into account in the remainder of the paper, but is discussed in Ref. [27].

Formulation of an appropriate system of forces is obtained by considering a configuration in which a master surface triangle (facet of a tetrahedral finite element) is penetrated by several slave nodes. For each of the penetrating slave nodes, let δ be the normal depth of penetration to be corrected by the contact forces. The contact constraints determine a local problem with the normal slave force as an unknown, which is obtained as a direct function of the penetration δ and the master normal, $\mathbf{n}^M$,

$$\mathbf{N}_{n+1}^S = M^S \frac{2\delta}{\Delta t^2} \mathbf{n}_{n+1}^M. \tag{9}$$

D. Calculating the contact area

The contact algorithm just described identifies all slave nodes on the top surface that attempt to penetrate the flat bottom surface. We obtain the total contact area by multiplying the number of penetrating nodes n_p by the area of the square associated with each node. In most cases we report the fractional area $A/A_0 = n_p/L^2$ where L^2 is the total number of surface nodes.

In principle, the area associated with each node varies with the normal load if the Poisson ratio is nonzero. However, tests using a Voronoi tessellation to determine the area for each node gave equivalent results for the relatively
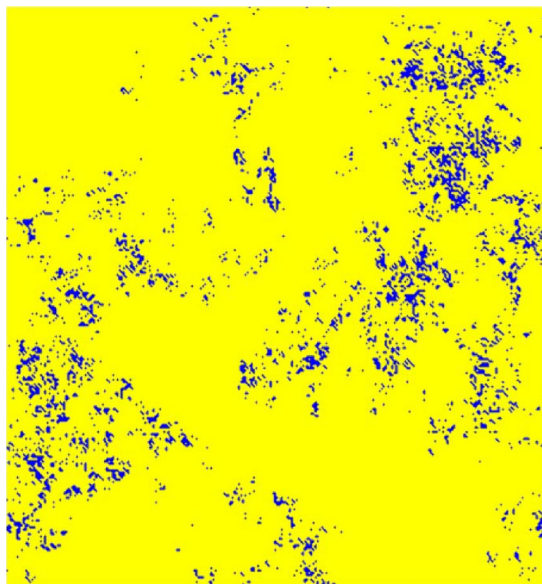

FIG. 3. (Color online) Regions of contact (dark) for a surface with $L=256$, $H=1/2$, $\nu=0$ and $\Delta=0.082$. The fraction of the area in contact is $A/A_0=0.1$.

smooth surfaces considered here. A Voronoi approach becomes important for rougher surfaces or irregular grids. A more fundamental ambiguity in the contact area comes from the fact that contacts (Fig. 3) contain many disjoint regions most of which contain only a few nodes. The consequences of this are discussed in Sec. III C.

There are also complications in defining the contact area in experimental systems. Dieterich and Kilgore's optical method [2,3] identifies any region where the surfaces are closer than some fraction of the wavelength as in contact. Due to the fractal nature of contacts (Fig. 3) this may overestimate the true area. At the atomic scale, contact is difficult to define. If one associates contact with a finite interaction, then A/A_0 would always be unity since the van der Waals interactions between surfaces extend to arbitrary distances. A more practical definition is to associate contact with the separation at which the net interaction becomes strongly repulsive due to the overlap of electrons on opposing surfaces. This leads to a range of separations where surfaces are in contact. As in the optical measurements, the contact area can be greatly enhanced relative to the penetration definition used here. Direct comparisons to atomistic models will be presented in another paper [28].

III. RESULTS

A. Contact area versus external load

As noted in Sec. I, analytic theories of contact between self-affine surfaces predict that the real area of contact A should be proportional to the applied load W at small loads [7,8,14,15]. To make the load dimensionless we normalize it by A_0 times the effective modulus E'. Figure 4 shows a plot of the fraction of the projected area that is in contact A/A_0 as a function of the normalized load for a system with $L=256$, $\Delta=0.082$, $\nu=0$ and $H=1/2$. As expected, the area is propor-
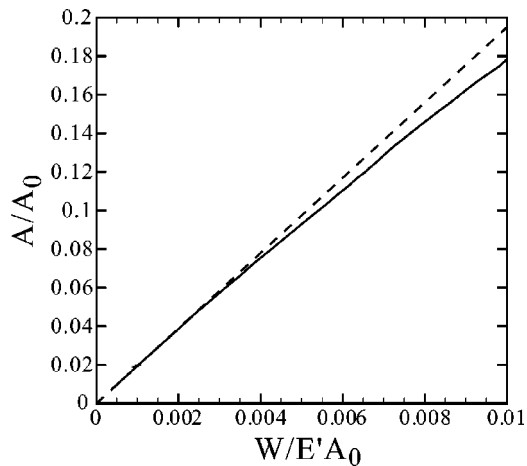
FIG. 4. Fractional contact area A/A_0 (solid line) as a function of the normalized load $W/E'A_0$ for $L=256$, $\Delta=0.082$, $\nu=0$ and $H=1/2$. The dashed line is a fit to the linear behavior at small areas.

tional to the load at small loads. A growing deviation from this proportionality is evident as A/A_0 increases above 4 or 5%.

To emphasize deviations from linearity, the dimensionless ratio of true contact area to load AE'/W is plotted against $\log_{10}A/A_0$ in Fig. 5. Results for L between 64 and 512 fall onto almost identical curves. The small variation between curves is comparable to that between different random surfaces of the same size. In each case AE'/W is nearly constant when from 1% to 8% of the surface is in contact. This implies that the mean pressure in contacting regions $\langle p \rangle \equiv W/A$ is also constant. The ratio AE'/W drops as A/A_0 increases further because A/A_0 is bounded by unity while the normalized load keeps increasing. The ratio has decreased by a factor of two by $A/A_0=70\%$.
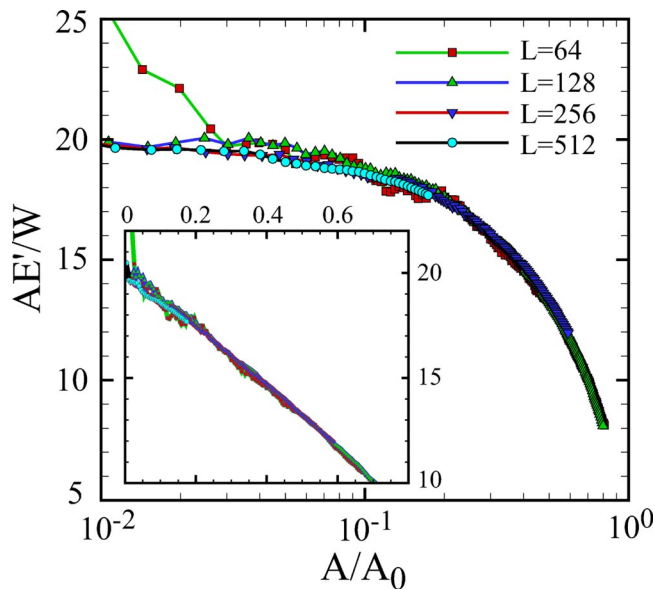

