Rock Mech Rock Eng (2017) 50:2891–2909
DOI 10.1007/s00603-017-1281-4

ORIGINAL PAPER

Comparison of Surface Properties in Natural and Artificially Generated Fractures in a Crystalline Rock

Daniel Vogler[1] · **Stuart D. C. Walsh**[2] · **Peter Bayer**[1,3] · **Florian Amann**[1]

Received: 14 July 2016 / Accepted: 11 July 2017 / Published online: 25 July 2017

Abstract This work studies the roughness characteristics of fracture surfaces from a crystalline rock by analyzing differences in surface roughness between fractures of various types and sizes. We compare the surface properties of natural fractures sampled in situ and artificial (i.e., man-made) fractures created in the same source rock under laboratory conditions. The topography of the various fracture types is compared and characterized using a range of different measures of surface roughness. Both natural and artificial, and tensile and shear fractures are considered, along with the effects of specimen size on both the geometry of the fracture and its surface characterization. The analysis shows that fracture characteristics are substantially different between natural shear and artificial tensile fractures, while natural tensile fracture often spans the whole result domain of the two other fracture types. Specimen size effects are also evident, not only as scale sensitivity in the roughness metrics, but also as a by-product of the physical processes used to generate the fractures. Results from fractures generated with Brazilian tests show that fracture roughness at small scales differentiates fractures from different specimen sizes and stresses at failure.

Keywords Fracture roughness · Surface topography · Roughness scale effects · Natural fracture

✉ Daniel Vogler
davogler@ethz.ch

[1] ETH Zurich, Sonneggstr. 5, 8092 Zurich, CH, Switzerland

[2] UNSW, Sydney, Australia

[3] Technische Hochschule Ingolstadt, Ingolstadt, Germany

1 Introduction

Correctly characterizing the mechanical and hydraulic properties of rock fractures, and the coupling between these properties is a crucial part of many subjects in the applied geosciences. Examples include, but are by no means limited to: estimating the productivity of oil, gas and geothermal reservoirs (Rutqvist and Stephansson 2003), understanding microseismic events in faults and joints (Derode et al. 2013), exploring the efficiency of groundwater remediation, and predicting radionuclide migration in the excavation damage zone around nuclear waste repositories (Zhu et al. 2007; Bear et al. 2012). Fracture topography is a major determinant for the permeability and stiffness of rock in deep geological systems. However, due to the variability of natural media, limited accessibility, and the influence of transient mechanical deformation, exact characterization of in situ fractures is often difficult, if not impossible.

Instead, laboratory-scale investigations are frequently employed to provide insight into characteristic values and ranges of fracture properties from which model parameters can be derived. In particular, in many cases where natural fractures cannot be sampled directly, fractures are generated through artificial means—for example, by performing a Brazilian test on an intact specimen. However, the properties of fractures used in laboratory studies depend on many factors: specimen size, fracture origin, stress path and sampling methods—which poses the question: What experimental bias is introduced in such artificially generated fracture surfaces?

Fractures generated under laboratory conditions are often considered as a proxy for in situ fracture geometries (Witherspoon et al. 1980; Esaki et al. 1999; Nicholl et al. 1999; Belem et al. 2000; Lee and Cho 2002; Jiang et al.

2006; Watanabe et al. 2008; Nemoto et al. 2009; Xiong et al. 2011; Faoro et al. 2012; Li et al. 2014; Lee et al. 2014). However, the methods used to produce such "artificially" generated fractures may affect the surface characteristics and thereby alter the associated flow and stress-response properties away from those of their "natural" (i.e., in situ) fracture counterparts. These differences can arise due to: differences in the forces used to generate the fractures from those experienced in the field; differences in weathering or erosion experienced by the fracture surfaces; and finally, the very size of the specimen itself may bias both the creation and the characterization of the fracture geometry.

In this study, we compare the surface properties of natural fractures sampled in situ and artificial (i.e., man-made) fractures created in the same source rock under laboratory conditions. The topography of the various fracture types is compared and characterized using a range of different measures of surface roughness. Both natural and artificial, and tensile and shear fractures are considered, along with the effects of specimen size on both the geometry of the fracture and its surface characterization.

The relationship between fracture surface topology and the fracture characteristics such as fracture conductivity and shear and normal stiffness has been investigated at the laboratory scale in several experimental studies (Yeo et al. 1998; Jiang et al. 2006) and numerical simulations (Brown 1987; Xiong et al. 2011) attempting to replicate natural environments as closely as possible. While some studied natural fractures extracted by over-coring existing joints (Chen et al. 2000; Vogler et al. 2016a) or employed synthetic surfaces created by extrapolating characteristics of natural fractures (Ogilvie et al. 2003), others used artificial fractures obtained by splitting (i.e., tensile or Mode I fractures) (Watanabe et al. 2008; Nemoto et al. 2009; Xiong et al. 2011; Faoro et al. 2012; Vogler et al. 2016b), or shear fracturing solid rock (i.e., shear or Mode II fractures) (Esaki et al. 1999; Watanabe et al. 2009; Li et al. 2014). As such a broad range of different fracture types is used to investigate the physical behavior of rock properties, it is important to understand how the surface topography differs between these fracture types to better gauge the effects these differences might have on the measured quantities.

Much of the diversity among natural and artificial mode I and II fracture topologies can be attributed to the complexity of fracture formation, which involves micro-crack growth and micro-crack coalescence (Mosher et al. 1975; Kranz 1983; Mardon et al. 1990; Fujii et al. 2007). The form of crack growth mechanism—intragranular, intergranular or transgranular crack growth—depends on rock composition/minerology, stress field (normal or shear stresses as well as stress magnitude) and specimen dimensions. The close relation between crack growth and surface topography (Mosher et al. 1975; Kranz 1983; Morgan et al. 2013) highlights the importance of the link between crack growth mechanism and stress field, fracture mode and specimen size. For example, Mosher et al. (1975) analyzed both tensile and shear fractures microscopically. He showed that tensile fracturing in granite is dominated by intergranular over intragranular fractures. He also showed that during tensile fracturing of fine-grained granite almost twice as many fractures formed as compared to medium grain-sized granite. In compression, the fracturing processes change and the majority of induced fractures are intragranular fractures. Kranz (1983) studied micro-cracking processes under compressive loading conditions in Barre Granite and found that the average crack length increases with rising applied stress. Furthermore, the ratio of grain boundary cracks to intragranular cracks decreases with higher uniaxial compressive stress, resulting in smoother fracture surfaces. The above studies (Mardon et al. 1990 in particular) show that on the grain-scale microcracks can grow in all directions within a macroscopic fracture plane. They also demonstrate that fracturing process and fracture propagation path depend on the mineralogical composition, the grain size of the minerals and the loading conditions that further affect the fracture topography and roughness. Morgan et al. (2013) found, for example, that natural shear fractures in granite are less rough than tensile fractures. These differences in roughness are associated with the crack propagation pathway. In case of shear fractures, fracture formation is dominated by grain breakage, while for tensile fractures, crack propagation follows predominantly grain boundaries.

Understanding the influence of fracture formation processes on the surface topography and associated fracture properties, and the process of upscaling these properties requires a description of fracture surfaces at various scales and various origins (i.e., tensile and shear fractures). For example, surface roughness was characterized with individual methods such as JRC and Z2 values by Tse and Cruden (1979) and Yu and Vayssade (1991), fractal measures by Power and Durham (1997) and Pyraknolte et al. (1992). Lee et al. (1990) and Li and Huang (2015) related the joint roughness coefficient (JRC) to fractal measures, while Tatone and Grasselli related the JRC to another 2D roughness measure (Tatone and Grasselli 2010). Nevertheless, these studies focus on a single description parameter (e.g., JRC, fractals or other parameters) and do not attempt an extensive evaluation of roughness measures or fracture nature (Pyraknolte et al. 1992; Power and Durham 1997; Babadagli and Develi 2000; Tatone and Grasselli 2010). Moreover, these prior studies used fracture surfaces of varying dimensions or only conducted their roughness analysis on a very limited number of fractures (Huang et al.

1992; Belem et al. 2000; Fardin et al. 2001; Jiang et al. 2006; Tatone and Grasselli 2010, 2013), making it difficult to determine which parameters are capable of differentiating between fracture modes and size.

The objective of this work is a comparison of a large number of fracture surfaces (i.e., natural tensile and shear as well as artificial tensile fractures) with numerous surface roughness measures. This contributes valuable insights into the relationship between fracture formation process, origin (e.g., natural or artificial), mode type (e.g., tensile or shear) and specimen size, and the observed fracture characteristics of a given rock type. By contrasting these factors all in one study, most critical aspects for experimental fracture characterization can be identified. A main question will be whether (and to what extent) the origin and mode of fracture are reflected in the surface characteristics. Given the large number of previous studies performed on a wide range of fracture types, this will improve comparability of the findings from different studies. Furthermore, this will facilitate to optimize experimental fracture characterization, and this will ultimately support better transferability of experimental results to the field.

In the following sections, we first describe the sampling, preparation and scanning of the rock specimens. The fracture surfaces are then analyzed in terms of fracture roughness (JRC, Z2) and fracture topography (fractal dimensions, correlation functions). Special attention is given to the role of specimen size and scale effects associated with the surface measures. Finally the correspondence between the different surface measures and their variation across the different classes of fracture surfaces is discussed.

2 Methods and Analysis

2.1 Fracture Sampling

The specimens used to investigate the natural and artificial fractures in crystalline rock were obtained from the CRIEPI fractured rock study (Takana et al. 2014) at the Grimsel Test Site (GTS), Switzerland. The rock specimens were sourced from core material extracted from seven boreholes each with a diameter of 10.2 cm. The core material consists of granodiorite with uniformly distributed grains with sizes between 3 and 7 mm. One set of natural fractures was obtained by over-coring pre-existing fractures in the original 10.2 cm diameter cores with a 2.5 cm drill bit (denoted by EP_i in Table 1). The surface profiles from this first set of natural fractures were recorded as part of an earlier study of hydro-mechanically coupled processes in natural fractures (Vogler 2016; Vogler et al. 2016a). Additional natural fractures of varying size were sampled without subcoring the original core material (denoted by NG_i in Table 1). All of the natural fractures were classified as either tensile (Mode I) or shear (Mode II) fractures based on an inspection of their surface characteristics (Table 1). The artificial fractures were created by subjecting intact cylinders of the remaining core material to a Brazilian-strength test (Hatheway 2009). This produces tensile fractures that are formed in the cylinders when the stresses applied by the Brazilian tests exceed the tensile strength of the rock. To capture the effects of total fracture length and specimen size on the surface parameters, a range of core sizes was used to generate these "artificial" fractures: specimen cores were created with diameters of 2.5, 5.1, 10.2 and 30 cm. The specimens with 2.5 cm diameter and 6 cm length (denoted $A_{2.5}$) were obtained to provide a direct comparison to natural specimens EP_i used in earlier core holder experiments (Vogler et al. 2016a). The other artificial specimens with 5.1 cm ($A_{5,\mathrm{std}}$), 10.2 cm ($A_{10,\mathrm{std}}$) and 30 cm ($A_{30,\mathrm{std}}$) diameters were created with dimensions in accordance with standardized Brazilian test procedures (Hatheway 2009). Additional specimens with a diameter of 10.2 cm (A_{10}) and 17 cm length were fractured to obtain fractures with larger surface areas. Finally, specimen with 2.5 cm diameter and 1.25 cm length ($A_{2.5,\mathrm{std}}$) was fractured in Brazilian tests to investigate fracture paths on a subgrain scale. It should be noted that although the artificial specimens with 2.5 cm diameter were obtained using standard sampling aspect ratio, they do not constitute a standard Brazilian test due to the small specimen size. In Table 1, they are denoted with $A_{2.5,\mathrm{std}}$ as the cylinder length was chosen according to standard. In total, more than 60 fractures with sizes ranging from 1 to 40 cm edge length were scanned and analyzed under this study: a summary of their properties is provided in Table 2. Example pictures showcasing the size differences between the artificial tensile specimens and natural shear and tensile fractures are shown in Fig. 1.

