Effective stress, friction and deep crustal faulting

December 3, 15

N. M. Beeler, US Geological Survey, Cascades Observatory, Vancouver, Washington
Greg Hirth, Brown University, Providence, Rhode Island
Amanda Thomas, University of Oregon, Eugene, Oregon
Roland Bürgmann, University of California, Berkeley, California

key points:

The real area of contact determines the effective pressure coefficient in the deep crust

The effective stress coefficient transitions to near zero at the brittle ductile transition (BDT) for wide shear zones

Below the BDT reactivating friction may require localization in addition to elevated pore pressure

agu index terms: 8004, 8034, 8163

Abstract. Studies of crustal faulting and rock friction invariably assume the effective normal stress that determines fault shear resistance during frictional sliding is the applied normal stress minus the pore pressure. Here we propose an expression for the effective stress coefficient α_f at temperatures and stresses near the brittle ductile transition (BDT) that depends on the percentage of solid-solid contact area across the fault. α_f varies with depth and is only near 1 when the yield strength of asperity contacts greatly exceeds the applied normal stress. For a vertical strike-slip quartz fault zone at hydrostatic pore pressure and assuming 1 mm and 1 km shear zone widths for friction and ductile shear, respectively, the BDT is at ~13 km. α_f near 1 is restricted to depths where the shear zone is narrow. Below the BDT $\alpha_f = 0$ due to a dramatically decreased strain rate. Under these circumstances friction cannot be reactivated below the BDT by increasing the pore pressure alone and requires localization. If pore pressure increases and the fault localizes back to 1 mm, then brittle behavior can occur to a depth of around 35 km. The interdependencies among effective stress, contact scale strain rate and pore pressure allow estimates of the conditions necessary for deep low frequency seismicity seen on the San Andreas near Parkfield and in some subduction zones. Among the implications are that shear in the region separating shallow earthquakes and deep low frequency seismicity is distributed and that the deeper zone involves both elevated pore fluid pressure and localization.

1. Introduction

Studies of crustal faulting and rock friction nearly always assume the effective normal stress σ_n^e that determines fault shear resistance during frictional sliding is the difference between applied normal stress, σ_n, and pore pressure, p,

$$\sigma_n^e = \sigma_n - p \tag{1a}$$

[*Terzaghi*, 1936; 1943]. This effective stress principle is known to hold at low confining stress and low temperature in laboratory experiments [*Handin et al*., 1963; *Brace and Martin*, 1968] and provides an important explanation for the apparent weakness of some natural faults, particularly low angle reverse faults [*Hubbert and Rubey*, 1959; *Mandl*, 1988; *Wang and He*, 1994]. Nonetheless, there is a limit to (1a), a depth below which rocks undergo ductile flow regardless of the value of effective stress. While often the depth limit is equated with the 'percolation threshold', the point at which porosity transitions from an interconnected network to a series of isolated pores [*Zhu et al*., 1995], some high temperature, high confining pressure experiments with interconnected but lithostatic pore pressure deform by ductile creep [*Hirth and Kohlstedt, 1995*], suggesting that the limit is not uniquely related to percolation. Thus, there is no comprehensive laboratory data or theory that allows estimates of the limit of the effective stress principle in the Earth's crust. The purpose of the present study is to develop methods with which to estimate effective stress throughout the lithosphere using friction theory and published results from laboratory rock deformation. The resulting model for effective stress was suggested schematically by *Thomas et al*. [2012] (see their Figure 15) and is a refinement of the qualitative development of *Hirth and Beeler* [2015]. Throughout we use the adjective 'deep' to mean near and below the transition between brittle faulting and ductile flow (BDT). In particular to understand the role of pore fluid pressure, we focus on its mechanical role in controlling brittle faulting and the location of the BDT.

Limited understanding of the physical processes that influence effective pressure affects depth estimates of the BDT, the rheological transition that determines the depth limit of shallow crustal seismicity. It is the role of effective stress in determining the depth extent of brittle faulting and seismicity that is the primary application in our study. Typically the BDT is estimated as the intersection of a ductile flow law whose strength decreases strongly with increasing temperature and a frictional fault whose shear strength is $\tau = \mu\sigma_n^e$, where μ is the friction coefficient and σ_n^e obeys equation (1a) (**Figure 1a**) [*Goetze and Evans*, 1979]. In this classic approach [also see *Brace and Kohlstedt*, 1980; *Kirby*, 1980], the transition from brittle to ductile deformation is assumed to be abrupt; this ignores intermediate behaviors seen in some laboratory experiments such as a switch between rate weakening and rate strengthening friction in the brittle regime [*Stesky*, 1978; *Blanpied et al*., 1995; *Chester*, 1995; *Handy et al*., 2007] and distributed semi-brittle flow [*Evans et al*., 1990] spanning the BDT. These 'transitional' regimes are omitted to simplify the analysis, allowing the possible role of pore fluid pressure in the switch between purely brittle to fully ductile flow to be emphasized. As shown here, typically the shear resistance resulting from friction is assumed to be proportional to depth such as due to both normal stress and pore pressure increasing following lithostatic and hydrostatic gradients, while μ is constant. Depth estimates therefore rely on (1a) and the case shown in **Figure 1a** for San Andreas-like conditions will be used as a reference example later in this paper.

In other cases where pore fluid pressure is elevated above hydrostatic in the deep crust, implying an increase in the depth of the BDT, physical limits on effective stress may also be important in determining the transition depth. Indeed at plate boundaries, where most of the Earth's earthquake hazard resides, geophysical evidence of deep elevated pore fluid pressure is widespread. For example, in both the Nankai and Cascadia subduction zones, high fluid pressures are inferred from V_p/V_s ratios [*Shelly et al.*, 2006; *Audet et al.*, 2009]. Similarly using magnetotelluric data *Becken et al.* [2011] image a region of low resistivity adjacent to the San Andreas fault in central California that they attribute to interconnected fluid at elevated pore pressure. In all three cases (Nankai, Cascadia, San Andreas) the regions of inferred elevated pore pressure are associated with non-volcanic tremor, long duration seismic signals with highest signal-to-noise ratios in the ~2-8 Hz band [*Obara*, 2002]. This tremor also has properties that seem to require elevated pore pressure, particularly occurrence rates that are very sensitive to small stress perturbations. Studies of static stress changes from regional earthquakes report both an aftershock-like response of deep NVT and LFEs on the SAF to increases of 6 and 10 kPa in shear stress from the 2003 M_w 6.5 San Simeon and the 2004 M_w 6.0 Parkfield earthquakes respectively, and quiescent response to decreases in stress [*Nadeau and Guilhem*, 2009; *Shelly and Johnson*, 2011]. Several studies report triggering of NVT on the SAF and elsewhere by teleseismic surface and body waves that imposed stress transients as small as a few kilopascals [*Gomberg et al.*, 2008; *Miyazawa and Brodsky*, 2008; *Peng et al.*, 2009; *Hill*, 2010; *Ghosh et al.*, 2009; *Shelly et al.*, 2011]. Additionally, studies of tidal stress perturbations conclude that NVT is sensitive to stress changes as small as fractions of a kilopascal [*Nakata et al.*, 2008; *Lambert et al.*, 2009; *Thomas et al.*, 2009; *Royer et al.*, 2015]. On the basis of laboratory determined material strength, such sensitivity to small amplitude stress change is thought to arise only for weak faults, moreover, those that have shear strengths similar to the amplitude of the stress perturbation [e.g., *Beeler et al.*, 2013], which is most easily accomplished at these depths by elevated pore fluid pressure.

In the case of Nankai and Cascadia, as well as in some other subduction zones, NVT is spatially and temporally associated with quasi-periodic intervals when fault slip accelerates well above the long-term rate over a portion of the deep extension of the subduction zone, down-dip of the inferred locked zone [e.g., *Dragert et al.*, 2001]. In Cascadia these episodic slow slip events are also sensitive to small stress changes [*Hawthorne and Rubin*, 2010], providing additional evidence of elevated pore pressure over a large areal extent of the deep fault. Because these events show recurring accelerating slip they are often modeled with modified brittle frictional earthquake models [*Liu and Rice*, 2005; *Segall and Bradley*, 2012]. To produce episodic slip with realistic recurrence intervals, slip and slip speeds, the models require elevated pore fluid pressure, providing consistency with the tidal and dynamically triggered seismicity datasets. Collectively these observations of deep NVT and slow slip with tidal correlation, indicate that in at least a portion of deep crust equation (1a) applies and that brittle frictional sliding is the predominant faulting mechanism.