FIG. 5. (Color online) Dimensionless ratio of the area to load AE'/W vs $\log_{10}(A/A_0)$ for the system sizes indicated. In all cases $\Delta=0.082$, $\nu=0$ and $H=1/2$. The inset shows a linear plot of the same data.

Most of the analytic theories mentioned above explicitly assume that there is a statistically significant number of asperities in contact, and that only the tops of asperities are in contact. The latter assumption breaks down as A/A_0 approaches unity, and this contributes to the decrease at large A/A_0 in Fig. 5. The first assumption must break down for our systems when the total number of nodes in contact, L^2A/A_0, is small. This explains the rise in the $L=64$ data for $A/A_0<2\%$ in Fig. 5. This rise is dependent on the specific random surface generated, and is particularly dramatic for the case shown. Examination of this and other data indicates that proportionality between load and area is observed when there are more than 100 contacting nodes ($L^2A/A_0>100$).

Batrouni *et al.* [10] considered the same range of system sizes for $\nu=0$. They fitted all data from $A/A_0 \lesssim 0.2$ to a power law and found $W \propto A^\gamma$ with $\gamma>1$. Their numerical results clearly rule out an earlier prediction [12] that $\gamma=(1+H)/2$, but we do not believe that their results are inconsistent with an initially linear relation between load and area ($\gamma=1$). Their fit included regions where the number of contacting nodes is below the minimum threshold just described. In addition, their value of γ decreased steadily towards unity with increasing L, varying from 1.18 at $L=32$ to 1.08 at $L=256$. We believe that this small difference from unity is within the systematic errors associated with the limited scaling range. Note that $\gamma=1.1$ would imply a 25% decrease in Fig. 5 from $A/A_0=10^{-2}$ to 10^{-1}, which is much larger than the observed change ($\sim 5\%$). The inset of Fig. 5 shows that data up to $A/A_0=0.7$ can be described by a very simple linear decrease in A/W with an increase in A. We conclude that $\gamma=1$, but that leading quadratic corrections to scaling give a larger apparent exponent when a large range of data is fit.

The lack of system-size dependence in Fig. 5 may appear surprising in the context of some previous results for self-affine surfaces. These studies [8,19] considered a fixed roughness at large scales and examined changes in contact area with increasing resolution. They found that the contact area decreased as the number of nodes increased because the local slope of the surface became rougher at higher resolution. Our Fig. 5 compares results for the same small scale roughness and shows that there is a well-defined thermodynamic limit as one increases the size of the system. This result is not obvious, since the rms roughness at the scale of the entire contact rises as $L^H\Delta$. Apparently this increase in large scale roughness is irrelevant because it rises sufficiently slowly with L. As noted in Sec. II A, the large scale roughness is sensitive to the first few random numbers chosen in generating the self-affine surface. If surfaces with the same large scale roughness are compared, substantial differences are found because the small scale roughness varies. These fluctuations are absent when results from the same small scale roughness are compared.

The existence of a well-defined thermodynamic limit allows us to consider results for a single system size in the remainder of Sec. III, and to extrapolate the results to other cases. In the following we will focus on the constant region in Fig. 5 and examine variation of this constant value with the statistical properties of the interface. Unless noted, all
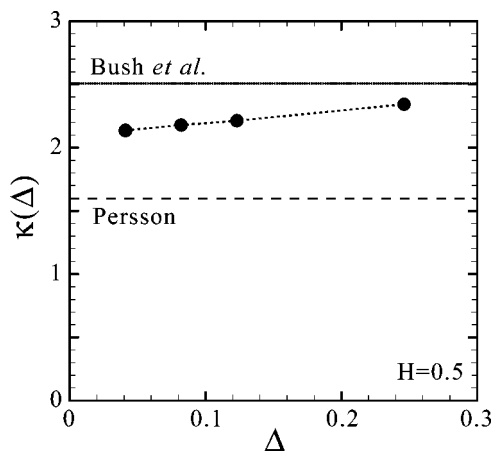
FIG. 6. Product κ [Eq. (10)] as a function of roughness Δ for $H=1/2$ and $\nu=0$, and the constant values predicted by Bush *et al.* (solid line) [15] and by Persson (dashed line) [8]. The dotted line is a guide to the eye.

results are for $L=256$ and uncertainties due to statistical fluctuations are less than 5%.

B. Comparison to analytic theories

Bush *et al.* [15] found that the ratio plotted in Fig. 5 should increase inversely with the root-mean-square slope of the surface $\sqrt{\langle|\nabla h|^2\rangle}$. More specifically, they predicted that the quantity,

$$\kappa \equiv \sqrt{\langle|\nabla h|^2\rangle}AE'/W, \qquad (10)$$

should have a constant value $\kappa=\sqrt{2\pi}$. Persson arrived at a rather different looking expression in terms of the height–height correlation function [Eqs. (1) and (2)] [8]. However, it can be reduced to a prediction that $\kappa=\sqrt{8/\pi}$ using the fact that $q^2C(q)$ is the Fourier transform of $|\nabla h|^2$. Note that both predictions have a well-defined thermodynamic limit that is independent of large scale roughness, just as observed in our simulations.