Table 1 Characterization of fracture origination by Mode I (tensile), Mode II (shear) and Mode I* (tensile with partial shear)

Sample	Mode	Sample	Mode	Sample	Mode	Sample	Mode
NG_{01}	II	NG_{41}	I	EP_{13}	I	EP_{23}	I
NG_{02}	I	NG_{42}	I	EP_{14}	I	EP_{24}	I
NG_{05}	II	EP_{01}	I	EP_{15}	I	EP_{25}	II
NG_{06}	II	EP_{02}	I	EP_{16}	I	EP_{26}	II
NG_{07}	I	EP_{03}	I*	EP_{17}	I	EP_{27}	I
NG_{08}	I	EP_{04}	II	EP_{18}	II	EP_{28}	I
NG_{09}	I	EP_{06}	I	EP_{19}	I*	EP_{29}	II
NG_{10}	I	EP_{10}	I	EP_{20}	I		
NG_{31}	II	EP_{11}	I	EP_{21}	II		
NG_{32}	II	EP_{12}	I	EP_{22}	I		

Table 2 Summary of core specimens, indicating nature of fracture (natural or artificial), diameter D, length L, approximate fracture area A, whether artificial fractures from Brazilian test were created according to standard (BT-std) (Hatheway 2009) and the number of specimens in each category n

Samples	Type (–)	Diameter (cm)	Length (cm)	Area (cm^2)	BT-std (–)	Number (–)
$N_{2.5}$	Natural	2.5	6.0	15.2	–	25
N_L	Natural	Variable	Variable	Variable	–	12
$A_{2.5,\mathrm{std}}$	Artificial	2.5	1.25	3.1	No	5
$A_{2.5}$	Artificial	2.5	6.0	15.2	No	5
$A_{5,\mathrm{std}}$	Artificial	5.1	2.5	12.9	Yes	5
$A_{10,\mathrm{std}}$	Artificial	10.2	5.1	51.6	Yes	5
$A_{10,17}$	Artificial	10.2	17.0	172.7	No	2
$A_{30,\mathrm{std}}$a	Artificial	30.0	15.0	450.0	No	2

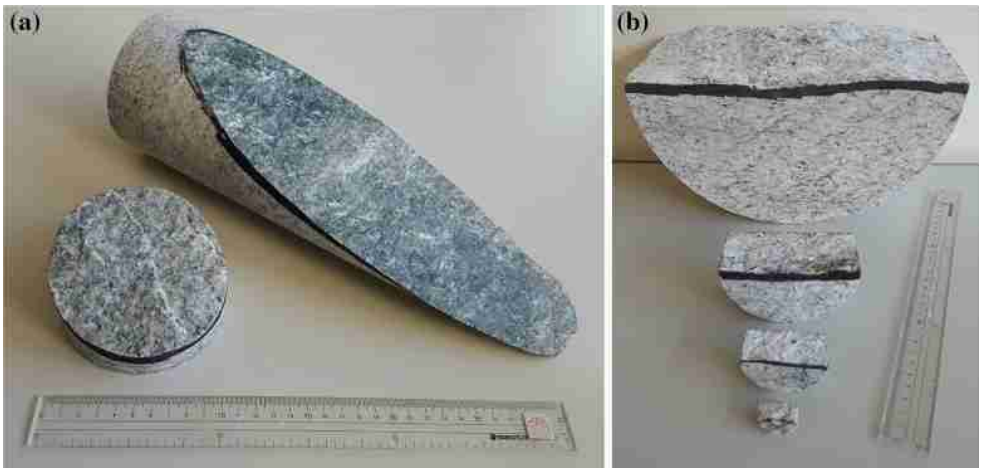


Fig. 1 Example specimen: **a** surfaces for shear (*back right*) and tensile (*front left*) natural fracture; **b** artificial fractures obtained by Brazilian tensile tests with specimen diameters of 30.0, 10.2, 5.1 and 2.5 cm (*back to front*)

2.2 Photogrammetric Scan

The fracture surfaces were evaluated with surface scans using the ATOS Core 3D scanner from GOM (GOM mbH 2015). The ATOS Core sensor projects fringe patterns on the object surface, which are recorded by two cameras. The patterns form a phase shift based on a sinusoidal intensity distribution that enables measurement of the 3D surfaces. The ATOS Core was calibrated with length deviation errors between 0.009 and 0.027 mm and optimized calibration deviations of 0.014 ± 0.001 Pixel. This means that for measurements of standardized objects (e.g., the diameter of a perfectly round sphere, or the distance between two spheres which are mounted on one bar), the geometric measure in question (e.g., diameter or distance) can be measured with accuracies between 9 and 27 µm. For more complex surfaces, accuracy varies, as different areas of a surface may reflect the projected fringe patterns differently, due to the angle of incoming light or changes in reflectivity. To address the issue of different reflectivities among minerals, the fracture surfaces were coated with a white spray before scanning, which adds a few micrometers in thickness.

Surface profiles produced by the scanner are represented as three-dimensional point clouds. The original resolution of the scanned fracture surfaces (without material on the sides of the specimens) depends slightly on the individual surface and is roughly between 30 and 60 unique vertices per square millimeter. Post-processing was conducted on the scans to remove any extraneous details from either the specimen edges or the specimen holder, and to align the axes of the specimen in the x–y plane, with the normal of the fracture surface oriented along the z-axis. Here, the x–y plane is defined as the plane providing a least square fit for any given fracture surface. The best-fit plane through the surface was found by aligning the eigenvectors of the plane with the primary coordinate system. The orientation of the uppermost surface was inverted to allow the fracture aperture to be calculated directly. After the fracture roof and floor were properly aligned, the surfaces were mapped onto a regular grid with a resolution of 50 µm along both the x–y axes of the plane. This gridded representation of the

surface was then employed to calculate surface roughness measures.

Examples of the three fracture classes investigated (natural tensile & shear and artificial tensile) are shown in Fig. 2. Figure 2a shows a natural shear fracture, displaying long range variation, with few short wave oscillations. The natural tensile fracture in Fig. 2b displays the individual grain or conglomerate sizes, with three visible asperity peaks and four areas of low asperity height, representative of large asperity height on the opposite fracture surface. The artificial tensile fracture (Fig. 2c) displays a lot of short wave length oscillations, with long wave lengths

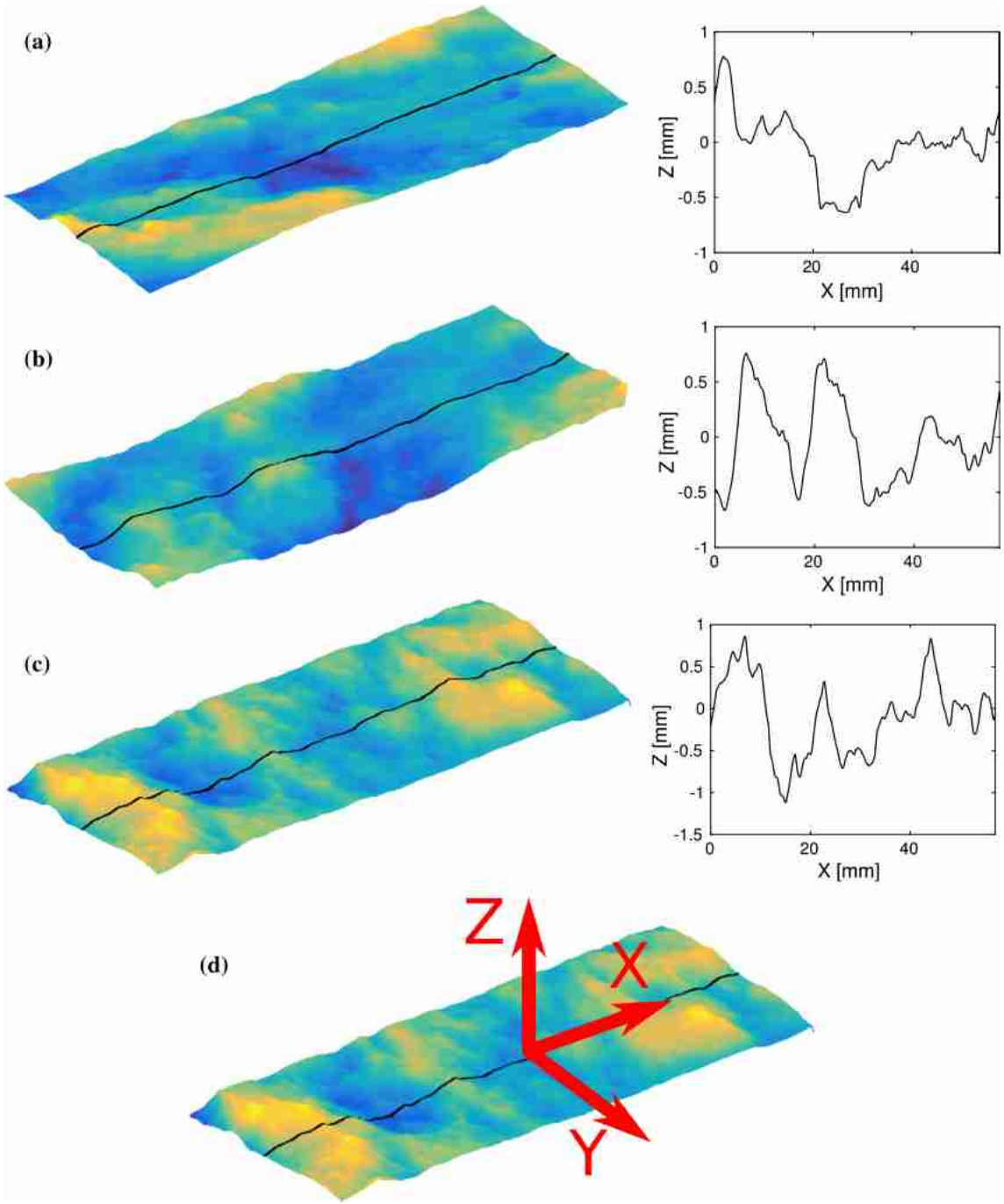


Fig. 2 Examples of typical scans and cross sections: **a** natural fracture surface (EP_{04}—shear); **b** natural fracture surface (EP_{20}—Mode I); **c** artificial fracture surface ($A_{2.5,6\,cm}$ specimen 6—Mode I); **d** example specimen illustrating the best-fit plane (x–y plane) and the direction of asperities (z-direction). The *presented scan lines* do not represent the sampling procedure employed during roughness computation, which was instead based on a large number of *scan lines* covering each surface. *Color* coding for the fracture surfaces corresponds to the maximum and minimum asperity height, respectively. The fracture width in y-direction is 25 mm for **a**–**c** (color figure online)

underlying the scan line. Scan lines similar to Fig. 2a–c were found across many fracture surfaces, indicating differences in surface roughness due to variable fracture propagation mechanisms. It should be noted that the presented scan lines (Fig. 2) across the surfaces do not represent the sampling employed during roughness computation. Instead, a large number of scan lines covering the entirety of the surface in both the x- and y-directions were used to compute the roughness measures presented in this study. The average values computed across all scan lines are presented here for brevity. While some directional dependence of roughness on the scan line direction was found, the effect was small in comparison to the differences observed between specimen of different fracture type and size, and is therefore not discussed in detail in this study.