Most relevant to our interest in the BDT in the present study, seismicity in these locations is not continuous with depth and the distribution provides key constraints on fault rheology. Seismicity is partitioned into two separate and distinct seismic zones. On the San Andreas there is seismicity above 10 km with typical earthquake source properties and a deeper region between 15 km and 30 km depth with low frequency earthquakes and tectonic tremor [*Shelly and Hardebeck*, 2010]. A perhaps related structure is suggested by collected work in Cascadia on the composition and mechanical properties of the fault [*Wang et al.*, 2011], non-volcanic tremor [*Wech and Creager*, 2008] and geodetic inversions for the megathust earthquake locking depth [*McCaffrey et al.,* 2007; *Burgette et al.*, 2009; *Schmalze et al.*, 2014]. In that body of literature, there is separation between the estimated extent of the locked zone of the megathrust earthquake and the region of active deep episodic slip that is accompanied by tectonic tremor. Studies of borehole strain [*Roeloffs et al*., 2009; *Roeloffs and McCausland*, 2010] and GPS [*Bartlow et al*., 2011] show that in deep slip events in northern Cascadia between 2007 and 2011, the up-dip limit of episodic slip is around 50 km east-northeast of the estimated down-dip limit of the locked zone [*Yoshioka et al.*, 2005; *McCaffrey et al.,* 2007; *Burgette et al.*, 2009]. Notably slip in these episodic events produces a shear stress concentration on the fault up-dip of the slip zone, but generates no post-slip event seismicity on this most highly stressed shallow extension. This suggests that the region between 10 and 15 km depth is ductile.

So, again using the San Andreas as an example, instead of a single BDT as in **Figure 1a**, seismicity defines a shallow BDT at around 10 km depth, a transition back to brittle behavior at around 15 km (DBT) and a second BDT at approximately 30 km. This distribution of seismicity obviously reflects varying mechanical properties. In other examples of double seismic zones, the separation is attributed to a rheological contrast at the crust mantle boundary [*Chen and Molnar*, 1983]; that interpretation does not apply here. More likely the second seismic zone that hosts NVT on the San Andreas is a region of frictional sliding following the effective stress principle, equation (1a), activated by elevated pore fluid pressure. Those are the conditions used in **Figure 1b** to calculate a double brittle zone, for which the pore fluid pressure gradient is elevated to 27.6 MPa/km for depths below 16 km. This second reference case for San Andreas-like conditions is used later in this paper to consider the role of effective stress in transitions between brittle and ductile faulting in the lithosphere.

In this paper, the model developed to estimate effective stress is constructed by combining a contact-scale force balance in which effective stress is controlled by the fractional contact area across faults [*Scholz*, 1992; *Skempton*, 1960] with experimental observations from static friction tests that relate the fractional contact area to the ratio of the material yield strength to the applied normal stress [*Dieterich and Kilgore*, 1994; 1996]. The pore fluid pressure in the fault zone at any depth is assumed to be constant. This approach that was developed in an earlier study [*Hirth and Beeler*, 2015] using a uniaxial stress state (consistent with the *Dieterich and Kilgore* [1996] experiments) is expanded here to the stress state associated with frictional sliding by using the assumptions of contact-scale yielding and a constant macroscopic friction coefficient. This portion of the analysis is found in section **3** (**A general effective stress relation**) and follows a brief review of laboratory constraints on effective stress for frictional sliding and rock fracture (section **2. Experimental constraints on effective stress**). For the model, effective stress depends on the rate of contact scale yielding and thus is related to the macroscopic strain rate. Since fault slip rates during the seismic cycle vary from much less than the plate rate (~0.001 μm/s on the San Andreas) to ~ 1 m/s during seismic slip, to make the analysis tractable we consider slip at the plate rate with a steady-state shear resistance and a constant shear zone thickness. This approach follows from the previous studies of crustal stress and strength [*Goetze and Evans*, 1979], as in **Figure 1**. Using data on dilatancy and compaction from room temperature friction experiments we assume a dynamic balance between on-going contact-scale yielding and shear induced dilatancy to relate macroscopic shear strain to contact-scale strain and thus to the yield stress at contacts, as discussed in section **4** (**Relations between contact scale and macroscopic strain rates**). The necessary laboratory data and flow laws for quartz yield stress as a function of temperature and strain rate are assembled in section **5** (**Yield strength of asperity contacts**). Finally, effective pressure is calculated throughout the lithosphere for comparison with the two reference cases (**Figures 1a and 1b**) in section **6** (**Results**). Our analysis suggests that a highly efficient effective stress is restricted to portions of the crust where the yield strength of asperity contacts within fault zones greatly exceeds the applied normal stress. Because yield strength decreases with increasing temperature and decreasing strain rate, a highly efficient effective pressure coefficient is more difficult to maintain at depths where temperature is high and deformation is distributed. Accordingly, the effective stress in the deep crust tends to the applied normal stress unless both the shear strain rate and pore pressure are elevated.

2. Experimental constraints on effective stress

The concept of effective stress,

$$\sigma^e = \sigma - \alpha p, \tag{1b}$$

was discovered in soil mechanics experiments by *Terzaghi* between 1919 and 1925, [e.g., *Terzaghi*, 1936; 1943]. Here σ^e is the effective stress, σ is applied stress, p is pore pressure and α is the effective pressure coefficient, $0 \le \alpha \le 1$. The underlying principle is that for materials with interconnected porosity, fluid pressure within the pore space works in opposition to the applied stresses. Stress dependent properties (frictional strength, elastic compressibility, poroelasticity) are changed relative to fluid-absent values. The α coefficient characterizes the efficiency of the pore fluid in opposing the applied stress. There are many different specific effective stress relationships [*Skempton*, 1960; *Nur and Byerlee*, 1971; *Robin*, 1973]. For example, for a particular material at specified normal stress, temperature, and pore pressure, effective stress for poroelasticity (Biot's effective stress) [*Rice and Cleary*, 1976; *Cheng*, 1997], volumetric strain [*Geertzma*, 1957; *Skempton*, 1960; *Nur and Byerlee*, 1971], seismic velocity [*Gurevich*, 2004], friction [*Hubbert and Rubey*, 1959; *Mandl*, 1988; *Hirth and Beeler*, 2015], and pore strain [*Robin*, 1973], all have the form of (1b) with different values of α. Like *Terzaghi,* in the present study we are interested strictly in effective stress for shear failure, in which case σ is stress normal to the shear zone, σ_n, and (1b) is the effective stress law for frictional sliding with an effective pressure coefficient denoted α_f throughout.

In many previous low temperature studies of natural faulting and laboratory rock friction where effective normal stress is considered, α_f is found or assumed to be 1, leading to the standard effective normal stress relation for faulting (1a) [e.g., *Hubbert and Rubey*, 1959; *Mandl*, 1988] sometimes referred to as *Terzaghi*'s effective stress. Equation (1a) well characterizes intact rock failure in experiments on granite, diabase, dolomite, gabrro, dunite, and sandstone at room temperature [*Brace and Martin*, 1968] and on dolomite, limestone, sandstone, siltstone and shale at temperatures up to 300°C [*Handin et al*., 1963]. There are known limitations to (1a) that the rock must be inert in the pore fluid, and the fluid is drained and pervasive. High strain rate loading tests [*Brace and Martin*, 1968] show an apparent breakdown of (1a) when the rate of dilatancy exceeds the rate that fluid flows into the incipient fault, resulting in undrained conditions and a dilatancy hardening contribution to the failure strength. In this case the externally measured pore pressure is not the pore pressure in the fault and the effective normal stress is unknown (but can be inferred from the observed shear stress). To meet the requirement of drained deformation and pervasive saturation, the rock must be sufficiently porous and permeable. *Handin et al*.'s [1963] experiments show breakdown of $\alpha_f = 1$ in presumed cases of low permeability (undrained deformation, shales) and low porosity (non-pervasive fluid, dolomite, marble, limestone). Because rock failure at low temperature involves dilatancy that favors high permeability and pervasive fluid distribution [*Brace et al.*, 1966], the requirements for (1a) to apply are expected at typical laboratory faulting conditions where strain rates are intermediate between tectonic and seismic rates. Limited stick-slip failure and frictional sliding experiments on preexisting faults at room temperature on a range of materials, e.g., on sawcut surfaces of granite [*Byerlee*, 1967] and simulated gouges of illite and montmorillonite [*Morrow et al*, 1992], also confirm (1a).