For our surfaces, Δ is the rms change in height between adjacent nodes in each of the two spatial directions. Thus $\sqrt{\langle|\nabla h|^2\rangle}=\sqrt{2}\Delta$. Numerical results for κ are plotted against Δ in Fig. 6 for $\nu=0$ and $H=1/2$. The value of κ only changes about 10%, while the roughness changes by almost an order of magnitude. If Δ is increased to larger values than considered here, the local slope of the surface exceeds unity in some regions. This regime was not studied because it requires a different meshing algorithm and most treatments of contacts assume that the local slope remains less than unity.

We also evaluated κ over the range of roughness exponents typically observed on real surfaces, $0.3 \leqslant H \leqslant 0.7$ [17,18], and at the higher values of 0.8 and 0.9. Figure 7 shows results for a fixed value of $\Delta=0.082$. Even though the large scale roughness, $L^H\Delta$, varies by more than an order of magnitude, κ changes by less than 30%. There is a nearly linear decrease of κ with an increase of H that may be influenced by uncertainties in determining the true contact area. As shown in Sec. III C, increasing H also increases the popu-
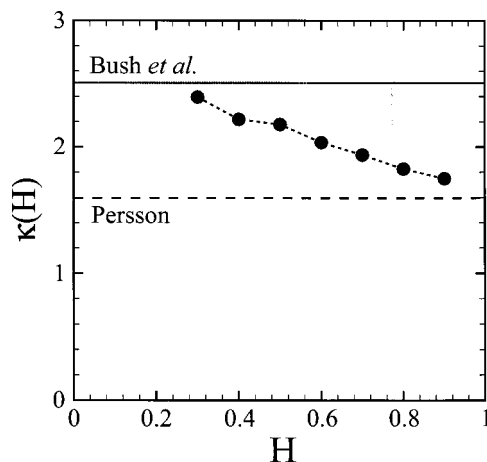
lation of large clusters and may reduce uncertainties associated with assuming that the entire square region around contacting nodes is in contact.

Figure 8 shows that the Poisson ratio also has relatively little effect on κ. Results for each value of H are normalized by the value $\kappa(0)$ obtained at $\nu=0$. In every case there is a nearly linear rise in κ at small ν that appears to saturate as ν approaches the limiting value of 0.5. The total change of around 10% is comparable to the change found with Δ. The increase in κ with ν appears to be related to increased interactions between nearby asperities. The lateral expansion in response to a normal stress increases with ν. This reduces the local curvature, making it easier for adjacent regions to come into contact. Detailed analysis of neighboring asperities shows that a smaller peak between two contacts may be brought up in to contact at high Poisson ratios.

All of the values of κ in Figs. 6 and 7 lie between the analytic predictions of Bush *et al.* and Persson. Our results suggest that using a value of $\kappa=2.2$ should predict the ratio of area to load within about 10% over a wide range of surface geometries at $\nu=0$. Figure 8 indicates that the value of κ

FIG. 7. Product κ [Eq. (10)] as a function of H for $\Delta=0.082$, and the constant values predicted by Bush *et al.* (solid line) [15] and by Persson (dashed line) [8]. The dotted line is a guide to the eye.

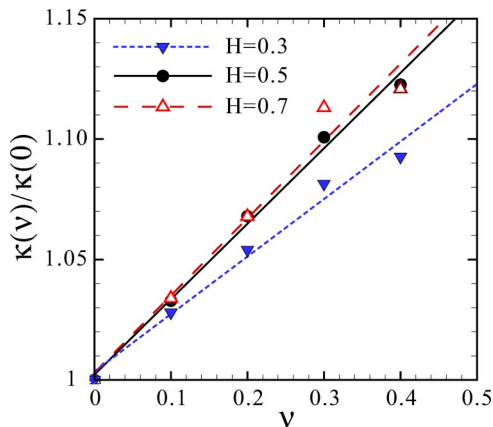

FIG. 8. (Color online) Ratio of κ to its value at $\nu=0$ as a function of ν. Results for H between 0.3 and 0.7 show nearly the same linear rise with ν. Lines are linear fits to the data.
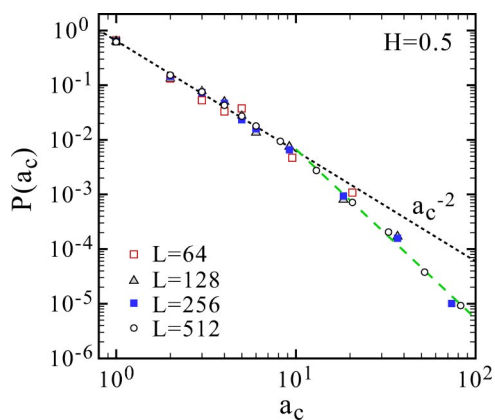
FIG. 9. (Color online) Probability P of a connected cluster of area a_c as a function of a_c for $\nu=0$, $H=1/2$, $\Delta=0.082$ and the indicated system sizes. All results follow a power law, $P(a_c)\sim a_c^{-\tau}$, with $\tau=3.1$ (dashed line) at large a_c. The dotted line corresponds to $\tau=2$.

should be increased linearly to about 2.5 as ν increases to the limiting value of 0.5.

The agreement with these analytic predictions is quite good considering the ambiguities in discretization of the surface. Both analytic models assume that the surface has continuous derivatives below the small length scale cutoff of the roughness. Bush *et al.* [15] considered contact between elliptical asperities and Persson [8] removed all Fourier content above some wave vector. While we use quadratic shape functions, the contact algorithm only considers nodal heights and assumes that contact of a node implies contact over the entire corresponding square. One might expect that this assumption would lead to larger areas of contact, and our results do lie above those of Persson's. Discretization would not be important if the spacing between nodes were much smaller than the typical size of asperity contacts. However, as we now show, the majority of the contact area consists of clusters containing only a few nodes. The number of large clusters grows as $H\rightarrow 1$, which may explain why our numerical results approach Persson's prediction in this limit.

C. Distribution of connected contact regions

Most continuum theories approximate the real contact area by summing over many disconnected asperity contacts, each of which has a circular [14] or elliptical [15] shape. The connected regions in our calculated contacts (in, e.g., Fig. 3) are considerably more complicated. We consider two nodes to be connected if they are nearest neighbors on the square lattice of interfacial nodes [29]. All clusters of connected contacting nodes are then identified for each load. The area of each cluster a_c is just the number of connected nodes, since each represents a square region of unit area.