2.3 Roughness Measures

Two common measures of surface roughness are the standard deviation of the asperity heights (STD) and the joint roughness coefficient (JRC). As discussed herein and elsewhere (Fardin et al. 2001; Vogler et al. 2017), while simple, the standard deviation suffers from scale dependence if not pegged to an underlying length scale. The JRC or joint roughness coefficient is commonly used to characterize rock surface properties and correlate the roughness of a fracture with its mechanical and hydraulic properties (Barton and Choubey 1977). Properly applied, the JRC value has an implicit associated scale, as the JRC value is evaluated by comparing the specimens of prescribed length (10 cm) to a hierarchy of "characteristic" surface profiles. Nevertheless, the arbitrary nature of this comparison makes the JRC value a somewhat qualitative measure of roughness. Instead, several groups have sought to correlate the JRC scale to more quantitative metrics and different empirical approaches have been outlined for estimating JRC values. From these, we chose the Z2 measure (Tse and Cruden 1979; Yu and Vayssade 1991) originally proposed by Myers (1962). The dimensionless Z2 roughness measure is defined as:

$$Z2 = \left[\frac{1}{L}\int_0^L \left(\frac{\mathrm{d}z}{\mathrm{d}x}\right)^2 \mathrm{d}x\right]^{1/2}, \tag{1}$$

where $z(x)$ is the profile height of the fracture surface. In practice, the Z2 value is determined from a discretization of the surface profile

$$Z2 = \left[\frac{1}{L}\sum \frac{(z_{i+1}-z_i)^2}{x_{i+1}-x_i}\right]^{1/2}, \tag{2}$$

where x_i and z_i are the coordinates of the fracture surface, typically taken at regular sampling intervals, Δx, such that $x_{i+1} - x_i = \Delta x$ for all i and L is the total length of a scan line along which Z2 is measured. As with the standard deviation, roughness parameters like the Z2 measure are contingent upon the scale at which they are measured. One method to overcome this sensitivity is to fix the sampling interval. However, the choice of interval is somewhat arbitrary. Yu and Vayssade (1991) derived empirical equations relating Z2 and JRC values using sampling intervals of 0.25, 0.5 and 1.0 mm with the recommendation that the shortest interval should be used whenever possible. Accordingly, for this paper, a sampling interval of 0.25 mm was used to determine all Z2 values, and the JRC value was then determined from:

$$\begin{aligned} \mathrm{JRC} &= 60.32 \\ &gtZ2 - 4.51. \end{aligned} \tag{3}$$

Fixing the sampling interval filters out geometric characteristics below the sampling length. It might be speculated that crucial information is lost as a result, and thus a fractal measure of the surface (i.e., one capable of recognizing self-similarity in the surface properties across length scales) may better serve to distinguish different surface types. Fractal dimensions are frequently used to characterize surface roughness, though there is debate over their efficacy, particularly if used as a sole measure to characterize the surface (Huang et al. 1992). Thus, estimates of the fractal Hausdorff and Box count dimensions were obtained for each surface to investigate their ability to distinguish the different fracture types. Using the Hausdorff dimension, we obtain the fractal dimension d_{HD}, for the number of line subsets N_{HD} of length l_{HD} that are required to constitute the scan line segment S_{SL}.

$$d_{\mathrm{HD}} = -\lim_{l_{\mathrm{HD}}\to 0} \frac{\ln(N_{\mathrm{HD},l_{\mathrm{HD}}}(S_{\mathrm{SL}}))}{\ln(l_{\mathrm{HD}})}. \tag{4}$$

While the equation for the Box count dimension (Eq. 5) is fundamentally the same to the Hausdorff dimension (Eq. 4), the measuring (e.g., counting) approach differs. A grid of boxes is overlaid with the scan line, and the number of occupied grid cells N_{BC} (e.g., boxes) needed to cover the scan line segment S_{SL} (with S_{SL} being a non-empty bounded subset of $\mathbb{R}^n$) is counted for a given box side length l_{BCD}

$$d_{\mathrm{BC}} = -\lim_{l_{\mathrm{BCD}}\to 0} \frac{\ln(N_{\mathrm{BCD},l_{\mathrm{BC}}}(S_{\mathrm{SL}}))}{\ln(l_{\mathrm{BCD}})}. \tag{5}$$

The above roughness measures all provide a single scalar value describing the surface—they give little detail regarding the spatial relationships of different surface features. To obtain a richer description of the surface features, we use two-point correlation functions (namely the two-point probability function and the lineal-path function

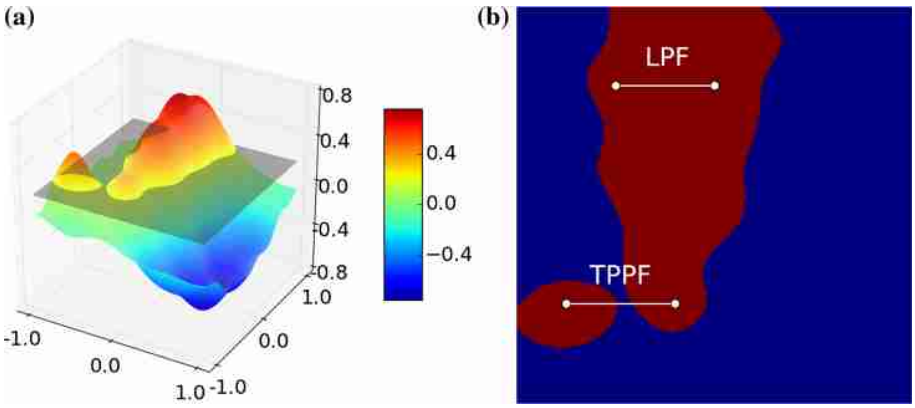


Fig. 3 Two-point correlation functions are used to characterize the distribution of surface features as follows: **a** first the surface is cut by a plane at a given height relative to the median asperity height. **b** The two-point probability function (TPPF) gives the probability that two points within the cutting plane separated by a given distance fall inside the fracture surface, and the lineal-path function (LPF) gives the probability that the entire path between the two points lies within the fracture surface

(Jiao et al. 2007)) to characterize the geometric distribution of surface asperities. Under the approach considered here, the surfaces are transformed such that a best-fit plane through the surface lies in the x–y plane, with a mean z value (e.g., asperity height) of zero. By shifting the intersecting plane up and down the z-axis, the asperity distribution is sampled at different reference heights (Fig. 3). For a given reference height (z value), the two-point correlation functions are used to describe the spatial distribution of the intersecting asperities. Specifically, the two-point probability function $\mathrm{TPPF}^{(i)}(\mathbf{x}_1, \mathbf{x}_2)$ denotes the probability that two points on the intersecting plane $\mathbf{x}_1$ and $\mathbf{x}_2$ are both inside the fracture surface (Fig. 3b). The lineal-path function $\mathrm{LPF}^{(i)}(\mathbf{x}_1, \mathbf{x}_2)$ measures the probability that the entire line segment S between the two points $\mathbf{x}_1$ and $\mathbf{x}_2$ lies within the asperities, without intersecting with the x–y fracture surface (Fig. 3b). The surfaces are characterized by recording the two-point correlation functions for a range of different line segment lengths ($l = |\mathbf{x}_1 - \mathbf{x}_2|$) at a set of predetermined surface heights (z). To investigate correlation of surface topography independently of total fracture size, smaller parts of the fracture surfaces were subsampled to study scale-independent effects.

3 Results and Discussion

We first compare the chosen surface roughness measures applied to the scanned fracture surfaces. The results are further scrutinized for differences in roughness based on fracture size. Subsequently, a scale-independent measure is elaborated, to examine if scale effects (introduced by fracturing specimens of different size) can be captured within a small sampling windows on the fracture surfaces. To highlight the crucial role of fracture formation process when discussing scaling effects, the impact of tensile stress during crack growth on the resulting surface roughness is inspected. Finally, the suitability of correlation functions to capture spatial roughness on small scales is discussed.

3.1 Surface Roughness

Figure 4 shows a scatter plot comparing the standard deviations of surface heights (based on the entire fracture surface) with the Z2 measure from Eq. (2). For reference, the corresponding JRC values obtained from Eq. (3) are given on the upper axis. It is evident from the plot that the standard deviation of asperity heights (calculated from the total fracture area) does not readily distinguish the three fracture classes: while the tensile fractures have some outliers between 1.2 and 3 mm, the values for all three cases are mainly distributed between 0.3 and 1.2 mm. As

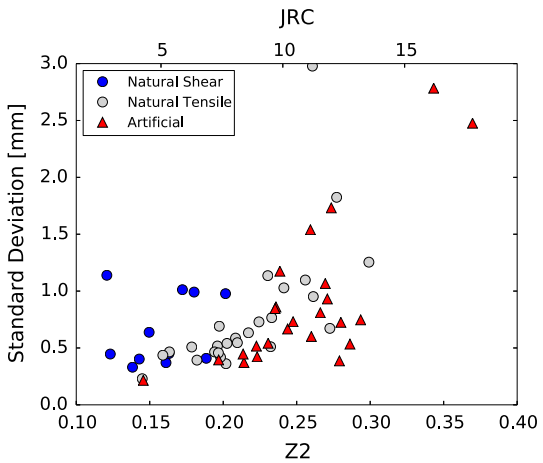


Fig. 4 Standard deviation of surface height versus Z2 measure and JRC value as obtained from Eq. (3) for natural shear fractures (*blue circles*), tensile fracture (*gray circles*) and artificial (*red triangles*) fractures (color figure online)

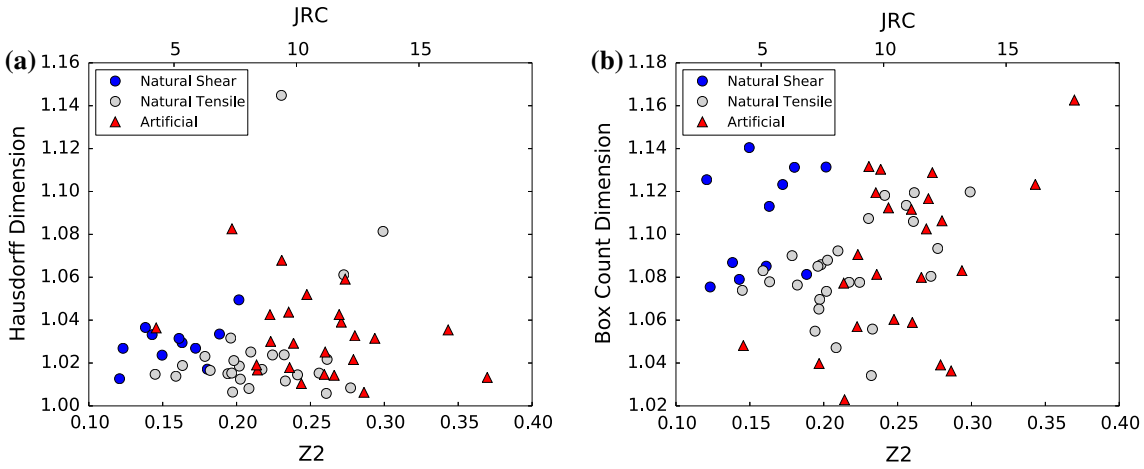


Fig. 5 Fractal dimension measures versus Z2 measure from the natural (*blue* shear, *gray* tensile) and artificial (*red*) fracture surfaces: **a** Hausdorff dimension; **b** Box count dimension (color figure online)

will be discussed in greater detail below, the standard deviation fails to distinguish between the different fracture classes due the size dependence of the measure combined with a lack of an intrinsic length scale. However, the same is not true of the Z2 values. Artificial fractures tend to have higher Z2 values, while the natural fractures are somewhat uniformly distributed over a wide range from 0.12 to 0.3. This indicates that natural surfaces tend to be smoother than artificial ones, and this is in line with the examples shown in Fig. 2. When separating natural tensile and shear fractures, natural shear fractures provide lower Z2 values (0.12–0.21), indicating the smoothest surfaces. Thus, the Z2 measure appears to provide a method to distinguish fracture surface characteristics in such fractures sets, especially between the two extremes, the diverse and relatively rough artificial tensile fractures and the smoother natural shear fractures. This differentiation between different fracture types is especially remarkable, given that crack paths and thus created fractures yield surfaces which are each highly unique.