However, near the BDT ductile deformation tends to reduce porosity and permeability, leading to an expected breakdown of (1a) in the form of a reduction in α_f, as seen in low porosity rocks by *Handin et al*. [1963] and references therein. Similarly, in more recent high temperature, high pressure laboratory experiments some rocks exhibit ductile deformation in the presence of near-lithostatic pore pressure [*Chernak et al.*, 2009] or near-lithostatic melt pressure [*Hirth and Kohlstedt, 1995*], rather than brittle failure at near zero shear resistance as required by (1a) [*Hirth* *and Beeler*, 2015]. There are some natural counterparts of these experiments, mylonites with near lithostatic pore pressure inferred from fluid inclusions [*Axen et al.*, 2001]. These observations suggest that under some conditions the BDT is associated with an effective stress relation with $\alpha_{\square}$ near zero, instead of the fully efficient coefficient (1a) and that the change in $\alpha_{\square}$ is expected as porosity decreases in the deep crust.

In contrast to these scattered laboratory observations that suggest an "ineffective" effective pressure at some mid-crustal conditions, observations of microseismicity and tectonic tremor on the deep extent of some subduction zones and the San Andreas fault (detailed in the **Introduction)**, particularly the modulation of fault slip and tectonic tremor by kPa or smaller tidal stresses [e.g., *Hawthorne and Rubin,* 2010; 2013, *Thomas et al.*, 2009; 2012], are difficult to explain without allowing friction to operate in the presence of elevated pore pressure with $\alpha_{\square}$ near one. In light of conflicting seismic, field and laboratory evidence, some of which suggests limits on (1a), collectively the observations suggest that the effective pressure coefficient α_f can be near zero or near 1 depending on the circumstances. Though cause-effect relations are unknown, likely controls on α_f involve material properties such as ductile strength, and environmental variables such as pore pressure, temperature, normal stress, and strain rate. To develop a model for effective stress, in the following section we extend to crustal temperatures and stresses a physical model of effective stress derived from a contact scale force balance [*Skempton*, 1960; *Scholz*, 1990].

3. A general effective stress relation

Imagine a representative asperity contact surrounded by fluid at pore pressure *p* on a fault surface or within a shear zone (**Figure 2**). Here and throughout this paper, pore fluid pressure in the fault zone is assumed to be constant, in full communication with the surroundings (drained). The macroscopic force applied normal to the asperity *N* is balanced by the normal force at the solid-solid asperity contact N_c and the pressure in the pore space [*Skempton*, 1960]:

$$N = N_c + \left(A - A_c\right) p \tag{2a}$$

where A_C is the solid-solid contact area and A is the total area measured in the plane parallel to the contact. Normalizing by the total area, defining the macroscopic normal stress, $\sigma_n = N/A$, leads to a definition of effective normal stress, $\sigma_n^e = N_C/A$, as

$$\sigma_n^e = \sigma_n - \left(1 - \frac{A_c}{A}\right)p, \tag{2b}$$

an equation of the form (1b) with $\alpha_f = 1 - \frac{A_c}{A}$ [*Skempton*, 1960; *Scholz*, 1990]. Noting that the contact normal stress is $\sigma_C = N_C/A_C$, the ratio of σ_n^e to σ_c for this model is the fractional contact area,

$$\frac{\sigma_n^e}{\sigma_c} = \frac{A_c}{A}, \tag{2c}$$

similar to classic plastic and elastic models of friction [c.f., *Bowden and Tabor*, 1950; *Greenwood and Williamson*, 1966]. In (2b), the effective stress for friction is thus related to the area along a shear plane that is supported by pressurized pore space relative to area of asperity contact across the plane. When the area of contact is small a change in pore pressure acts in nearly exact opposition to the applied fault normal stress. Conversely when the pore space is small and equi-dimensioned, changes in pore pressure produce nearly no opposition. Here and throughout this report we assume that the contact stresses are limited by plastic yielding [*Bowden and Tabor*, 1950] and that the contacts between grains are not wetted by the pore fluid.

To get a qualitative idea of how α_f estimated from (2) might vary with depth in the Earth's crust, first consider a rough fault surface uniaxially loaded in true static contact (no resolved shear stress onto the fault) with no confining pressure ($\sigma_3 = 0$) and dry as in the experiments of *Dieterich and Kilgore* [1996]. The macroscopic principal stresses are coincident with the fault normal and in-plane directions; the fault normal stress is $\sigma_1 = \sigma_n$ (**Figure 3a**). The corresponding stress state at a representative contact on the fault is in the same orientation as the macroscopic stress (**Figure 3b**); the contact normal stress is the greatest principal stress and also is the differential stress at the asperity contact. Plasticity on the contact scale requires the contact normal stress is also the yield stress, $\sigma_c = \sigma_1^c = \sigma_\Delta^c = \sigma_y$ (**Figure 3b**). Fractional contact area is

$$\frac{A_c}{A} = \frac{\sigma_n}{\sigma_y}. \tag{3a}$$

Direct measurements of contact area for minerals and analog materials at room temperature show this to be valid [*Dieterich and Kilgore*, 1996]. Though (3a) is only strictly applicable to true static conditions of no shear stress on the fault, using (2c), the implied effective pressure coefficient is

$$\alpha_f = 1 - \frac{\sigma_n^e}{\sigma_y}. \tag{3b}$$

[*Hirth and Beeler*, 2015]. Observations in laboratory tests on strong materials such as granite and quartz at a few to hundreds of MPa normal stress at room temperature are qualitatively explained by (3b). α_f = 1 is found at room temperature regardless of confining pressure [*Byerlee*, 1967] or rock type [*Morrow et al.*, 1992]. σ_y for quartzofeldspathic minerals at room temperature is several GPa [*Dieterich and Kilgore*, 1996]. Even extrapolating to normal stresses of 500-800 MPa appropriate for the deep crust, we still expect $\alpha_f \approx 1$ at room temperature. So at low temperature faults the fractional area of contact is very small.

The uniaxial compression contact scale stress state used to derive (3b) is not consistent with that expected during frictional sliding. To include a macroscopic applied shear stress during slip at elevated confining stress we make an additional explicit assumption of steady-state frictional sliding $\mu = \tau / \sigma_n^e$. Because fluid in the pore space supports no shear stress, applying a shear force balance to the contact model (**Figure 2**) requires the macroscopic applied shear force S equals the contact shear resisting force, S_c. This leads to the same type of proportionality between the macroscopic shear stress, $\tau = S/A$, and the contact scale shear stress, $\tau_c = S_c/A_c$, seen in equation (2c) for the normal stresses, namely, $\tau = \tau_c A_c/A$. One consequence is that the ratio of the contact shear and normal stresses is the macroscopic friction coefficient, $\tau_c/\sigma_c = \mu$, again consistent with familiar assumptions from friction theory [*Bowden and Tabor*, 1950; *Skempton*, 1960; *Greenwood and Williamson*, 1966]. A more general consequence is that all of the macroscopic stress components on the fault such as the effective normal stress (σ_n^e), the effective confining stress (σ_3^e) and the greatest principal stress (σ_1^e) (**Figure 3c**), scale from the analogous contact stresses (**Figure 3d**) by the area ratio. Similarly, the macroscopic stresses relate to the material yield stress via the area ratio and a constant, χ, specific to the stress component of interest, as

$$\frac{A_c}{A} = \frac{\sigma^e}{\chi\sigma_y}. \tag{3c}$$

□□□ particular value of χ can be determined from the Mohr construction shown in **Figure 3d**. For example the contact-scale normal stress is $\sigma_c = \sigma_y \cos\left(\tan^{-1}\mu\right)/2\mu$. From equation (2c), then, $\chi = \cos\left(\tan^{-1}\mu\right)/2\mu$.