Figure 9 shows the probability of finding a cluster of a given area $P(a_c)$ as a function of a_c for $H=0.5$, $\Delta=0.082$ and $\nu=0$. Results for different system sizes collapse onto a common curve. For $a_c>8$ the curve can be described by a power law $P(a_c)\sim a_c^{-\tau}$ with $\tau=3.1$ (dashed line). This rapid falloff ($\tau>1$) means that the integral of P is dominated by small
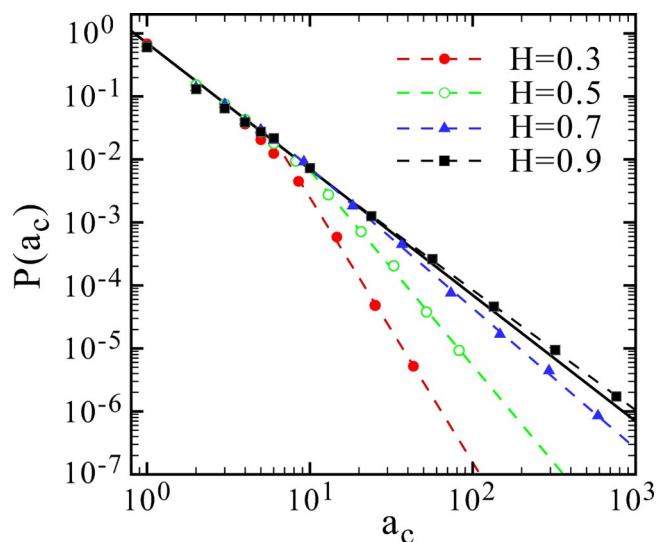

FIG. 10. (Color online) Probability P of a connected cluster as a function of area a_c for $\nu=0$, $\Delta=0.082$, $L=512$ and the indicated values of H. Dashed lines indicate asymptotic power law behavior with $\tau=4.2$, 3.1 and 2.3 for $H=0.3$, 0.5 and 0.7, respectively. The solid line corresponds to $\tau=2$.

clusters. Thus even though the maximum observed cluster size grows with L, the value of P at small a_c is unaffected. All of the data shown in Fig. 9 are for A/A_0 between 5% and 10%, but we find that the distribution of clusters is nearly constant for $A/A_0\lesssim 10\%$. This is the same range where A and the load are nearly linearly related. The probability of large clusters rises markedly for $A/A_0>0.3$, as clusters begin to merge and eventually percolate across the interface. The following data are all for $A/A_0\lesssim 0.1$.

The model considered by Greenwood and Williamson also gives a load independent $P(a_c)$ and this is a central reason for the linear relation between the load and area in their model. As the load increases, each existing cluster grows larger and new small clusters are generated in a way that maintains a stationary distribution of cluster sizes. Only the total number of clusters changes, and it rises linearly with the load. Our calculated $P(a_c)$ and the total number of clusters both follow this behavior. However, the distribution and the shapes of the clusters are very different than assumed by Greenwood and Williamson.

The variation of $P(a_c)$ with H is shown in Fig. 10. Results for very small clusters ($a_c<8$) are nearly independent of H, but the asymptotic power law behavior at large a_c changes dramatically. The range of scaling behavior is too small for precise determination of τ, but our data are consistent with $\tau=3.1\pm 0.2$ for $H=0.5$. When $H<0.5$ there is an anticorrelation between the surface slopes in nearby regions [18]. This leads to more rapid upward and downward fluctuations that make large contacts unlikely and yields a larger $\tau=4.2\pm 0.4$ for $H=0.3$. When $H>0.5$ there is a positive correlation between local slopes, yielding larger clusters. Fits give $\tau=2.3\pm 0.2$ for $H=0.7$ and $\tau=1.9\pm 0.1$ for $H=0.9$. For $H<0.9$ we find $\tau>2$, which implies that the mean cluster size $\langle a_c\rangle$ is dominated by small clusters. Directly calculated sizes are indeed independent of both L and A/A_0. We find $\langle a_c\rangle=1.8$,
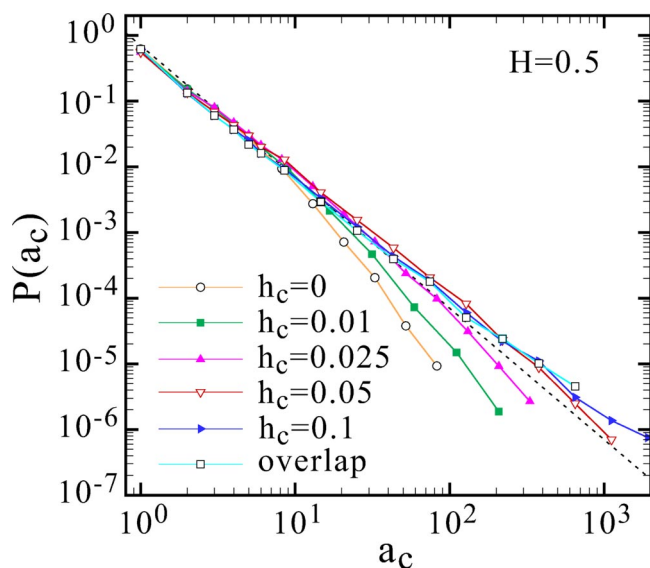
FIG. 11. (Color online) Probability P of a connected cluster as a function of area a_c for $\nu=0$, $H=1/2$, $\Delta=0.082$, $L=512$ and different criteria for contact. The probability distribution for the rigid overlap model (open squares) falls off more slowly than a_c^{-2} (dotted line). When only contacting nodes are included ($h_c=0$), P falls off more rapidly (open circles). As h_c increases, results from the full calculation approach the overlap results.

2.5 and 4.0 for $H=0.3$, 0.5 and 0.7, respectively. For $H=0.9$, τ appears to be slightly smaller than 2, and the mean cluster size grows weakly with L.