Less frequent small-scale oscillations in natural fractures are likely attributed to abrasion and small-scale damage after fracturing. Stress changes (loading and unloading of the fractures), changes of the directions of principal stresses, small shear deformation and other weathering processes could lead to substantial abrasion of fracture surface asperities, and thereby lower quantitative surface roughness measures. Besides these abrasive processes, which are difficult to reconstruct or quantify, fracture surface topography could also be influenced by different crack growth mechanisms during fracturing, which will be discussed further below.

The higher Z2 values of artificial fractures result mainly from small-scale oscillations in form of individual asperities, which cause larger surface roughness on that scale. However, these asperities may have a low shear strength and could fail at low shear stress. While natural shear fractures are smooth (indicating low shear strength) on the small scale, stronger large-scale variations increase shear strength more significantly on a larger scale. This could be caused by crack formation, which may travel around large conglomerates or geological formations, if this proves to be the path of least resistance. Due to the confined nature in scale and stress field of fractures created with Brazilian tests, these longer wave length roughnesses are observed less frequently for artificial specimens.

In Fig. 5, the Z2 (and JRC) value of each fracture surface is plotted versus the fractal Hausdorff dimension and the box count dimension of the surface. Results for the Hausdorff dimension yield a fractal dimension between 1.00 and 1.05 for all surface areas and fracture modes (Fig. 5a). Individual ranges are 1.01–1.05 for natural shear, 1.005–1.15 for natural tensile and 1.005–1.08 for artificial tensile fractures. The Box Count dimension results in a wider distribution of fractal dimension with values between 1.01 and 1.13 (Fig. 5b). Here, fractal dimension ranges from 1.07 to 1.14 for natural shear, 1.03–1.12 for natural tensile and 1.02–1.16 for artificial tensile fractures. For both fractal measures, the natural shear fractures show a more narrow distribution of the fractal dimension. However, the plot reveals no evident correlation between the fracture class and either measure of fractal dimension.

Nevertheless, particularly with large numbers of data points, visual inspection may not be sufficient to reveal a correspondence between input parameters and categories. Thus, a Mann-Whitney U test (Mann and Whitney 1947) was applied to obtain a more quantitative answer to the question, as to whether natural and artificial fractures show significant differences for JRC, Z2, Box count and Hausdorff dimension (Figs. 4, 5). The findings for this test are

Table 3 Mann–Whitney *U* test results, indicating the probability that the roughness parameters of the natural and artificial fractures are drawn from the same population

Metric	*p* value		
	NT/AT (–)	NS/AT (–)	NS/NT (–)
STD	2.4×10^{-1}	1.7×10^{-1}	2.0×10^{-1}
Z2/JRC	1.3×10^{-3}	5.7×10^{-6}	6.2×10^{-5}
Hausdorff	1.2×10^{-2}	3.9×10^{-1}	6.1×10^{-3}
Box Count	2.6×10^{-1}	7.0×10^{-2}	8.0×10^{-3}

Compared are NT/AT—natural tensile versus artificial tensile fractures; NS/AT—natural shear versus artificial tensile fractures; NS/NT—natural shear versus natural tensile fractures

listed in Table 3. A smaller *p* value indicates a lower likelihood that a roughness measure distribution of a given fracture type could be obtained with another fracture type as well (i.e., the smaller the *p* value, the greater the ability of the metric to distinguish between classes). Values are calculated for all pairings between Natural Tensile (NT), Natural Shear (NS) and Artificial Tensile fractures (AT). Especially large *p* values are obtained for the STD for NT/AT, the Hausdorff dimension for NS/AT and the Box count dimension for NT/AT. This result can be compared visually with Figs. 4 and 5, where ranges of the roughness measures for the respective specimen types are virtually indistinguishable. Likewise, the *p* values show small differences between the three specimen types for the fractal dimensions. This supports the findings that the roughness measure Z2 (and by implication the derived JRC measure) are most suitable for distinction of the three fracture classes.

3.2 The Influence of Fracture Scale on Surface Roughness

Previous studies investigating the effect of fracture length on crack propagation have suggested that different fracture formation mechanisms will influence roughness parameters on scales significantly smaller than the total fracture length (Mosher et al. 1975; Mardon et al. 1990). In this study, fracture length during crack propagation is known for the artificial fractures (as it is restricted to the specimen size) but not for the natural fractures sampled in the field. As the scale of the natural fractures is unknown, the dependence of surface roughness on the fracture size is investigated here by considering the artificial fractures alone.

The dependency of the standard deviation on specimen size is removed by subsampling the standard deviation of the surface asperity heights within regions of fixed size at 20 randomly chosen locations on the surfaces. In Fig. 6a, the sample region size is increased incrementally from 1 cm^2 to the total size of the surface, and the average standard deviation of all surface points is compared to the edge length of the subsampling region. The resolution of the respective subsampled regions was thereby kept constant, which was shown to be crucial for roughness comparison by Tatone and Grasselli (Tatone and Grasselli 2013). When sampling from a small scale to the whole fracture area, the apparent roughness approaches a plateau once the box edge length approaches the global specimen dimensions (e.g., the whole fracture area). To avoid this effect, the sampling window side length is restricted to less than 5 cm in Fig. 6a, and in Fig. 6b the sample regions considered are restricted to 1 cm × 1 cm to 2 cm × 2 cm to study roughness effects on a scale found across all specimens. Comparing sampling windows of the same size, the surface roughness as measured by the asperity height standard deviation increases faster with increasing specimen size (Fig. 6b). Even on a small but fixed sampling window (Fig. 6b, c) of 1 cm^2, the averaged standard deviations show a dependence on the total specimen size. The averaged standard deviation over the total fracture size for a box edge length of 1 cm can be seen in Fig. 6c. For small specimens, the standard deviation is significantly smaller than for larger specimens, even when sampling a uniform box of 1 cm^2. However, the increase in standard deviation occurs sublinear to the total fracture area size. Here, this difference may reflect the effect of the physical specimen size on the surface roughness as the specimen dimension and diameter are also representable of the total crack length during propagation.

These observations also suggest that tensile fracturing in a Brazilian test of small specimen is likely dominated by intragranular cracks. The change in the dominant crack propagation mechanisms across specimen sizes raise an interesting question when considering fractal measures. If fracture surfaces are self-affine across multiple length scales, this would indicate that different fracture mechanisms lead to self-affine cracks, independent of fracturing. However, intragranular cracks lead to surfaces that are smoother on a large scale as compared to larger specimen since cracks propagate in a straighter line and do not circumvent individual grains. But, on a small scale they are jagged, since cracks do not follow grain boundaries—instead splitting individual grains through their plane of weakness. For small specimens, single grains in the tensile stress zone of a Brazilian test specimen are split if they lay on the straight line of largest tensile stresses in the middle of the specimen. The short wave oscillations resulting from intragranular crack propagation more prominently found in artificial tensile fractures can be observed for instance in Fig. 2c.

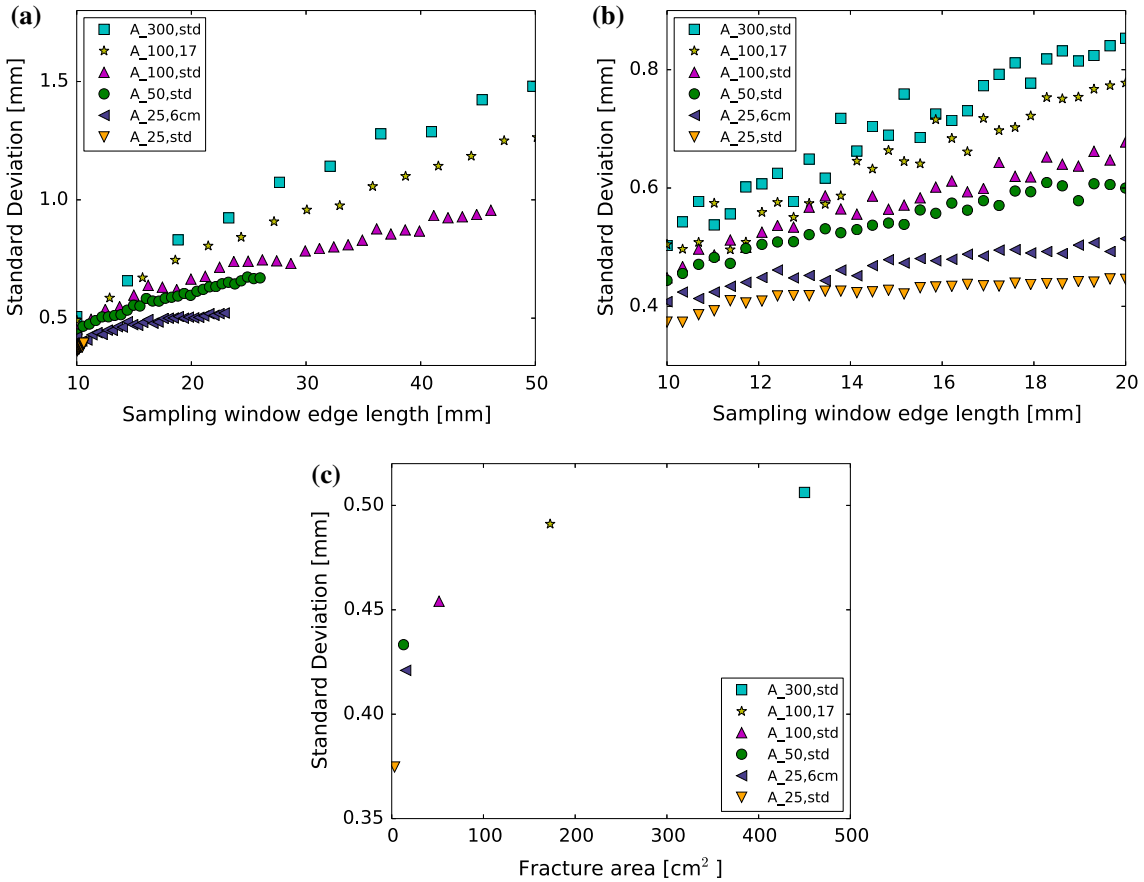


Fig. 6 Specimen size effects: **a** standard deviation of asperity height versus the edge length of the sampling rectangle for artificial fractures from Brazilian tests with diameters of 25.0 cm ($A_{25.0,\mathrm{std}}$, according to standard), 2.5 cm ($A_{2.5}$), 5.0 cm ($A_{5.0,\mathrm{std}}$, according to standard), 10.0 cm (A_{10}), 10.0 cm ($A_{10,\mathrm{std}}$, according to standard) and 30.0 cm ($A_{30.0,\mathrm{std}}$, according to standard). **b** Mean of the standard deviation for small box edge lengths between 10 mm × 10 mm and 20 mm × 20 mm with a square root fit. **c** The averaged standard deviation versus total fracture area when sampling 10 mm box edge lengths

While Fig. 6a, b showcase the scale dependence of the sampling interval, these figures also reveal that the total fracture length of a specimen influences the surface roughness on all scales.