The contact stress state, derived from the force balance and the assumptions of contact yielding and steady-state sliding at a macroscopic, constant friction coefficient differs in detail from the expected stress state at a representative contact on a sliding frictional interface. For example in Hertz's solution for a uniaxially loaded elastic contact, normal stress varies within the contact from zero at the edges to approximately 1.3 (4/π) times the mean at the contact center [*Johnson*, 1987]. Imposed sliding further alters the stress distribution to be asymmetric about the contact center with relative tension and compression at the trailing and leading edges, respectively. An example of these complications, that are ignored in our representative contact model, are described in more detail in the Supplement 4. There, a solution for a sliding contact from the contact mechanics literature is developed and compared with that from our model. A primary concern is whether the average stress model adequately characterizes the stress state at yield. The supplementary analysis suggests that if spatial variation and asymmetry in the contact stress are considered, differential stress at yielding during slip is within 10% of the representative contact model. Nevertheless, that analysis should be considered just one example of the possible contact stresses during slip, and the size and distribution of deviations from the average stress state during sliding requires further laboratory and theoretical research, especially at high-temperature conditions where crystal plastic deformation mechanisms become kinetically more efficient. Additional considerations and guidance in future work relating contact stress state to macroscopic shear resistance during frictional sliding may be found in the study of *Boitnott et al.* [1992] and references therein.

Throughout the remainder of this paper, we use the representative contact model (**Figure 2**) to characterize the average shear and normal stresses at the contact. Issues that arise in true contact mechanics models such as spatial variability of shear and normal stresses within the contact, asymmetry of the stresses about the contact [*Johnson*, 1987] and interactions between contacts are not considered. The general form for the resulting effective stress coefficient is

$$\alpha_f = 1 - \frac{\sigma^e}{\chi \sigma_y}, \tag{3d}$$

Accounting for physical limits on $\alpha_{\square}$, the general form of a bounded ($0 \le A_c/A \le 1$, $0 \le \alpha_f \le 1$) effective stress law for faulting is

$$\begin{array}{ll} \alpha_f = \dfrac{\chi\sigma_y - \sigma}{\chi\sigma_y - p} & \chi\sigma_y > \sigma \\ \alpha_f = 0 & \chi\sigma_y \le \sigma \end{array}, \tag{4a}$$

which follows from combining (1b) with (3d) and solving for α_f. From inspection, at low values of σ_y relative to t□□□□□ress component of interest, $\alpha_f \approx 0$, and at high values $\alpha_f \approx 1$. Physically, once the macroscopic differential stress reaches the yield stress, the contact area is equal to the total area ($A_c/A = 1$). This limiting condition on effective stress ($\alpha_f = 0$) at elevated temperature and stress occurs when $\chi\sigma_y \le \sigma$. The limit is independent of pore pressure and implies that in porous and permeable materials there is a depth below which friction cannot determine fault strength, even when the pore fluid pressure approaches lithostatic, consistent with the limited laboratory data [*Chernak et al.*, 2009; *Hirth and Kohlstedt, 1995*]. The general relation for effective stress is

$$\begin{array}{ll} \sigma^e = \dfrac{(\sigma - p)}{\left(1 - \dfrac{p}{\chi\sigma_y}\right)} & \chi\sigma_y > \sigma \\ \sigma^e = \sigma & \chi\sigma_y \le \sigma \end{array} , \tag{4b}$$

which results from combining (1b) with (3d) and solving for effective stress.

Accordingly, to calculate effective stress requires specified values of the environmental variables, pore pressure and applied stress, and knowledge of the material yield stress. The yield stress also depends on the environment via temperature and fundamentally on the strain rate. Since fault slip rates during the seismic cycle vary from much less than the plate rate (~0.001 μm/s on the San Andreas) to ~ 1 m/s during seismic slip, to make the analysis tractable in this study we consider slip at the plate rate at a steady-state shear resistance and constant shear zone thickness. Thus, in the calculations the strain rates are constant. This approach follows from previous studies of crustal stress and strength inferred from experimental data [*Goetze and Evans*, 1979; *Brace and Kohlstedt* 1980; *Kirby,* 1980] (**Figure 1**). While the dependences of yield stress on temperature and strain rate have been established in laboratory tests at controlled temperatures and macroscopic strain rates, the appropriate strain rate for use in (4b) is the fault normal strain rate due to yielding at the asperity contacts. In the next section we apply friction theory at steady state to determine a relation between the macroscopic steady-state shear strain rate and the macroscopic fault normal strain rate. Then we use the macroscopic normal strain rate to determine the contact-scale normal strain rate due to yielding.

4. Relations between contact scale and macroscopic strain rates.

Following our assumption of steady-state deformation we assume that during frictional sliding the shear zone has constant volume and that there is no change in thickness or porosity with slip. This assumption is reasonably well approximated in large displacement friction experiments [e.g., *Beeler et al*., 1996]. To estimate the necessary value of the contact scale normal strain rate due to yielding that determines the area of contact we use friction theory and laboratory observations made far from steady-state. During frictional sliding at room temperature, fault zone porosity varies with sliding rate [e.g., *Morrow and Byerlee*, 1989; *Marone et al.*, 1990]. When the fault is sliding at steady state, there is essentially no displacement normal to the fault. If the imposed sliding velocity is changed, the fault dilates or compacts as observed in the single asperity study of *Scholz and Engelder* [1976] due to changes in the contact area. Although quartz has a yield strength of more than 10 GPa at room temperature [*Evans*, 1984], indentation studies show that the contact scale creep rate is easily measurable, and even at 25°C the observations of dilation and compaction during frictional sliding can be interpreted to result from a dynamic balance between time-dependent compaction (due to fault normal yielding at the asperity contacts) and shear-induced dilatancy. These two opposing effects have been observed in lab faulting tests on initially bare rock surfaces, notably by *Worthington et al.* [1997] (**Figure 4**). Since during steady-state sliding the fault normal displacement δ_n is constant, $d\delta_n = 0$, the dynamic balance between opposing time-dependent normal yielding and shear-dependent dilation can be written in terms of the macroscopic normal and shear strains, ε_n and γ, as

$$\left(\frac{\partial \varepsilon_n}{\partial \gamma}\right)_t^{ss} = -\frac{1}{\dot{\gamma}}\left(\dot{\varepsilon}_n\right)_\gamma^{ss},$$

or in terms of slip δ_s and fault normal displacement as

$$\left(\frac{\partial \delta_n}{\partial \delta_s}\right)_t^{ss} = -\frac{1}{V}\left(\frac{\partial \delta_n}{\partial t}\right)_{\delta_s}^{ss}, \tag{5a}$$

[*Beeler and Tullis*, 1997]. Here V is the imposed sliding velocity.

The nature of the competition makes it difficult to measure either of the steady-state rates in (5) directly. However, a minimum rate of shear-induced dilatancy may be inferred from measurements during frictional sliding in which the competing rate of fault normal creep has been induced to be very low. Such a situation arises during reloading following a long duration stress relaxation test. During the relaxation test, the loading velocity is zero, however the fault continues to slip under the shear load, and as the fault slips, the measured strength decreases. This is accompanied by compaction that is logarithmic in time [e.g., *Beeler and Tullis*, 1997]

(**Figure 4a**). The compaction is presumed to be due to fault-normal creep at asperity contacts. At the end of the long relaxation the normal creep rate is very low. In the subsequent reloading the fault dilates with displacement (**Figure 4b and 4c**). The measurements are made at large displacements >100 mm and large shear strains, typically > 1000. Dilatancy and compaction measured in those experiments have no known displacement dependencies, however, there are no comprehensive studies of these effects. The examples shown in **Figures 4** are from initially bare surfaces of granite and quartzite at room temperature and 25 MPa normal stress. The displacement rate of dilation, $d\delta_n/d\delta_s \approx 0.1$ for granite and is ~0.06 for quartzite. Because there may be contributions from time dependent compaction during these reloading tests, we can infer that the steady-state rate $\left(\partial\delta_n/\partial\delta_s\right)_t^{ss}$ is no smaller than 0.06. These values are similar to those inferred by theoretical treatments of the kinematics of frictional sliding [*Sleep*, 2006] that yield values between 0.04 and 0.11 for quartz and a preferred value in the range 0.04 to 0.05. The approaches of *Sleep* [1997; 2006] and *Sleep et al.* [2000] are similar to (5a) in that during steady-state sliding time-dependent compaction is balanced by shear induced dilatancy.