Some approximate treatments of contact begin by assuming that the two surfaces do not deform and then determine the regions where the two solids would interpenetrate [2,3,6]. Scaling arguments [18] and simulations [18,30] show that this rigid overlap model gives a power law distribution of connected areas at large a_c with $\tau=2-H/2$. This is qualitatively consistent with Dieterich and Kilgore's experiments where τ varied from 1 to 2 and tended to decrease with an increase of H. However, it is qualitatively different from our results where τ is always greater than 2. One consequence is that the mean cluster size from the rigid overlap model diverges with the system size as L^H, while it remains of order the discretization size in our calculation [29].

A possible explanation for the discrepancy between our results and experiment is that the latter identifies regions that are within some small fraction of the wavelength of light as being in contact. Figure 11 illustrates the dramatic effect that this can have on $P(a_c)$. The uppermost curve shows the cluster distribution obtained by applying the rigid overlap model to our surfaces. The asymptotic slope is consistent with the analytic prediction for $H=0.5$: $\tau=2-H/2=1.75$. The lowermost curve is our result for the actual contact area. The intermediate curves were obtained by changing our definition of contact to include all nodes that are separated by less than some value h_c. As h_c increases, the number of nodes in the contact region rises and the probability of large clusters grows. When h_c is comparable to or larger than $\Delta=0.082$, $P(a_c)$ follows the rigid overlap prediction quite closely. It is likely that the optical experiments were in this limit. However, it is also likely that plastic deformation is important in
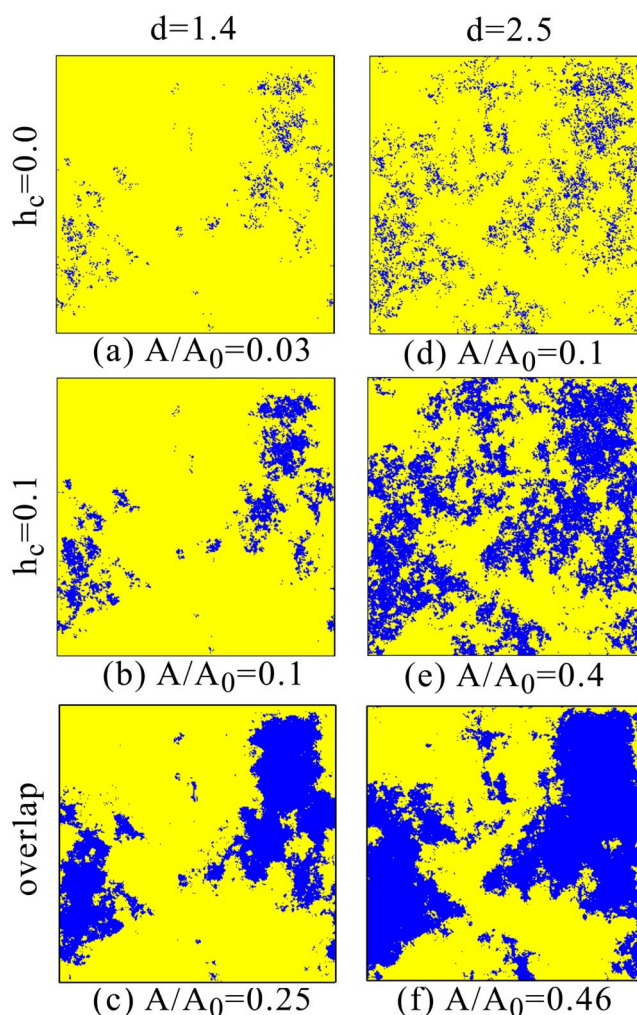

FIG. 12. (Color online) Contact morphology for $L=256$ for two different values of displacement d following contact: $d=1.4$ and 2.54. (a), (d) Regions in true contact ($h_c=0$); (b), (e) regions where the surface separation is less than $h_c=0.1$. (c), (f) Contacts predicted by the rigid overlap model. The fractional contact area (A/A_0) is indicated for each case.

these experiments. This effect will be explored in future work.

D. Contact morphology

The contact morphologies produced by different models are contrasted in Fig. 12. Results for two values of the interpenetration d are shown, where d is the downward displacement applied to the top of the elastic solid (Fig. 2) after the surfaces first touch. The top panels show the results of the full calculation, the middle panels are for $h_c=0.1$, and the bottom panels show the contacts obtained from the rigid overlap model. The first obvious difference between the results is that the rigid overlap model grossly overestimates the fraction of area in contact. For small d the actual area is roughly eight times smaller than that given by the overlap model.

As seen above, the rigid overlap model also gives many more large clusters. Moreover the shapes of the clusters are
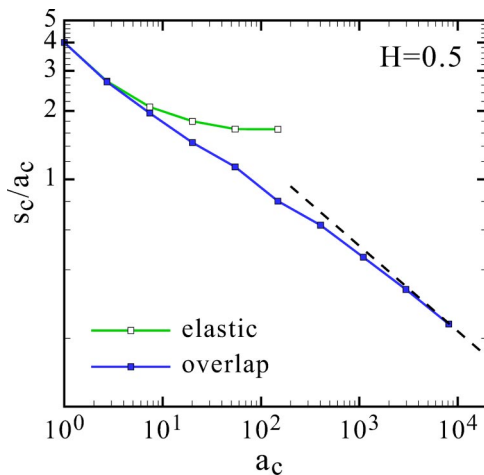
FIG. 13. (Color online) Ratio of perimeter s_c to area a_c vs the area for the full calculation (open squares) and rigid overlap model (closed squares). The dashed line shows the asymptotic prediction for the overlap model.

FIG. 14. (Color online) Probability distribution of the local pressure at contacting nodes for different system sizes with $\Delta=0.082$, $\nu=0$, $H=0.5$ and A/A_0 between 5% and 10%.

quite different. Analytic studies predict that the overlap model should give clusters with a nonfractal interior, but with a fractal interface. More specifically, if R is the diameter of a cluster, then the area $a_c \propto R^2$, but the perimeter length $s_c \propto R^{D_f}$ where the fractal dimension $D_f=(3-H)/2$ [18,30]. Thus as the clusters grow in size, the perimeter becomes smaller and smaller relative to the area: $s_c/a_c \propto a_c^{-(1+H)/4}$. Figure 13 shows that our results for the overlap model are consistent with this scaling prediction. However, the results for the full calculation are quite different. The value of s_c/a_c approaches a constant at large a_c, indicating that the perimeter and area are both fractals and have the same fractal dimension. Plots of s_c and a_c as a function of R show D_f is roughly 1.6 for $H=0.5$. Note that s_c is actually larger than a_c because it is defined as the number of missing nearest neighbors along the periphery. Thus it would be four for a cluster containing a single node.