Comparison of the artificial tests shows smoother surfaces for smaller specimens, which could be explained by a higher ratio of intragranular cracks under higher tensile stresses (Fig. 10). Roughness as measured by the STD increases faster for larger specimens when the same sampling window is compared (Fig. 6b). This suggests that measuring the surface roughness of a fixed sampling window may provide an indication of the stress at failure and total crack length. The decrease in roughness with smaller artificial specimen could also indicate that surface roughness on natural fractures might be tied to the domain size which experienced failure or rupture. At the very least, by fixing the sampling window, the size-dependent behavior observed by taking the standard deviation across the whole surface of each specimen can be avoided.

3.3 Size-Independent Measures of Fracture Roughness

The scale dependence of the standard deviation on the whole fracture surface is depicted in Fig. 7. This figure repeats the earlier plot of the standard deviation for the whole surface versus the measured Z2 value; however, in this case the points are scaled to reflect the size of the sampled area. The Z2 values for natural shear fractures do not demonstrate a size dependency due to the fixed sampling interval. However, the standard deviation clearly increases with fracture area, with most small-scale fracture areas displaying a standard deviation between 0.4 and 0.7 mm while fractures with larger surface areas show a standard deviation above 0.8 mm.

To eliminate the effects of size dependence, each fracture surface was divided into subregions of a fixed size and those sample regions analyzed independently. Based on the analysis given in the previous section, subregions 1 cm^2 in size

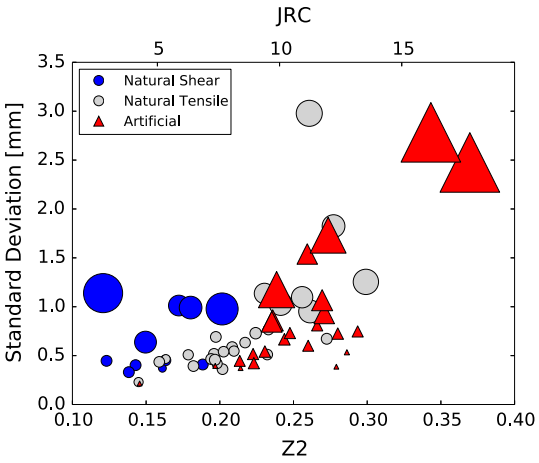


Fig. 7 Sample size effects: standard deviation of surface height versus Z2 measure for both the natural shear (*blue circles*), natural mode I (*gray circles*) and artificial tensile (*red triangles*) fracture surfaces, with points scaled proportional to the specimen size (color figure online)

(Fig. 8) were chosen, as this allowed a reasonable number of subregions to be obtained from each surface, while maintaining a large number of sample points within each subregion. Prior to calculating the roughness measures, each subregion was reoriented to align the surface normal parallel to the z-axis and the surface shifted to align the origin with the mean asperity height. Figure 8 shows sampling windows of 1 cm^2 on the fracture surfaces, sorted by the different fracture classes (natural tensile and shear and artificial tensile fractures). These images reveal the characteristic roughness patterns for the different fracture classes on a small scale, as marked in the respective graphs. Natural shear fractures display the shear orientation of the fracture and slickensides (Fig. 8a), while natural tensile fractures tend to show grain shaped surface undulations (Fig. 8c). The small-scale oscillations commonly observed on artificial tensile fracture surfaces (Fig. 2c) are visible on small scales (Fig. 8e), with jagged behavior of surface roughness indicating crack propagation through individual grains. In all cases, however, the feature size correlates strongly with the average grain scale.

Figure 9 shows the effect of subdividing the surface into 1 cm^2 areas on the Z2 and standard deviation measures. The subdivision of the surfaces into 1 cm square regions mitigates the scale dependency of surface measures and avoids skewing of results due to large-scale fluctuations—effectively performing a high-pass filter on the surface properties. The results from all 1 cm^2 subregions on each surface in Fig. 9 are averaged over each surface. This plot reveals distinct clustering for the three fracture classes, with natural shear fractures showing small Z2 values and standard deviations, natural tensile fractures having a larger standard deviation, and artificial tensile fractures displaying small variation in standard deviation as well as in Z2.

Figure 9 also reveals a strong linear trend between the average standard deviation and the average Z2 value for the surface classes. Notably, however, the two trends, while roughly parallel, are slightly offset from one another—with the natural specimen exhibiting a larger standard deviation for the same Z2 value. The offset between the two clusters in Fig. 9 can be attributed to differences in the autocorrelation of the natural and artificial surfaces at the given 0.25 mm sampling interval. Assuming an equal sampling interval of $\Delta x = 0.25$ and a sufficiently large number of sample points ($n \gg 1$), Eq. (2) can be expressed as

$$\begin{aligned} Z2 &\simeq \left[\frac{1}{n\Delta x^2}\left(\sum_i y_{i+1} - y_i\right)^2\right]^{\frac{1}{2}} \\ &\simeq \left[\frac{1}{n\Delta x^2}\sum_i y_i^2\right]^{\frac{1}{2}}\left[1 - \sum_i y_i y_{i+1} \Big/ \sum_i y_i^2\right]^{\frac{1}{2}}. \end{aligned} \tag{6}$$

As the mean asperity height is zero, the term on the left approximately equals the standard deviation scaled by the sampling interval. The term "approximately" is used as the standard deviation is obtained directly from the surface rather than from the 0.25 mm sampling interval employed for the Z2 approximation. The term on the right can be expressed in terms of the autocorrelation for the surface at the sampling interval:

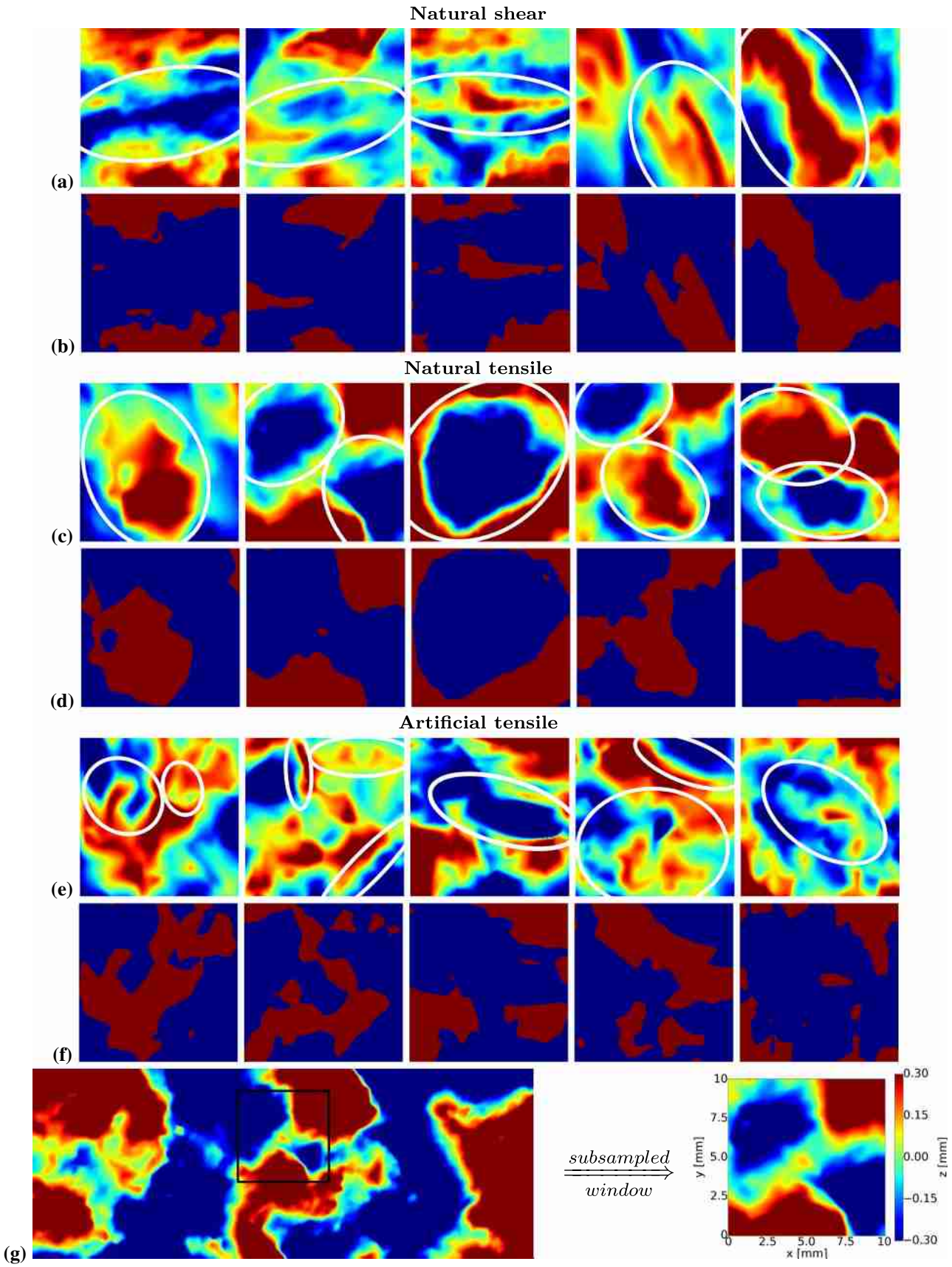
$$\mathrm{ACF}(\tau) = \int y(\mathbf{x}) y(\mathbf{x}+\tau)\,\mathrm{d}\mathbf{x} \Big/ \int y^2\,\mathrm{d}\mathbf{x}, \tag{7}$$

when $\tau = \Delta x$. Thus, the standard deviation and Z2 values are related by:

$$Z2 \simeq \frac{\mathrm{STD}}{\Delta x}[1 - \mathrm{ACF}(\Delta x)]^{\frac{1}{2}}. \tag{8}$$

With Eq. (8), the trends observed in Fig. 9 can be explained as follows: First, both classes show a strongly linear relationship between the standard deviation and the Z2 measure, implying that there is relatively little change in the autocorrelation function of the surfaces for the sample distance of 0.25 mm. However, there is a slight difference in the slopes of the standard deviation versus Z2 plots for the two sets (natural vs. artificial) of surfaces. The artificial surfaces have a shallower slope than the natural surfaces when plotting the standard deviation versus Z2. This second observation implies that the artificial surfaces are less correlated (i.e., are rougher) than the natural surfaces at the sample distance, which can be attributed to crack propagating either through or around individual grains, without circumventing larger conglomerates.