Using the data in **Figure 4** and equation (5a), the macroscopic normal strain rate $\dot{\varepsilon}_n$ due to yielding at asperity contacts is assumed to be ~10% of the shear strain rate $\dot{\gamma}$. The contact-scale normal strain rate $\dot{\varepsilon}_n^c$ is greater than or equal to the macroscopic normal strain rate, and varies systematically with percent contact area as $\dot{\varepsilon}_n^c = \dot{\varepsilon}_n A/A_c$. Combining with (5a), the contact scale fault normal strain rate due to yielding is

$$\dot{\varepsilon}_n^c = 0.1\dot{\gamma}\frac{A}{A_c},$$

or, equivalently

$$\dot{\varepsilon}_n^c = 0.1\dot{\gamma}\left(1-\alpha_f\right), \tag{5b}$$

the strain rate with which to determine the yield stress. Much of the variation in the effective stress coefficient (4a) illustrated in the calculations described later in this paper arise directly from assumed changes in the shear zone thickness (strain rate). The other primary variations in the effective stress (4b) and the effective stress coefficient (4a) are due to the temperature dependence of the yield stress, which we describe next.

5. Yield strength of asperity contacts.

The yield strengths of crustal minerals typically have a very the strong temperature dependence which implies a strong depth dependence in the effective pressure relation (4). For example, at the base of the seismogenic zone where the temperature is several hundreds of degrees C, the yield stress of quartz approaches the applied confining stress [*Evans and Goetze*, 1979; *Evans,* 1984]. For our purposes to estimate the asperity yield strength at low temperature (red dashed) we use quartz data from indentation (solid symbols) and triaxial (open) tests (**Figure 5**) [*Evans*, 1984; *Heard and Carter*, 1968]. These experiments were conducted at strain rates on the order of 1 x 10^{-5}/s. At the lowest temperatures, the data are represented by a flow law for low-temperature plasticity (LTP) from *Mei et al.* [2010] that is described in more detail in the Appendix. *Evans'* [1984] experiments were conducted dry. A complication is that while quartz undergoes some kind of plastic yielding at low temperature [*Masuda et al.*, 2000], the mechanism is not strictly the dislocation glide assumed in the *Mei et al*. [2010] flow law at low temperature. Nonetheless the flow law can fit the data quite well and we use it empirically. To account for weakening due to the presence of water in the Earth's crust, in the absence of experimental data at saturated, low stress conditions, the wet strength (blue dashed) is somewhat arbitrarily assumed to be half the dry strength in the low temperature regime. At around 800°C the data depart from the trend of low temperature plasticity. This is the onset of dislocation creep. The dislocation creep flow law for dry deformation (red dotted line in **Figure 5**) used is of the standard form [*Hirth et al.,* 2001]. As with the low temperature plasticity data, it is necessary to consider the effect of water on the creep flow strength; in this case there are data from wet creep tests, represented by the flow law (blue dotted) using parameters from *Hirth et al.* [2001]. To produce a combined flow law for contact yielding (solid curves) we use a standard assumption that the combined differential strength is $\sigma_\Delta^c = \left(1/\sigma_\Delta^{LTP} + 1/\sigma_\Delta^{DC}\right)^{-1}$. To extrapolate the indentation data to the Earth we use the wet flow laws at the appropriate contact scale strain rate. Application of these flow laws on the asperity scale implicitly ignores any transitional semi-brittle deformation mechanisms that are observed in large strain experiments [*Evans et al*., 1990]

6. Estimating α_f and the position of the BDT

The objective of this study is to estimate the position of the BDT while accounting for effective stress using equation (4). As described in the immediately preceding sections, effective stress depends on material properties, thermal structure, strain rate, and stress regime. The BDT depends on these same variables directly [*Goetze and Evans*, 1979; *Brace and Kohlstedt*, 1980] and also via the effective stress. Our strategy is to assume a thermal structure, stress regime, pore pressure, depth variations in shear-zone thickness, and a particular material (quartz). There are two example calculations in this section. The calculations correspond to the same thermal structure, stress state and material as the cases shown for the standard effective stress assumption ($\alpha_f = 1$) in **Figure 1;** these previous plots serve as the two reference calculations for comparison with the examples with equation (4). Furthermore, between the two following calculations, only the pore pressure and thickness distributions differ; all other environmental variables and material properties are the same. Pore pressure at any depth within the fault zone is assumed to be constant. The calculations do not consider the percolation threshold and it is assumed that the pore space is interconnected for all porosities greater than zero. While this is not ideal - some of the related issues are described in the Discussion section. The calculations are for a vertical strike-slip faulting environment with a lithostat that is typical for the continental crust. Overburden is 28 MPa/km and is assumed equal to the average of the greatest and least principal stresses, $\sigma_m = (\sigma_1 + \sigma_3)/2$. The temperature distribution is from *Lachenbruch and Sass* [1973] (Model A) for the San Andreas. Fault normal stress for constant friction and an optimally oriented fault (**Figure 3c**) is

$$\sigma_n = \alpha_f p + \left(\sigma_m - \alpha_f p\right)\frac{\sin\left(\tan^{-1}\mu\right)\cos\left(\tan^{-1}\mu\right)}{\mu}. \qquad (6a)$$

The differential stress is

$$\sigma_\Delta = 2(\sigma_m - \alpha_f p)\sin\left(\tan^{-1}\mu\right)$$

or

$$\sigma_\Delta = \frac{2\tau}{\cos\left(\tan^{-1}\mu\right)}. \tag{6b}$$

Combining equation (6a) and (4a) for normal stress ($\sigma = \sigma_n$) results in a compact expression for the effective pressure coefficient for friction in strike slip,

$$\begin{aligned} \alpha_f &= \frac{\sigma_y - 2\sin\left(\tan^{-1}\mu\right)\sigma_m}{\sigma_y - 2\sin\left(\tan^{-1}\mu\right)p} && \sigma_y > 2\sin\left(\tan^{-1}\mu\right)\sigma_m \\ \alpha_f &= 0 && \sigma_y \le 2\sin\left(\tan^{-1}\mu\right)\sigma_m \end{aligned}. \tag{7}$$

The shear zone differential stress is given by the same flow laws used to estimate the contact asperity yield strength. The position of the BDT is estimated as the intersection of the friction and flow stress relations, assuming failure at the lower of the differential strength of friction or flow, $\sigma_\Delta = \min\left(\sigma_\Delta^{friction} + \sigma_\Delta^{flow}\right)$. The long term macroscopic shear strain rate $\dot{\gamma}$, is the plate rate, for which we use a San Andreas-like value, V_L = 0.001 μm/s (corresponding to 31.5 mm/yr), divided by the shear zone thickness *w,* which we take to be ~1 mm in the brittle regime [*Chester and Chester*, 1998] and 1 km below the BDT [*Bürgmann and Dresen*, 2008]. These thickness choices are intended to produce illustrative results but unfortunately they are poorly constrained. These applied strain rates of 1 x10^{-6} /s and 1 x 10^{-12} /s result in macroscopic fault-normal strain rates of $\dot{\varepsilon}_n$= 1 x 10^{-7}/s and 1 x 10^{-13}/s, following the discussion in section 4 above. The strain rates for friction assuming a 1 mm thick shear zone are similar to those in the laboratory tests.

In the first calculation, pore pressure is hydrostatic (10 MPa/km) throughout the lithosphere. **Figure 6** shows α_f □□□□□□ and differential stress (black) from friction (red) and from ductile flow (green). At the BDT there is a large change in the assumed shear zone thickness resulting in a large corresponding change in the fault zone strain rates. This produces a large change in fractional contact area (right panel) and a corresponding change in α_f from high values associated with localized, dilatant frictional slip (grey) to zero associated with non-dilatant distributed ductile shear (yellow).