Results for other values of H look very similar to those for $H=0.5$. In each case s_c/a_c saturates at large a_c, indicating the area and perimeter have the same fractal dimension. The limiting value of s_c/a_c decreases from about 1.7 for $H=0.3$ to 1.1 for $H=0.9$. The large number of perimeter nodes leads to some ambiguity in the total area obtained from our calculation. If only a fraction of the square around each node were actually in contact, then the true contact area would be smaller, moving our results in Figs. 6 and 7 closer to Persson's result [8,9]. It is interesting that our results approach Persson's prediction as $H \rightarrow 1$, and the perimeter becomes less important.

Despite the above differences, the rigid overlap model does provide information about where real contacts may occur. The distance between surfaces is always larger than that given by the overlap model because $\nu=0$. Thus all of the contacts in Figs. 12(a) and 12(d) are part of the overlapping regions shown in Figs. 12(c) and 12(f). Only a fraction of the overlapping regions is in contact, because a local peak can screen a neighboring valley from contact. As pointed out by Greenwood [5], a large fraction of points are local maxima,
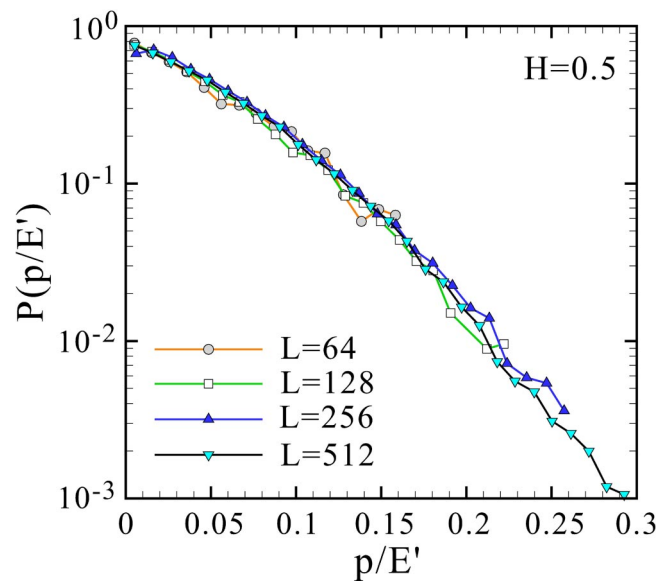
so the average cluster size is comparable to the lattice resolution. On the other hand, a small local maximum can only screen a small local region. Thus there tend to be many small contacts in the regions where overlap first occurs. These points lie in the middle of the large clusters in Figs. 12(c) and 12(f). A higher density of clusters and clusters of larger size are found in these regions of Figs. 12(a) and 12(d). As noted above, when nodes that are within a distance comparable to Δ are considered in contact, the distribution of clusters approaches that for the rigid overlap model. Figures 12(b) and 12(e) show the contact morphology produced in this limit ($h_c=0.1$). Note that the clusters in Figs. 12(a) and 12(d) have been connected into larger clusters that still lie within those of Figs. 12(c) and 12(f). Growth is most pronounced in regions where overlap is greatest. These regions carry a greater share of the load and flatten more.

Greenwood and Wu [6] have recently reconsidered when a local peak should be considered an asperity. They concluded that one should think of each cluster in the rigid overlap model as a single asperity, and use the diameter and height of the overlap to determine the dimensions of an effective ellipsoidal asperity. Our results indicate that the original view [5] that almost all points are asperities and that the typical asperity diameter is comparable to the lattice size provides a more accurate description of the contact. However, the revised approach of identifying overlapping regions with asperities may give a better description of subsurface stresses, because it captures correlations in the location of load bearing regions. Since the maximum shear stress is usually below the surface, the overlap model [6] may be useful in modeling wear.

E. Distribution of local pressures

Plastic deformation at the interface will be influenced by how pressure is distributed within the contact area. We find
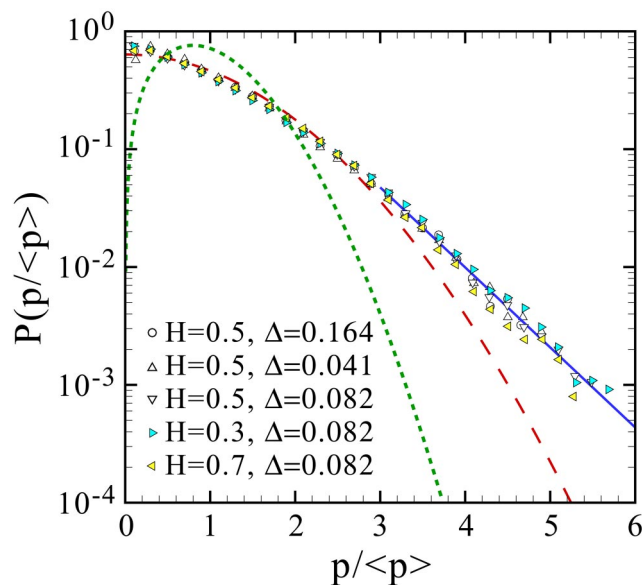
FIG. 15. (Color online) Probability distributions for $p/\langle p\rangle$ at the indicated values of Δ and H all collapse onto a universal curve. Here $\nu=0$ and A/A_0 is between 5% and 10%. The solid line is a fit to the exponential tail of the distribution, the dotted line shows Eq. (14), and the dashed line shows a Gaussian with the appropriate normalization and mean.

that this distribution has a strikingly universal form. Figure 14 shows that the probability $P(p)$ for a contacting node to have local pressure p is independent of the system size. Since the contact area increases linearly with the load, the mean local contact pressure $\langle p\rangle=W/A$ is independent of the contact area, and the entire distribution also remains unchanged for A/A_0 between about 0.01 and 0.1.