This finding is in line with studies by Kranz (1983) and Mardon et al. (1990), which pointed out the dependency of


Natural shear
(a)
(b)
Natural tensile
(c)
(d)
Artificial tensile
(e)
(f)
(g)
subsampled
window
10
7.5
5.0
2.5
0
y [mm]
0
2.5
5.0
7.5
10
x [mm]
0.30
0.15
0.00
−0.15
−0.30
z [mm]

◀**Fig. 8** Examples of subsampled windows of 1 cm^2 on the fracture surfaces, sorted by fracture mode, with marking of characteristic regions. **a**, **c**, **e** *color bar* goes from −0.3 (*blue*) to 0.3 (*red*) mm. Note that significant portions are outside this range; **b**, **d**, **f** surface height above (*red*) and below (*blue*) 0.1 mm; **g** example surface patch with dimensions and colorbar representative of **a**–**f** (*right*) sampled from whole fracture surface (*left*) and reoriented to the *x*–*y* plane (color figure online)

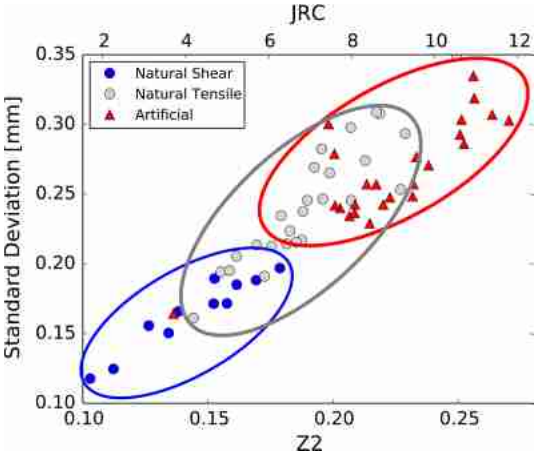


Fig. 9 Plots of standard deviation of asperity height versus Z2 measure for 1 cm^2 subregions of the fracture surface averaged for each surface

fracture length and most common fracture propagation mechanism on total stress. This would suggest longer wavelengths for natural fractures. Additional factors contributing to smoother natural shear fractures are shearing off of first order asperities and ongoing mineralization. To extend on the topic of size dependency, the surface roughness of experimental specimens of constant size, commonly used in core holder experiments, is addressed in "Appendix".

3.4 Dependence of Surface Roughness on Tensile Stress for Brazilian Tests

As the artificial fracture surfaces were obtained through Brazilian testing, a comparison can be made between the tensile stresses during failure and the resulting fracture roughness. Measurement of fracture roughness as indicated by STD on 1 cm^2 patches suggests increasing roughness for larger Brazilian test specimen (Fig. 6c). To investigate this further, the impact of tensile stress at failure on surface roughness is investigated. Figure 10a shows the standard deviation averaged over both fracture surfaces per specimen and 30 squares with 10 mm edge length per fracture surface and for the whole fracture surface (Fig. 10b).

Due to the small sampling intervals (1 cm^2), the differences in standard deviation are not significant, but a trend of decline in surface roughness as measured with the standard deviation with increasing tensile stress during failure can be observed. Higher tensile stresses favor intragranular crack propagation, which will result in a smoother fracture surface since crack propagation occurs through individual grains without having to circumvent minerals of high strength as is the case for low tensile stresses. The trend of decreasing surface roughness with rising tensile stress becomes more pronounced when comparing the averaged standard deviation of the two specimen fracture surfaces for A_{300} in Fig. 10b. Here, the specimens have a size that is significantly larger than inhomogeneities within the material, which makes them less susceptible to heterogeneity influences on the surface roughness. The two tests show most clearly a reduced roughness as measured with the asperity standard deviation with increased tensile stress. Generally, with increasing specimen size the roughness of a tensile fracture that develops under Brazilian testing conditions increases. This

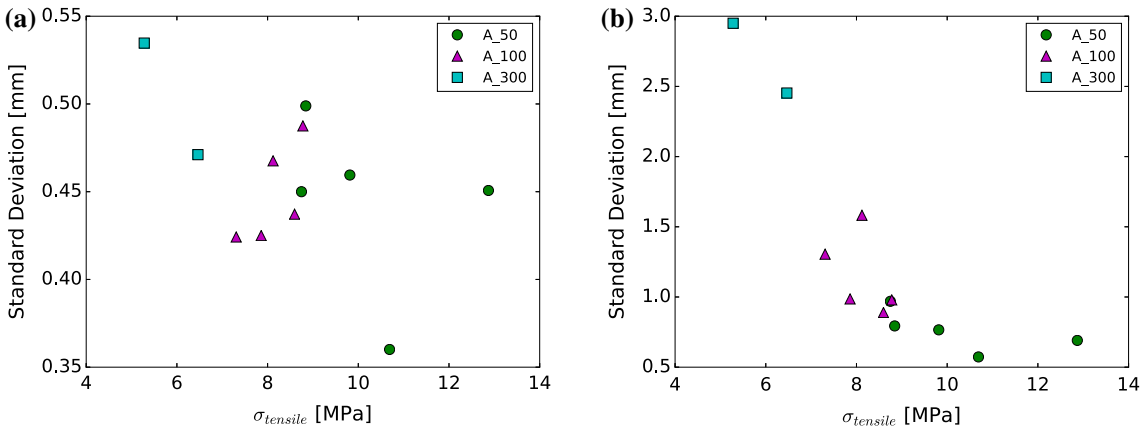


Fig. 10 Relationship of surface roughness to sigma tensile: **a** average standard deviation of asperity height within 1 cm squares versus the tensile stress at specimen failure. **b** Average standard deviation of asperity height for both fracture surfaces versus the tensile stress at specimen failure

increase in roughness is likely associated with the fracture propagation path. Fracture under Brazilian testing conditions emanate at the specimen center where the tensile stress is maximum (Fairhurst 1964). This initial fracture propagates radially toward the upper and lower loading plate where the stress conditions are compressive. The fracture propagation path in a Brazilian test is controlled by both, the stress state associated with a line load applied on both specimen sides (i.e., induced tensile stresses occur within a narrow area in the specimen center), and stress heterogeneities stemming from stiffness heterogeneities (Tapponnier and Brace 1976) that are associated with the

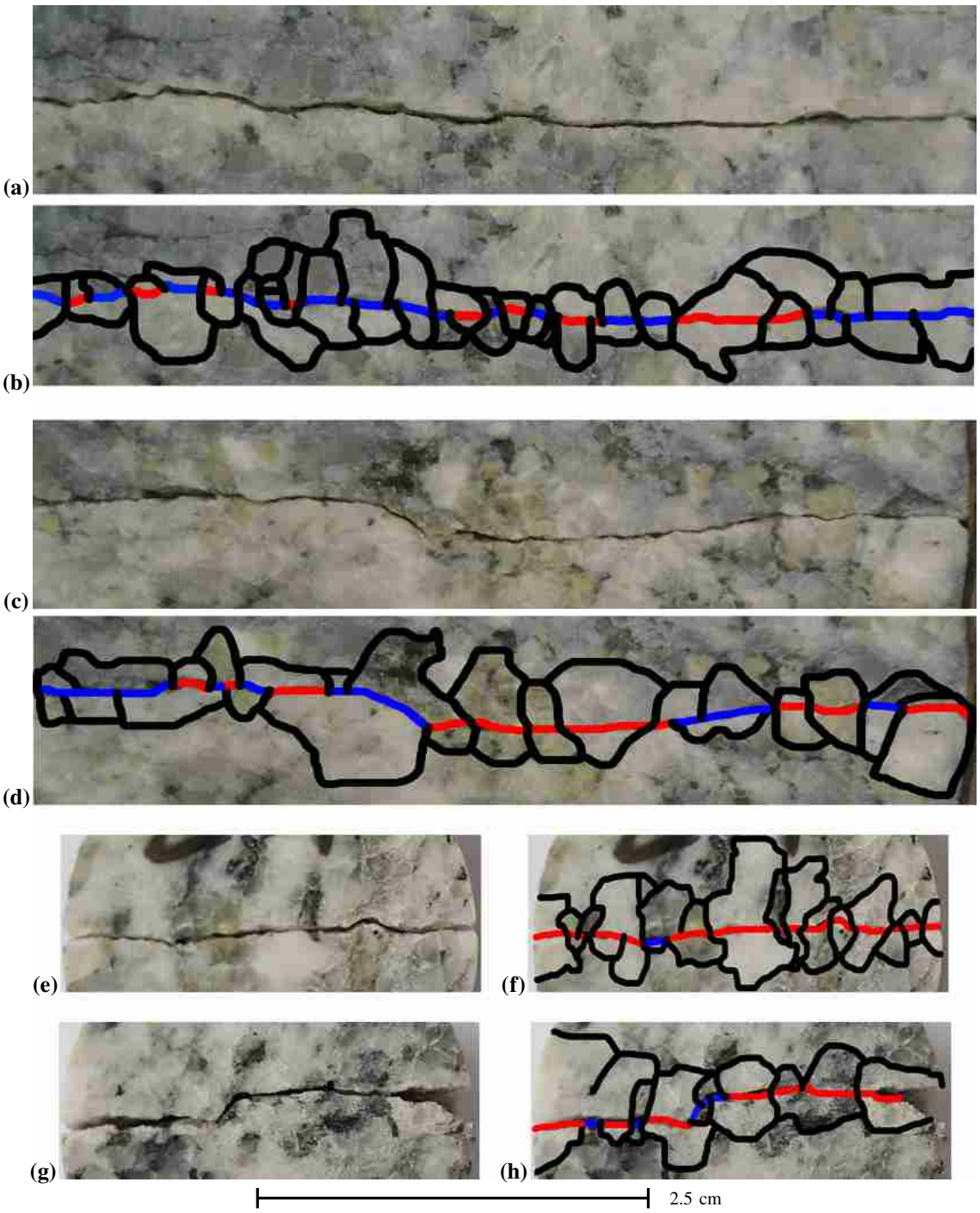


Fig. 11 Examples of crack propagation between two grains (*blue*) and through an individual grain (*red*) for: **a–d** artificial tensile fracture from specimens with 10 cm diameter; **e–h** artificial tensile fracture from specimens with 2.5 cm diameter. Individual grains near the crack are outlined in *black* (color figure online)

distribution of mineral grains of dissimilar stiffness (Tapponnier and Brace 1976). For smaller specimens, the tension area for potential fracture pathways is significantly smaller as compared to larger specimens. In case of smaller specimens, cracks are therefore forced to propagate through mineral grains. Increasing crack propagation around versus through individual grains with increasing specimen size can be observed in Fig. 11, where representative examples of crack propagation through specimen with 10.5 cm (Fig. 11a–d) and 2.5 cm diameter (Fig. 11e–h) are shown, with marked crack paths through (red) and around (blue) individual grains. If specimens are larger, a propagating fracture may bypass strong mineral grains and may also encounter a larger number of flaws within the vicinity of the ideal crack propagation path through the specimen (e.g., the straight line between top and bottom of the specimen). With increasing crack growth around grains, the asperity heights are more irregular and sizeable. This observation is consistent with Figs. 7 and 10, where larger roughness values were computed for artificial fractures from larger specimens. Since the total length of the cracks in the natural fractures is unknown, the natural fractures are not suitable for comparison here. This interpretation is strongly supported by the considerable decrease in measured tensile strength from 11 MPa (on average) for 54 mm specimens to 6 MPa for 300 mm diameter specimens. It should be noted that while this behavior was found to be consistent for the specimens used in this study, conclusions for fractures of larger size remain hypothetical.