When compared with the results from the standard assumption about effective stress (**Figure 1**) there are both strong similarities and significant differences: 1) α_f is close to 1 very near the Earth's surface and decreases progressively but weakly with depth; 2) α_f remains relatively large immediately above the BDT because the asperity scale deformation is controlled by low temperature plasticity and the asperities are very strong; 3) because of the small difference between α_f compared with the standard assumption, the brittle ductile transition depth of ~13 km is only very weakly influenced by effective stress; 4) however, at and below the BDT $\alpha_f = 0$. This is a consequence of the much lower strain rate due to ductile flow within the assumed 1-km-wide shear zone and a transition to the much weaker dislocation creep regime on the asperity scale. The large difference between effective stress for localized frictional slip (w=1 mm, grey) and for ductile distributed shear (w= 1 km, yellow) highlights the shear strain rate effect on effective stress. Because α_f is zero on the deep extent of the fault, it is impossible to reactivate friction at these depths by raising pore pressure to lithostatic without also invoking a mechanism that imposes localized slip, the shear strain rate increases and the effective stress coefficient increases. Such localization might occur by imposing a high slip rate on the deep extent of the fault, for example, due to propagation of earthquake slip through the BDT during large earthquakes [e.g., *King and Wesnousky*, 2007; *Rice, Rudnicki and Platt*, 2014] or during propagating afterslip. Simply increasing the slip velocity at constant shear zone width will produce a deepening of the BDT itself, an increase in α_f, and an increase in the limiting depth where $\alpha_f = 0$ (equations (4) and (7)). Thus, despite the implied barrier to reactivation of friction at depth, any 'dynamic' effective pressure coefficient will be higher than estimated in **Figure 6**.

Another way that localization might be encouraged on the deep extent below the BDT would be an increase in pore fluid pressure in a limited portion of the broader shear zone. Examples of increased pore pressure localized along a specific horizon might involve migration up the fault from depth [*Rice*, 1992] or from local dehydration as is thought to be common in subduction zones [*Peacock*, 2009; *Peacock et al*., 2011].

6.1 Elevated pore pressure in the deep crust. The second calculation follows **Figure 1b**, and examines the implication of the model effective stress relation (4) for generating rheological contrasts as pore pressure and localization are varied in the deep crust. As described in the introduction, evidence for elevated pore fluid pressure is widely observed and generally expected in the deep crust. Elevated pore fluid pressure will tend to significantly increase the effective pressure coefficient in (4a) by making the denominator smaller. This is the mechanical effect of increased pore pressure itself on the effective stress coefficient. Adding the region of elevated pore pressure and assuming localized frictional slip at depths greater than 16 km produces a second brittle region (**Figure 7**). In the crust above 16 km all properties are identical to the calculation shown in **Figure 6** where pore pressure is hydrostatic. Below 16 km the pore pressure is nearly lithostatic and the shear zone is 1 mm thick. In this calculation the lithostat is 28 MPa/km and the pore pressure below 16 km is 27.6 MPa/km. At 16 km depth the pore pressure is 6.5 MPa less than lithostatic. The increase in pore pressure and decrease in the shear zone thickness results in an increase in α_f from 0 to nearly one and a more than order-of-magnitude decrease in the differential stress. The increase in α_f is due to the large magnitude increase in the contact scale strain rate from narrowing the shear zone from 1 km width to 1 mm and also due to the increase in pore pressure in the denominator of equation (4). The decrease in macroscopic strength corresponds to a transition from ductile to brittle possibly allowing for seismicity in the otherwise ductile deep crust. The potentially seismic zone persists to around 30 km depth, in contrast to the standard calculation (**Figure 1b**) where brittle deformation extends to 35 km. Between 16 and 30 km the contact scale deformation follows the low temperature plasticity relation. The narrow 'gap' region between the two brittle regions is a zone of imposed distributed creep.

Figure 7 depicts a situation that is little different from scenarios proposed in prior modeling studies where elevated pore pressure is often invoked to reactivate friction on a portion of a fault below the BDT [e.g., *Segall and Bradley*, 2012]. The primary difference is that the transitions between brittle and ductile are calculated in the present study. Their locations reflect contact-scale strength based on laboratory data and its dependence on temperature, contact-scale strain rate, the degree of shear localization, and the pore fluid pressure. There is interplay between the macroscopic fault strength and the contact scale, for example the effective pressure coefficient is determined at the contact but influences the location of the macroscopic BDT. And while the pore pressure and degree of localization are imposed in this calculation, the rheological properties dictate the ranges of localization and pore pressure necessary to reactivate friction at depth. We consider this a modest step forward. Greater advances may come from considering time-dependent rather than steady-state deformation, including time-dependent evolution of hydraulic properties and fluid pressure in the vicinity of the rheological transitions, the influence of other minerals/rock types (including those rich in micas or clays) and most importantly allowing degree of localization to be a dependent variable [e.g., *Platt et al*., 2014].

While in the calculations both elevated pore pressure and localization are required to reactivate friction below the BDT, this is not the general requirement. It is possible that some fault zone rheologies and shear zone widths allow reactivation by increasing the pore pressure alone. So long as the ductile shear zone width is sufficiently narrow that α_f for ductile shear is non-zero ($\sigma_n < \chi\sigma_y$) then increasing the pore pressure to high levels can reactivate friction. This behavior does not arise in the example (**Figure 7**) because α_f for ductile shear of a 1 km width quartz fault is zero for all depths below about 12.5 km.

7. Limitations

Despite the physical basis (**Figure 2**) and its appearance in the earthquake fault mechanics literature [*Scholz*, 1990], effective stress relations for faulting of the type described by equations (2), (3) and (4), are disputed on theoretical grounds [*Hubbert and Rubey*, 1959, 1960; *Skempton*, 1960; *Bishop and Skinner*, 1977; *Mandl*, 1988, 2000]. The supplemental materials describe these concerns in detail and how they relate to our interpretation that equation (4) is appropriate in the deep crust. Nevertheless there remain fundamental differences between our analysis and those in the soil mechanics literature that should be resolved in future theoretical and experimental studies.

Similarly, while there are a number of experimental studies that are qualitatively consistent with the decrease in α_f at high contact area that arises in our calculations [*Handin et al.*, 1963; *Hirth and Kohlstedt, 1995; Chernak et al.*, 2009] there are important counter examples. In particular, are the deformation experiments conducted by *Bishop and Skinner* [1977] to understand effective stress that find no correlation between effective pressure and contact area. These are also described in Supplementary material where we contrast and reconcile them with our view of effective stress in the deep crust. The Bishop and Skinner experiments provide the best existing constraints on the physical basis of effective stress, albeit at very low nominal effective normal stresses. Keeping in mind that the deep crust is thought to be a zone of vanishing effective stress [*Audet et al.* 2009; *Thomas et al*., 2009], experimental procedures following Bishop and Skinner could be employed in future experimental studies of effective stress at transition zone conditions to resolve the physical basis of effective stress.

Among the deficiencies of our effective stress model is the assumption of non-wetted grain boundaries. While this is consistent with the properties of quartz at elevated temperature [*Watson and Brennan*, 1987; *Beeler and Hickman*, 2015], it is not universally expected and there are other considerations. Soils that include clay minerals may have a significant fraction of grain contacts that have some form of wetted, adsorbed or bonded water within the grain boundary, conditions that favor a fully efficient effective pressure coefficient. Similar wetting properties may be associated with other sheet silicates. Another material property that may influence effective stress in fault zones at great depth is rheological anisotropy. Sheet silicates are preferentially weak for shear parallel to the basal plane and therefore may not deform by dislocation creep at any temperature [e.g., *Escartin et al.*, 1997; 2008], owing to grain-scale strain compatibility requirements. So even though they are relatively weak in the shallow crust, microcracking at the grain scale may persist well into the deep crust, at conditions where quartz and other more isotropic phases deform by dislocation creep. A consequence is that $\alpha_f > 0$ may persist to greater depths in these materials. Notably in recent experiments on serpentinite near its breakdown temperature the effective stress relationship seems to be highly efficient with interconnected porosity consisting of cleavage plane microcracks [*Proctor et al.*, 2015]. At the same time because of the anisotropy, narrow shear zones persist in phyllosilicates even at high temperatures despite ductile or rate strengthening rheological properties [e.g., *Escartin et al.*, 2008]. Thus localization defined by mineral structure such as associated with sheet silicates, rather than strictly by rheology, may be required for friction to be activated at depths below the BDT (**Figure 7**).