Increasing the small scale roughness (Δ) leads to a proportional increase in $\langle p\rangle$. Yet Fig. 15 shows that results for all Δ and H collapse onto a universal function of the dimensionless variable $p/\langle p\rangle$. The probability decreases monotonically with increasing p, and for $p/\langle p\rangle>3$ follows an exponential decay (solid line), $P(p/\langle p\rangle)\propto\exp(-p/p_1)$, with $p_1\approx\langle p\rangle/1.6$. This exponential tail implies that some regions have stresses much higher than $\langle p\rangle$ and may undergo plastic deformation even when the mean stress is much less than the hardness.

Similar universal curves have been found for the stress distribution in a variety of "jammed" systems [22], including granular media [20,21], thermal glasses [31] and polymer crazes [32]. In each case the tail of the distribution follows a simple exponential rather than the Gaussian that might be expected from equilibrium arguments. Several explanations for the exponential form have been proposed [20,21,31,32], but most do not apply to our zero temperature, deterministic elastic system. However, it is possible that the power law correlations in interface height may lead to a hierarchical distribution of the load that is analogous to that in the q model [21].

The distribution of local pressures plays a central role in Persson's theory of contact between self-affine surfaces [8,9]. He defines a resolution ζ corresponding to the number of points along an axis at which the height of the surface is

known, and assumes smooth interpolation between these points. Increasing ζ corresponds to resolving more of the surface roughness and increases Δ if the surface is self-affine. The pressure distribution $P(p,\zeta)$ is a function of both resolution and pressure. Its derivatives satisfy

$$\frac{\partial P}{\partial \zeta}=G'(\zeta)p_0^2\frac{\partial^2 P}{\partial p^2}, \tag{11}$$

where $p_0=W/A_0$ is the apparent mean pressure, primes denote a derivative and

$$G(\zeta)=(E'/p_0)^2\langle|\nabla h|^2\rangle. \tag{12}$$

Persson obtained solutions for $P(p,\zeta)$ in the geometry considered here by starting from perfectly flat planes with $P(p,1)=\delta(p-p_0)$ and iterating to higher resolution. He also imposed the boundary condition that the probability goes to zero at zero pressure [9].

Given the results shown in Fig. 15, it is interesting to ask if Persson's equations have a universal solution in the limit of small loads. Figure 15 shows the probability distribution within the contact, while Persson's $P(p,\zeta)$ includes noncontacting regions and its integral over pressure is the fractional contact area. Thus it should be related to the universal distribution $\widetilde{P}$ by $P(p,\zeta)=\widetilde{P}(p/\langle p\rangle)p_0/\langle p\rangle^2$, where $p_0/\langle p\rangle=A/A_0$. Then use of Eqs. (10) and (11) leads to an equation for $\widetilde{P}$:

$$2\widetilde{P}(x)+x\widetilde{P}'(x)+\kappa^2\widetilde{P}''(x)/4=0. \tag{13}$$

There is a solution with unit norm and mean for Persson's value of $\kappa=\sqrt{8/\pi}$:

$$\widetilde{P}(x)=\frac{\pi}{2}\,x\exp\left(-\frac{\pi}{4}\,x^2\right). \tag{14}$$

As shown in Fig. 15, this solution (dotted line) is much more strongly peaked than our numerical results, decaying to zero linearly in the limit $p\to0$ and as a Gaussian at large p. Note that since the mean and norm are the same for all curves, the presence of extra weight at large p implies more weight at low p. Also shown in Fig. 15 is a pure Gaussian with unit norm and mean (dashed line). This solution provides a better fit to the numerical data at low p. However, as for jammed systems [20–22,31,32], the tail of the true distribution is much closer to a pure exponential than a Gaussian decay. This may reflect correlations in the loads carried by different asperities that are not fully captured in analytic theories [8,9,14,15].

IV. SUMMARY AND CONCLUSIONS

In this paper, we developed a numerical framework for analyzing frictionless, nonadhesive contacts between self-affine surfaces using the finite-element method. This method was applied to perfectly elastic contacts with a range of Poisson ratios, roughness amplitudes and roughness exponents. In each case the real contact area A rises linearly with load W until the fraction of the total area in contact reaches $5-10\%$ (Fig. 5). This implies that the average local pressure in the

contacts, $\langle p \rangle = W/A$, remains constant. The dimensionless pressure $\langle p \rangle / E'$ is independent of the system size even though the large scale roughness grows as L^H.

As predicted by analytic studies [8,15], the dimensionless pressure scales linearly with the small scale roughness Δ. The constant factor κ [Eq. (10)] that relates the roughness to the local pressure is always between the predictions of Bush *et al.* [15] and Persson [8] (Figs. 6–8). A value of $\kappa = 2.2$ reproduces the numerical results within about 10% for $\nu = 0$, and the best fit value rises linearly to about 2.5 as ν rises to the limiting value of 0.5. These results allow the mean pressure and fractional contact area to be predicted for any elastic self-affine surface with known small scale roughness.

The detailed morphology of the contact region and distribution of areas a_c of connected regions were also studied. As in early theories of contact [14,15], the increase in area with the load reflects a linear increase in the total number of contacts with no change in the probability distribution of contact areas. As W increases, each existing contact grows, and new contacts are formed at a rate that maintains a constant $P(a_c)$. At large a_c, the probability distribution falls off as a power law, $P(a_c) \propto a_c^{-\tau}$ (Fig. 10). Since $\tau > 2$ for $H < 0.9$, the mean cluster area is independent of L and comparable to the resolution of the calculation.

The above results for connected clusters are consistent with the conclusion that a large fraction of nodes on a self-affine surface are local maxima that should be treated as asperities [5]. However, recent experimental [2,3] and theoretical [6] papers have suggested a different view. They examined regions where undeformed surfaces would overlap and associated each with a contact. This model gives qualitatively different distributions of areas (Fig. 11). The value of τ is always less than 2, and the mean cluster area diverges as a power of the system size. The geometry of the clusters is also very different. The rigid overlap model gives two-dimensional clusters with fractal perimeters, whereas the full calculation gives fractal cluster areas with the same fractal dimension as the perimeter (Fig. 13). Including regions where the surfaces are separated by less than h_c as part of the contact leads to dramatic changes in the cluster distribution exponent τ and total area. The results approach the overlap distribution when h_c is comparable to the small scale roughness. Optical experiments will detect gaps that are much smaller than the wavelength as in contact and this may explain why small values of τ are observed.