3.5 Two-Point Correlation Functions

The use of two-point correlation functions is somewhat scale dependent, as once again small-scale roughness is convoluted with the curvature of the fracture on larger-scales. As was the case with the calculation of the standard deviation in Sect. 3.3, it is difficult to separate the two without fixing a length scale. Thus, we proceed as before and apply the two-point correlation functions to the same 1 cm^2 subregions used to measure the standard deviation. The results of these calculations are summarized in Fig. 12. The figure compares the median value of the two-point probability and lineal-path functions (along with the corresponding interquartile range) for the three distinct fracture surface classes. Figure 12 reveals that there is minor difference in the two-point probability function and lineal-path function for the surfaces. However, the lineal-path function is slightly less well correlated.

Natural tensile fractures show the highest correlation, followed by natural shear fractures, with artificial tensile fractures showing the fastest decline in correlation with l. This is likely related to the surface fabric. As shown in Fig. 8b, crack propagation in natural tensile fractures tends to grow around whole grains, which could lead to longer wave lengths of surface roughness on the small sample scales investigated by the correlation functions. Natural shear fractures are often well correlated in the direction of slickensides, but not necessarily in other directions (Fig. 8a), which explains the lower correlation in comparison to natural tensile fractures. Additionally, natural shear fractures show lower absolute asperity height magnitudes, favoring short wave length oscillations of small-scale roughness. Artificial tensile fractures are least well correlated, which can be attributed to the short wave length oscillations commonly observed on shear fracture surfaces (Fig. 8c), which favor intragranular crack growth. These findings suggest that the surface correlation is dominated by the fabric of the granodiorite crystals and the dominant crack growth mechanism.

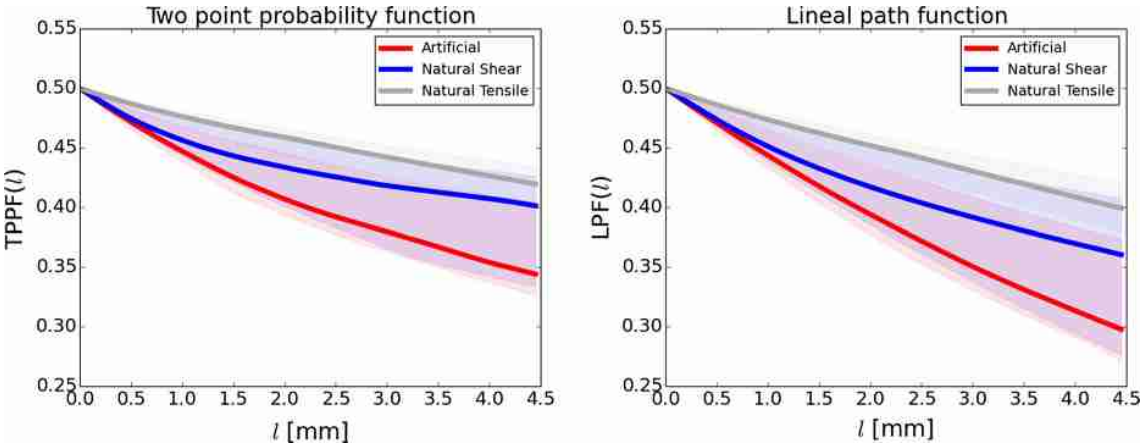


Fig. 12 Median values (*solid lines*) and interquartile ranges (transparent regions) of two-point correlation functions (two-point probability function *TPPF*; lineal-path function *LPF*) calculated on 1 cm × 1 cm regions. *Colors* indicate the three surface types—natural shear (*blue*), natural tensile (*gray*) and artificial (*red*) (color figure online)

While common surface roughness measurements mainly observe direction-dependent roughness along scan lines, the two-point correlation functions also deliver additional information about the spatial variation of these surfaces. Namely, that the differences in the roughness metrics across the different fracture types stem from differences in the extents of the asperities above and below the mean, rather than the frequency of changes in asperity height. These findings again reflect the importance of grain-scale features for surface roughness and strength properties on the laboratory scale. Namely, the grain-scale size dominates crack propagation and introduces characteristic shapes. It also helps to explain the divergent results observed by the fractal measures, as the grain scale introduces an intrinsic size that breaks the assumption of self-similarity across length scales.

While common methods such as the JRC or the Z2 can yield large roughnesses, if the global changes in surface asperity height are large (Eq. 2), correlation functions provide unique insights on roughness characterized by short wave lengths. The correlation function results showcase roughness induced by fundamental changes in crack growth, such as differences induced by changes between intra- and intergranular crack growth.

4 Conclusion

This study revealed a clear distinction between the roughness of natural fractures and those created artificially via the Brazilian tests, for crystalline rock specimen from the Grimsel Test Site, Switzerland. The Z2 value distinguishes well between different fracture types, with natural shear fractures showing the lowest roughness, artificial tensile fractures the highest roughness and natural tensile fractures spanning almost the whole range of values. Natural tensile and shear fractures both demonstrated stronger longer-wavelength fluctuations. In part this may be explained by the fact that, unlike their natural counterparts, the artificial fractures were not exposed to weathering or shearing, which remove small-scale asperities. However, in addition, fracture propagation is less constrained in a geological setting, than under laboratory conditions. Natural fractures are able to sample a larger set of potential fracture paths allowing it to exploit pre-existing planes of weakness within the rock, whereas the stress field generated by the Brazilian test constrains the artificial fractures to bisect a small region toward the center of the specimen.

This is supported by the scale dependency of roughness demonstrated in this study. For artificial fractures, where the total specimen dimensions during failure are determined, smaller specimens show less roughness. Observations of small-scale roughness indicate that this is linked to a higher ratio of intergranular cracks, due to increasing tensile strength with smaller specimen sizes. This increased ratio of intergranular cracks and short wave length oscillations of roughness also lead to less spatially correlated surfaces on artificial fractures. This serves to illustrate that roughness measurements are largely meaningless unless pinned to a particular scale, whether that be a sampling interval (as in the case of the Z2/JRC metrics) or restricted to a sample region (as in the case of the standard deviation of surface asperities). To address this problem, the presented study provides strong evidence that fracture type and size (e.g., natural or artificial tensile or shear fractures) can both be distinguished by computing surface roughness on subsampled regions of fixed size on the fracture surfaces.

This emphasizes the influence of the scale at which a fracture is created. While this illustrates the multi-scale nature of these surface properties, the fractal measures of roughness used in this study do not yield any conclusive distinction between the natural and artificial specimens, suggesting that these metrics are inappropriate for surface characterization on specimens of this size. Indeed, the sensitivity of the surfaces to different processes on different scales would suggest that the surfaces are not fractal (i.e., self similar), but vary in their properties across scales. This is especially important as fractal methods are commonly employed to generate synthetic fracture surfaces for numerical studies.

This study presents evidence that results from hydro-mechanical laboratory tests, which are strongly linked to fracture topography, need to be clearly distinguished with surface roughness measures, both regarding specimen size and nature. Furthermore, varying results in computed roughness—depending on the employed method and studied scale—suggest that novel methods in surface roughness need to be compared with established methods across a wide range.

Acknowledgements The authors thank the Federal Rock Laboratory Grimsel, Switzerland and the CRIEPI fractured rock study (Takana et al. 2014) for providing us with the core material for the rock specimen of this study. The authors further want to thank the Geosensors and Engineering Geodesy group at ETH Zürich and especially Robert Presl for their support with the photogrammetry scanner used for this study. This work was partially supported by the GEOTHERM II project, which is funded by the Competence Center Environment and Sustainability of the ETH Domain. Finally the authors would like to thank the anonymous reviewers for their valuable feedback and input to improve this manuscript.

Appendix

Supplementary Material: Comparison of Specimen Roughness for Experiment Purposes

Another way to investigate scale-independent roughness effects is the comparison of specimen sizes commonly used

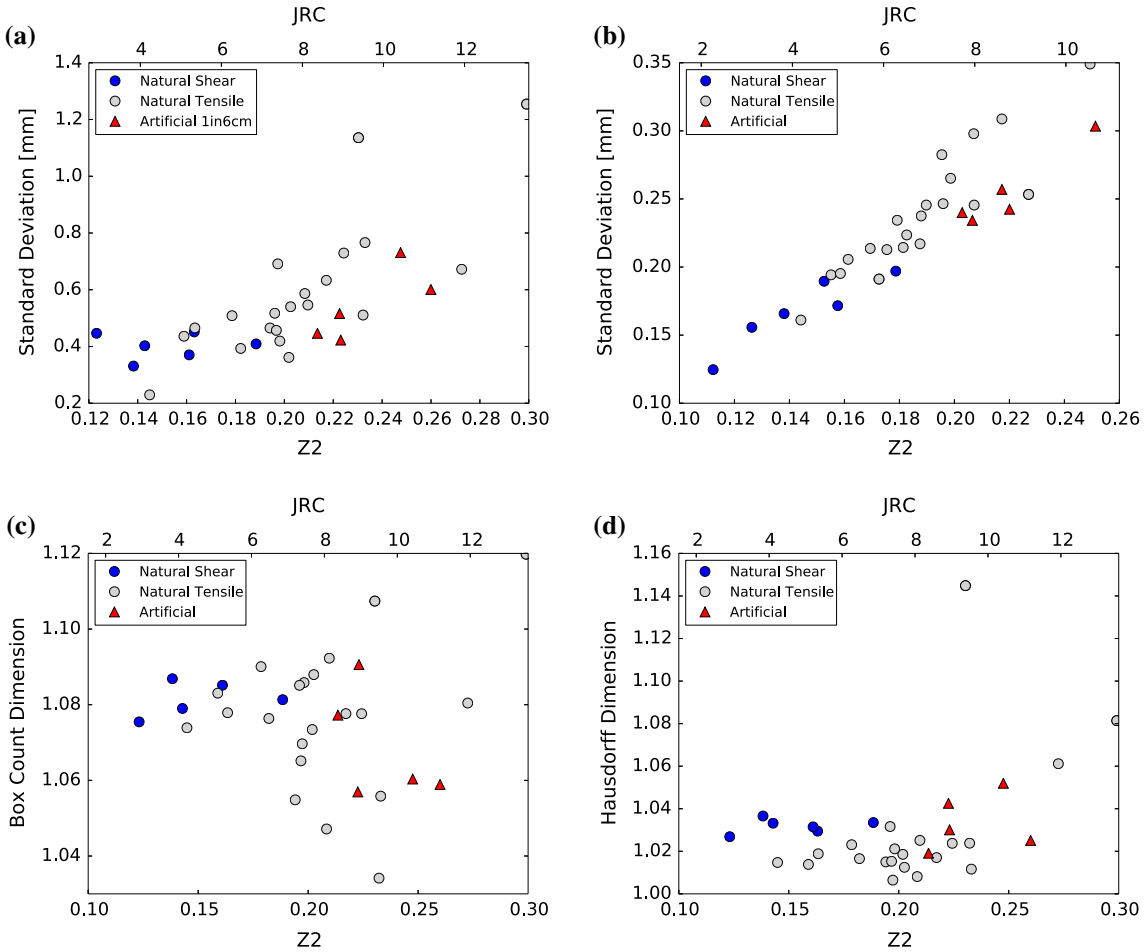


Fig. 13 Surface roughness measures for cylindric specimen sizes of 2.5 cm diameter and 6.0 cm length. Differentiated are natural shear (*blue*), natural tensile (*gray*) and artificial tensile fractures (*red*). **a** Standard deviation versus Z2 value for the entire fracture surface; **b** standard deviation versus Z2 value for 1 cm^2 subregions of the fracture surface (averaged values); **c** Box count dimension versus Z2 value; **d** Hausdorff dimension versus Z2 value (color figure online)

in core holders. Core holders apply a radial and axial pressure on cylindrical specimens, while establishing fluid flow through rock specimens to measure permeability. Not only do specimens have the same size, specimens for core holders have a size that allows investigation whether surface roughness characteristics can be distinguished for specimen used in laboratory tests.