The model (4) assumes that α_f can be estimated at porosity approaching zero whereas an expected experimental limit on $\alpha_f > 0$ is where the porosity remains interconnected. This is consistent with observations in quartz where this percolation threshold [e.g., *Zhu et al.*, 1995] at high temperature is approximately 1 volume percent or less [*Wark and Watson*, 1998], corresponding to a permeability of ~1 x 10 $^{-14}$ m^2. In contrast, a model sphere array of grains discussed in the Supplementary provides a counter example with which to estimate the porosity and area ratio where pore space becomes isolated. The associated area ratio at the threshold is $\pi/4$ and the associated $\alpha_f = 0.22$. Consequently, rather than the smooth variation to $\alpha_f = 0$ shown in **Figure 6** at > 30 km, we may expect a more abrupt transition and a somewhat shallower limit on effective stress than estimated with (4) if the percolation threshold is the appropriate limit on effective pressure. Differences between the sphere array and the *Wark and Watson* [1998] experimental observations are related to textural equilibrium and contributions of solid-liquid surface energy to determining the pore structure and fluid percolation threshold. An additional related consideration of pore structure is dependence of the effective pressure coefficient pore shape. Low aspect ratio pores (cracks) that are favored at low temperature in the brittle regime are more compliant and at fixed porosity will produce a higher value of af than stiffer equi-dimensioned pores. In contrast at high temperatures where diffusivity is high and surface energy can be rapidly minimized, pores will be more equant.

Our effective stress model also does not consider the possibility that pore pressure might exceed the least principal stress for materials with 'cohesion', resulting in a shear resistance at zero normal stress. As the model is for steady-state frictional sliding it is consistent with no cohesion. However, below the BDT, shear zones may well develop cohesion, super-lithostatic pore pressure, and hydrofacture may be a mechanism for producing localized shear deformation. For example en echelon tensile fracture arrays generated by pore pressure exceeding σ_3 plus cohesion could evolve into a localized dilatant shear zone and reactivate friction at elevated pore fluid pressure [*Sibson*, 1996].

By neglecting semi-brittle deformation or a transition to rate strengthening friction in the brittle regime, likely we over-estimate the crustal strength near the BDT [*Evans et al*., 1990; *Chester*, 1995]. Futhermore because the semi-brittle regime involves distributed fracturing it may play a significant role in maintaining interconnected porosity near the BDT. Semi-brittle flow may lead to an increase in the effective pressure coefficient through dilatancy, but since such flow results in distributed deformation its role is difficult to evaluate without more sophisticated modeling and experiments. Nonetheless, an obvious explanation for the gap between shallow seismicity and deep NVT/LFEs on the San Andreas and in subduction zones is that this is a region of semi-brittle flow with the associated dilatancy necessary to prevent significant elevation of pore pressure above hydrostatic. Accordingly the transition back to low frequency seismicity would occur when regional, fully ductile flow begins to dominate, promoting a collapse of the pore structure, a rise in pore fluid pressure and reactivation of frictional slip at low effective stress.

Finally, of course the Earth's crust is not mono-mineralic as is assumed in the calculations in **Figures 1, 6, and 7**. Instead, rheological variability associated with differences in lithology likely plays an important part in the observed depth dependent seismicity in the deep crust [*Chen and Molnar*, 1983; *Bürgmann and Dresen*, 2008], especially in plate boundary settings such as the San Andreas and in Cascadia. For example, on the San Andreas the limiting depth of LFE occurrence is similar to the depth of the Moho. So while the calculation shown in **Figure 7** in which friction is reactivated on the deep extent of the fault implies a depth distribution of seismicity that coincides with the natural observations, it does not consider the influence of mafic fault materials as suggested by surface observations [*Moore and Rymer*, 2012] and the tectonic history [*Wang et al.*, 2013; *Pikser et al.*, 2012] on the depth extent of frictional behavior.

8. Conclusions

For a model in which effective stress is determined by fractional contact area and controlled by contact-scale yielding, effective stress depends on temperature and shear strain rate. The resulting effective pressure coefficient α_f is near 1 when temperature is low or when the contact strain rate is high, as when shear is localized. When this model is applied to natural stresses and temperatures, α_f decreases with depth in the crust. In cases of low temperature or high strain rate, high strength mechanisms such as dislocation glide and subcritical crack growth determine the contact-scale stresses. At the transition to a weaker contact scale deformation mechanism such as dislocation creep, α_f tends rapidly towards zero with increasing temperature. For hydrostatic pore pressure and a brittle quartz shear zone with thickness of 1 mm in a vertical strike-slip faulting environment, the model BDT is at 13 km. Throughout the brittle portion of the crust above the BDT α_f is near 1. In the ductile regime immediately below the BDT the shear zone thickness is assumed to be 1 km and due to the strain rate dependence and the associated lower ductile contact-scale flow strength, the imposed delocalized slip requires α_f=0. For this wide shear zone, reactivating friction below the BDT requires both imposed localization and elevated pore pressure. To produce frictional slip at depths between 15 and 30 km, the depth range that hosts low frequency earthquakes on the San Andreas, requires pore pressure within 0.5 MPa of lithostatic if the shear zone is 1 mm thick. For this shear thickness friction can extend no deeper than 35 km.

Acknowledgements: There is no unpublished data in this paper. Access to the published data used in Figures 4 and 5 along with additional details of the calculations are available from the corresponding author (NMB). A number of helpful discussions of effective stress with Jim Rice, John D. Platt, Teng-fong Wong, and David Lockner are gratefully acknowledged. Teng-fong suggested the bounds used in equation (4) and the need to consider the percolation threshold. David pointed out issues with assuming non-wetted grain boundaries. Josh Taron and Ole Kaven of the USGS, and JGR referees Teng-fong Wong and Toshi Shimamoto provided reviews that significantly improved the manuscript. Thanks to the associate editor Alex Schubnel for assistance beyond the call in obtaining the journal reviews. This work was supported in part by a grant #12153 from the Southern California Earthquake Center to Brown University. SCEC is presently funded by NSF Cooperative Agreement EAR-0529922 and USGS Cooperative Agreement 07HQAG0008. The SCEC contribution number for this paper is 1971.

Appendix. Relationships for crystal plasticity

Dislocation creep follows a power law relation

$$\dot{\varepsilon} = \dot{\varepsilon}_0 \left(\frac{\sigma_\Delta}{\sigma_0} \right)^n \exp\left(-\frac{Q}{RT} \right). \tag{A1}$$

n is the stress exponent, σ_Δ is the differential stress, the difference between the greatest and least principal stresses, Q is an activation energy with units of Joules/ mol °K, and $\dot{\varepsilon}_0$ and σ_0 are arbitrary reference values of strain rate and differential stress such that $\dot{\varepsilon} = \dot{\varepsilon}_0$ when $\sigma = \sigma_0$. Flow law parameters used in the various calculations are shown in **Figures 1, 5**, **6, and 7** are listed in **Table 1**.

Table 1.

Reference	N	Q (kJ/mol)	$\dot{\varepsilon}_0/\sigma_0{}^n$ (MPa^{-n})
Evans (1984) (dry)	3	430	4.e3
Hirth et al. (2001) (wet)	4	135	1e-9

For low temperature plasticity, differential stress depends on the logarithm of the strain rate [e.g., *Evans and Goetze*, 1979]. The low temperature plasticity flow law of *Mei et al.* [2010] is

$$\dot{\varepsilon} = \dot{\varepsilon}_0 \left(\frac{\sigma_\Delta}{\sigma_0} \right)^2 \exp\left(\frac{-Q}{RT} \left[1 - \sqrt{\frac{\sigma_\Delta}{\sigma_p}} \right] \right), \tag{A2}$$

where R is the gas constant, T is temperature in °K, σ_p is the Peierls stress which is the yield strength at absolute zero and Q is activation energy at zero stress. The flow law parameters used in the various calculations that are shown in **Figures 1**, **5**, **6** and **7** are listed in **Table 2**.

Table 2.

Reference	Q (kJ/mol)	$\dot{\varepsilon}_0/\sigma_0{}^2$ (1/MPa^{2}s)	σ_p (MPa)
Evans (1984) (dry)	320	6.4e-5	15000
Estimated properties (wet)	320	2.6e-4	7500

References

Audet, P., M. G. Bostock, N. I. Christensen, and S. M. Peacock (2009), Seismic evidence for overpressured subducted oceanic crust and megathrust fault sealing, Nature, 457, 76–78, doi:10.1038/nature07650.