Plastic deformation will occur when the local pressure in a contact exceeds the hardness of the material. The linear relation between mean pressure and small scale roughness [Eq. (10)] can be used to estimate when this will happen. The largest experimental values of p/E' are of order 0.1 and are obtained in amorphous and nanocrystalline materials. Thus Eq. (10) implies that deformation can only be elastic when $\sqrt{\langle |\nabla h|^2 \rangle} < 0.1 \kappa \sim 0.2$. This condition is violated for many surfaces, and the much smaller hardness of macroscopic crystals will lead to even tighter constraints on the roughness. Our approach is readily extended to include plastic deformation, which will be the subject of future work.

Plastic deformation may occur well before the mean pressure reaches the hardness because some nodes have local pressures much larger than $\langle p \rangle$. Results for all parameters collapse onto a universal probability distribution $\tilde{P}(p/\langle p \rangle)$ (Fig. 15). Persson has presented approximate analytic equations for the pressure distribution. This analytic distribution drops as a Gaussian at large p, while the numerical results have an exponential tail that greatly increases the number of sites with large pressures. Similar exponential distributions are found in many jammed systems such as sand piles, glasses and crazes [20–22,31,32]. A common feature of these systems is a highly nonuniform distribution of stress. It is possible that the presence of small bumps on bigger bumps on still bigger bumps in our systems leads to a transmission of stress like that in the q model for sand piles [20,21]. This hierarchical structure may produce stress correlations that are not included in the analytic model [8].

There are many interesting avenues for future research. The approach outlined here can readily be extended to include more complex surface morphologies, plasticity, interfacial friction and tangential loading of the solids. More challenging issues include adhesion and the role of atomic scale roughness. These issues will require hybrid algorithms that include atomic information about interfacial interactions.

---

[1] F. P. Bowden and D. Tabor, *The Friction and Lubrication of Solids* (Clarendon, Oxford, 1986).

[2] J. H. Dieterich and B. D. Kilgore, Pure Appl. Geophys. **143**, 283 (1994).

[3] J. H. Dieterich and B. D. Kilgore, Tectonophysics **256**, 219 (1996).

[4] P. Berthoud and T. Baumberger, Proc. R. Soc. London, Ser. A **454**, 1615 (1998).

[5] J. A. Greenwood, Proc. R. Soc. London, Ser. A **393**, 133 (1984).

[6] J. A. Greenwood and J. J. Wu, Meccanica **36**, 617 (2001).

[7] A. Volmer and T. Natterman, Z. Phys. B: Condens. Matter **104**, 363 (1997).

[8] B. N. J. Persson, Phys. Rev. Lett. **87**, 116101 (2001).

[9] B. N. J. Persson, F. Bucher, and B. Chiaia, Phys. Rev. B **65**, 184106 (2002).

[10] G. G. Batrouni, A. Hansen, and J. Schmittbuhl, Europhys. Lett. **60**, 724 (2002).

[11] M. Ciavarella, G. Demelio, J. R. Barber, and Y. H. Jang, Proc. R. Soc. London, Ser. A **456**, 387 (2000).

[12] S. Roux, J. Schmittbuhl, J.-P. Vilotte, and A. Hansen, Europhys. Lett. **23**, 277 (1993).

[13] K. L. Johnson, *Contact Mechanics* (Cambridge University Press, New York, 1985).

[14] J. A. Greenwood and J. B. P. Williamson, Proc. R. Soc. London, Ser. A **295**, 300 (1966).

[15] A. W. Bush, R. D. Gibson, and T. R. Thomas, Wear **35**, 87 (1975).

[16] B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, New York, 1979).

[17] J. Krim and G. Palasantzas, Int. J. Mod. Phys. B **9**, 599 (1995).

[18] P. Meakin, *Fractals, Scaling and Growth Far From Equilibrium* (Cambridge University Press, Cambridge, 1998).

[19] M. Borri-Brunetto, B. Chiaia, and M. Ciavarella, Comput. Methods Appl. Mech. Eng. **190**, 6053 (2001).

[20] C. h. Liu, S. R. Nagel, D. A. Schecter, S. N. Coppersmith, S. Majumdar, O. Narayan, and T. A. Witten, Science **269**, 513 (1995).

[21] S. N. Coppersmith, C. h. Liu, S. Majumdar, O. Narayan, and T. A. Witten, Phys. Rev. E **53**, 4673 (1996).

[22] *Jamming and Rheology*, edited by A. J. Liu and S. R. Nagel (Taylor and Francis, London, 2001).

[23] R. F. Voss (Springer-Verlag, Berlin, 1985), p. 805.

[24] J. F. Molinari and M. Ortiz, Int. J. Numer. Methods Eng. **53**, 1101 (2002).

[25] T. Belytschko and T. J. R. Hughes, *Computational Methods for Transient Analysis* (North-Holland, Amsterdam, 1983).

[26] T. J. R. Hughes, *The Finite-Element Method: Linear Static and Dynamic Finite-Element Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1987).

[27] J. F. Molinari, M. Ortiz, R. Radovitzky, and E. A. Repetto, Eng. Comput. **18**, 592 (2001).

[28] J. A. Harrison, S. Hyun, J. F. Molinari, and M. O. Robbins (unpublished).

[29] We also considered connections between next-nearest neighbors. This increases the cluster sizes, leading to a distribution much like that for $h_c=0.025$ (Fig. 11). However, it does not change the fundamental difference between clusters produced by the full calculation and those from the rigid overlap model. Note that including next-nearest neighbor connections has no noticeable effect on $P(a_c)$ in the rigid overlap model.

[30] J. Kondev and C. L. Henley, Phys. Rev. Lett. **74**, 4580 (1995).

[31] C. S. O'Hern, S. A. Langer, A. J. Liu, and S. R. Nagel, Phys. Rev. Lett. **86**, 111 (2001).

[32] J. Rottler and M. O. Robbins, Phys. Rev. Lett. **89**, 195501 (2002).