The fracture surfaces investigated in this study vary greatly in size and origin (Tables 1, 2). A key motivation for this study is the impact of specimen origin when conducting experimental research. The natural specimen sizes of cylindrical shape with 2.5 cm diameter and 6 cm length were overcored for an experimental study investigating fracture surface deformation and permeability evolution during cyclic loading. Artificial specimens with the same dimensions were produced to compare natural and artificial fracture surfaces of specimen sizes representative for core holder experiments (Vogler et al. 2016a). The Z2 value distinguishes between natural shear and artificial tensile fractures for the whole specimen area and for 1 cm^2 subregions of the fracture surface (Fig. 13a, b). Natural tensile fractures span a wide range of Z2 and standard deviation values (Fig. 13a) that show values found for natural shear as well as artificial tensile fractures. The Box count and Hausdorff dimensions do not yield conclusive results to distinguish between specimens (Fig. 13c, d). However, while the Box count dimension is not consistent for individual fracture classes, the Hausdorff dimension for different fracture classes is more consistent with most values between 1.01 and 1.04, thereby showing a similar fractal dimension for fracture surfaces regardless of the respective class.

References

Babadagli T, Develi K (2000) Fractal analysis of natural and synthetic fracture surfaces of geothermal reservoir rocks. In: Proceedings

of the World Geothermal Congress, vol 28, Kyushu-Tohoku, Japan
Barton N, Choubey V (1977) The shear strength of rock joints in theory and practice. Rock Mech Rock Eng 10:1–54
Bear J, Tsang C, De Marsily G (2012) Flow and contaminant transport in fractured rock. Elsevier, New York
Belem T, Homand-Etienne F, Souley M (2000) Quantitative parameters for rock joint surface roughness. Rock Mech Rock Eng 33(4):217–242. doi:10.1007/s006030070001
Brown SR (1987) Fluid-flow through rock joints—the effect of surface roughness. J Geophys Res Solid Earth Planets 92(B2):1337–1347
Chen Z, Narayan S, Yang Z, Rahman S (2000) An experimental investigation of hydraulic behaviour of fractures and joints in granitic rock. Int J Rock Mech Min Sci 37(7):1061–1071
Derode B, Cappa F, Guglielmi Y, Rutqvist J (2013) Coupled seismo-hydromechanical monitoring of inelastic effects on injection-induced fracture permeability. Int J Rock Mech Min Sci 61:266–274. doi:10.1016/j.ijrmms.2013.03.008
Esaki T, Du S, Mitani Y, Ikusada K, Jing L (1999) Development of a shear-flow test apparatus and determination of coupled properties for a single rock joint. Int J Rock Mech Min Sci 36(5):641–650
Fairhurst C (1964) On the validity of the 'Brazilian' test for brittle materials. Int J Rock Mech Min Sci Geomech Abstr 1(4):535–546. doi:10.1016/0148-9062(64)90060-9
Faoro I, Elsworth D, Marone C (2012) Permeability evolution during dynamic stressing of dual permeability media. J Geophys Res Solid Earth. doi:10.1029/2011JB008635,b01310
Fardin N, Stephansson O, Jing LR (2001) The scale dependence of rock joint surface roughness. Int J Rock Mech Min Sci 38(5):659–669
Fujii Y, Takemura T, Takahashi M, Lin W (2007) Surface features of uniaxial tensile fractures and their relation to rock anisotropy in inada granite. Int J Rock Mech Min Sci 44(1):98–107
GOM mbH (2015) GOM: Atos core. http://www.atos-core.com/en/industries.php
Hatheway AW (2009) The complete isrm suggested methods for rock characterization, testing and monitoring; 1974–2006. Environ Eng Geosci 15(1):47–48
Huang SL, Oelfke SM, Speck RC (1992) Applicability of fractal characterization and modeling to rock joinnt profiles. Int J Rock Mech Min Sci Geomech Abstr 29(2):89–98. doi:10.1016/0148-9062(92)92120-2
Jiang Y, Li B, Tanabashi Y (2006) Estimating the relation between surface roughness and mechanical properties of rock joints. Int J Rock Mech Min Sci 43(6):837–846
Jiao Y, Stillinger F, Torquato S (2007) Modeling heterogeneous materials via two-point correlation functions: basic principles. Phys Rev E 76(3):031,110
Kranz RL (1983) Microcracks in rocks—a review. Tectonophysics 100(1–3):449–480
Lee HS, Cho TF (2002) Hydraulic characteristics of rough fractures in linear flow under normal and shear load. Rock Mech Rock Eng 35(4):299–318. doi:10.1007/s00603-002-0028-y
Lee YH, Carr J, Barr D, Haas C (1990) The fractal dimension as a measure of the roughness of rock discontinuity profiles. Int J Rock Mech Min Sci Geomech Abstr 27(6):453–464. doi:10.1016/0148-9062(90)90998-H
Lee YK, Park JW, Song JJ (2014) Model for the shear behavior of rock joints under CNL and CNS conditions. Int J Rock Mech Min Sci 70:252–263. doi:10.1016/j.ijrmms.2014.05.005
Li Y, Huang R (2015) Relationship between joint roughness coefficient and fractal dimension of rock fracture surfaces. Int J Rock Mech Min Sci 75:15–22
Li Y, Chen Y, Zhou C (2014) Hydraulic properties of partially saturated rock fractures subjected to mechanical loading. Eng Geol 179:24–31
Mann HB, Whitney DR (1947) On a test of whether one of two random variables is stochastically larger than the other. Ann Math Stat 18(1):50–60
Mardon D, Kronenberg AK, Handin J, Friedman M, Russell JE (1990) Mechanism of fracture propagation in experimentally extended sioux granite. Tectonophysics 182(3–4):259–278
Morgan SP, Johnson CA, Einstein HH (2013) Cracking processes in barre granite: fracture process zones and crack coalescence. Int J Fract 180(2):177–204. doi:10.1007/s10704-013-9810-y
Mosher S, Berger R, Anderson D (1975) Fracturing characteristics of two granites. Rock Mech 7(3):167–176
Myers N (1962) Characterization of surface roughness. Wear 5(3):182–189
Nemoto K, Watanabe N, Hirano N, Tsuchiya N (2009) Direct measurement of contact area and stress dependence of anisotropic flow through rock fracture with heterogeneous aperture distribution. Earth Planet Sci Lett 281(1–2):81–87. doi:10.1016/j.epsl.2009.02.005
Nicholl MJ, Rajaram H, Glass RJ, Detwiler R (1999) Saturated flow in a single fracture: evaluation of the reynolds equation in measured aperture fields. Water Resour Res 35(11):3361–3373
Ogilvie SR, Isakov E, Taylor CW, Glover PWJ (2003) Characterization of rough-walled fractures in crystalline rocks. Hydrocarb Cryst Rocks 214:125–141
Power WL, Durham WB (1997) Topography of natural and artificial fractures in granitic rocks: implications for studies of rock friction and fluid migration. Int J Rock Mech Min Sci 34(6):979–989
Pyraknolte LJ, Myer LR, Nolte DD (1992) Fractures—finite-size-scaling and multifractals. Pure Appl Geophys 138(4):679–706. doi:10.1007/bf00876344
Rutqvist J, Stephansson O (2003) The role of hydromechanical coupling in fractured rock engineering. Hydrol J 11(1):7–40
Takana Y, Miyakawa K, Fukahori D, Kiho K, Goto K (2014) Survey of flow channels in rock mass fractures by resin injection. In: Asian rock mechanics symposium 8
Tapponnier P, Brace W (1976) Development of stress-induced microcracks in westerly granite. Int J Rock Mech Min Sci Geomech Abstr 13(4):103–112
Tatone BSA, Grasselli G (2010) A new 2D discontinuity roughness parameter and its correlation with JRC. Int J Rock Mech Min Sci 47:1391–1400. doi:10.1016/j.ijrmms.2010.06.006
Tatone BSA, Grasselli G (2013) An investigation of discontinuity roughness scale dependency using high-resolution surface measurements. Rock Mech Rock Eng 46(4):657–681. doi:10.1007/s00603-012-0294-2
Tse R, Cruden D (1979) Estimating joint roughness coefficients. Int J Rock Mech Min Sci Geomech Abstr 16:303–307
Vogler D (2016) Hydro-mechanically coupled processes in heterogeneous fractures: experiments and numerical simulations. PhD thesis, ETH Zurich, Zurich, Switzerland
Vogler D, Amann F, Bayer P, Elsworth D (2016a) Permeability evolution in natural fractures subject to cyclic loading and gouge formation. Rock Mech Rock Eng 49(9):3463–3479. doi:10.1007/s00603-016-1022-0
Vogler D, Settgast RR, Annavarapu C, Bayer P, Amann F (2016b) Hydro-mechanically coupled flow through heterogeneous fractures. In: Proceedings, 41st workshop on geothermal reservoir engineering, pp SGP-TR-209
Vogler D, Walsh SDC, Dombrovski E, Perras MA (2017) A comparison of tensile failure in 3D-printed and natural sandstone. Eng Geol. doi:10.1016/j.enggeo.2017.06.011

Watanabe N, Hirano N, Tsuchiya N (2008) Determination of aperture structure and fluid flow in a rock fracture by high-resolution numerical modeling on the basis of a flow-through experiment under confining pressure. Water Resour Res. doi:10.1029/2006WR005411

Watanabe N, Hirano N, Tsuchiya N (2009) Diversity of channeling flow in heterogeneous aperture distribution inferred from integrated experimental-numerical analysis on flow through shear fracture in granite. J Geophys Res Solid Earth 114(B4):2156–2202. doi:10.1029/2008JB005959

Witherspoon PA, Wang JSY, Iwai K, Gale JE (1980) Validity of cubic law for fluid-flow in a deformable rock fracture. Water Resour Res 16(6):1016–1024

Xiong X, Li B, Y J, Koyama T, Zhang C (2011) Experimental and numerical study of the geometrical and hydraulic characteristics of a single rock fracture during shear. Int J Rock Mech Min Sci 48(8):1292–1302

Yeo IW, De Freitas MH, Zimmerman RW (1998) Effect of shear displacement on the aperture and permeability of a rock fracture. Int J Rock Mech Min Sci 35(8):1051–1070

Yu X, Vayssade B (1991) Joint profiles and their roughness parameters. Int J Rock Mech Min Sci Geomech Abstr 28:333–336

Zhu WC, Liu J, Elsworth D, Polak A, Grader A, Sheng JC, Liu JX (2007) Tracer transport in a fractured chalk: X-ray ct characterization and digital-image-based (dib) simulation. Transp Porous Media 70(1):25–42