Axen, G.J., J. Selverstone and T. Wawrzyniec, (2001), High-temperature embrittlement of extensional Alpine mylonite zones in the midcrustal ductile-brittle transition, *J. Geophys. Res.*, 106, 4337-4348.

Bartlow, N.M., S. Miyazaki, A.M. Bradley, and P. Segall, (2011), Space-time correlation of slip and tremor duing the 2009 Cascadia slow slip event, *Geophys. Res. Lett.*, 38, doi:10.1029/2011GL048714

Becken, M., O. Ritter, P.A. Bedrosian, and U. Weckmann, U., (2011), Correlation between deep fluids, tremors and creep along the central San Andreas Fault. *Nature 480*, 87–90.

Beeler, N. M., Tullis, T. E., Blanpied, M. L., and J. D. Weeks, (1996), Frictional behavior of large displacement experimental faults, *J. Geophys. Res.*, 101, 8697-8715.

Beeler, N. M., A. Thomas, R. Bürgmann, and D. Shelly (2013), Inferring fault rheology from low-frequency earthquakes on the San Andreas, *J. Geophys. Res*., 118, 5976–5990, doi:10.1002/2013JB010118.

Beeler, N. M., and S. H. Hickman (2015), Direct measurement of asperity contact growth in quartz at hydrothermal conditions, *J. Geophys. Res.*, 120, doi: 10.1002/2014JB011816.

Beeler, N.M., and T.E. Tullis, (1997), The roles of displacement in velocity dependent volumetric strain of fault zones, *J. Geophys. Res*., 102, 22,595-22, 609.

Bishop, A.W., and A.E. Skinner, (1977), The influence of high pore-water pressure on the strength of cohesionless soils, *Phil. Trans. Royal Soc. London, A*, 284, 91-130.

Blanpied, M. L., D. A. Lockner, and J. D. Byerlee, Frictional slip of granite at hydrothermal conditions, *J. Geophys. Res*., *100*, 13,045-13,064, 1995.

Boitnott, G. N., R. L. Biegel, C. H. Scholz, N. Yoshioka, and W. Wang (1992), Micromechanics of rock friction 2: Quantitative modeling of initial friction with contact theory, *J. Geophys. Res.*, *97*, 8965–8978, doi:10.1029/92JB00019.

Bowden, F. P., and D. Tabor (1950), *The Friction and Lubrication of Solids*, 374 pp., Oxford Univ. Press, New York.

Brace, W. F., B. W. Paulding, and C. H. Scholz (1966), Dilatancy in the fracture of crystalline rock, *J. Geophys. Res.*, 71, 3939 – 3953.

Brace, W.F. and Martin, R.J. (1968). A test of the effective stress law for crystalline rocks of low porosity. *Int. J. Rock Mech. Min. Sci*. 5, 415-426.

Brace, W. F., and D. L. Kohlstedt (1980), Limits on lithospheric stress imposed by laboratory experiments, *J. Geophys. Res*., 85(B11), 6248–6252, doi:10.1029/JB085iB11p06248.

Burgmann, R. and G. Dresen, (2008), Rheology of the lower crust and upper mantle: Evidence from rock mechanics, geodesy and field observations, *Annu. Rev. Earth Planet. Sci*, *36*, 531-567.

Burgette, R.J., R.J. Weldon II, and D.A. Schmidt (2009). Interseismic uplift rates for western Oregon and along strike variation in locking on the Cascadia subduction zone, *J. Geophys. Res*. **114**, doi:10.1029/2008JB005679.

Byerlee, J. D., (1967), Frictional characteristics of granite under high confining pressure, *J. Geophys. Res.,*72, 3639-3648.

Chen, W.-P., and P. Molnar, (1983), Focal depths of intracontinental and intraplate earthquakes and their implications for the thermal and mechanical properties of the lithosphere, J. Geophys. Res., 88(B5), 4183–4214, doi:10.1029/JB088iB05p04183.

Cheng, A. H.-D., (1997), Material coefficients of anisotropic poroelasticity, *Int. J. Rock Mech. Min. Sci.*, 34, 199-205.

Chernak, L., G. Hirth, J. Selverstone, and J. Tullis, (2009), The effect of aqueous and carbonic fluids on the dislocation creep strength of quartz, *J. Geophys. Res.*, 114, B04201, doi:10.1029/2008JB005884.

Chester, F. M., A rheologic model for wet crust applied to strike-slip faults, *J. Geophys. Res.*, *100*, 13,033-13,044, 1995.

Chester, F. M. and J. S. Chester, (1998), Ultracataclasite structure and friction processes of the San Andreas fault. *Tectonophysics*, *295*, 199-221.

Dieterich, J. H., and B. D. Kilgore (1994), Direct observation of frictional contacts: New insights for state-dependent properties, *Pure and Applied Geophys*., *143*, 283–302.

Dieterich, J.H. and B.D. Kilgore, (1996), Imaging surface contacts: power law contact distributions and contact stresses in quartz, calcite, glass and acrylic plastic. Tectonophysics, 256, 219–239.

Dragert, H., K. Wang, and T. S. James (2001). A silent slip event on the deeper Cascadia subduction interface, *Science* **292** 1525–1528.

Escartin, J., G. Hirth, and B. Evans, (1997), Nondilatant brittle deformation of serpentinites; implications for Mohr-Coulomb theory and the strength of faults, *J. Geophys. Res.*, *102*, 2897-2913.

Escartin, J., M. Andreani, G. Hirth, and B. Evans, (2008). Relationships between the microstructural evolution and the rheology of talc at elevated pressures and temperatures, Earth and Planetary Science Letters 268 (2008) 463–475

Evans, B., (1984), The effect of temperature and impurity content on indentation hardness of quartz, *J. Geophys. Res.*, *89*, 4213-4222.

Evans, B. and C. Goetze, (1979), The temperature variation of hardness of olivine and its implication for polycrystalline yield stress, *J. Geophys. Res.*, *84*, 5505-5524.

Evans, B., J.T. Fredrich, and T.-f. Wong, (1990), The brittle-ductile transition in rocks: recent experimental and theoretical progress, in *The Heard volume, Geophys. Mono. Ser. 56*, pps 1-20, A.G. Duba et al. eds, Am. Geophys. Un., Washington, D.C.

Geertzma, J., (1957), The effect of fluid pressure decline on volumetric changes of porous rocks, Petroleum Transactions of the AIME, 210, 331-340.

Ghosh, A., J. E. Vidale, Z. Peng, K. C. Creager, and H. Houston (2009), Complex nonvolcanic tremor near Parkfield, California, triggered by the great 2004 Sumatra earthquake, *J. Geophys. Res., 114*, doi:10.1029/2008JB006062.

Goetze, C., and B. Evans (1979), Stress and temperature in the bending lithosphere as constrained by experimental rock mechanics, *Geophys. J. R. Astron. Soc., 59*, 463 – 478.

Gomberg, J., J. L. Rubinstein, Z. Peng, K. C. Creager, J. E. Vidale, and P. Bodin (2008), Widespread triggering of nonvolcanic tremor in California, *Science, 319*, 173, doi:10.1126/science.1149164.

Greenwood, J. A., and J. Williamson, (1966), Contact of nominally flat surfaces, *Proc. R. Soc. London, Ser. A, 295*, 300 – 319.

Gurevich, B., (2004), A simple derivation of the effective stress coefficient for seismic velocities in porous rocks, *Geophysics*, *69*, 393-397.

Handin, J., R.V. Hager, M. Friedman, and J.N. Feathers, (1963), Experimental deformation of sedimentary rocks under confining pressure: pore pressure tests, Bull. Am. Assoc. Petroleum Geologists, 5, 716-755.

Handy, M. R., G. Hirth, and R. Bürgmann (2007), Continental fault structure and rheology from the frictional-to-viscous transition downward, in *Tectonic Faults: Agents of Change on a Dynamic Earth*, edited by M. R. Handy, et al., pp. 139-181, MIT Press, Cambridge, MA.

Hawthorne, J. C., and A. M. Rubin (2010), Tidal modulation of slow slip in Cascadia, *J. Geophys. Res., 115*, doi:10.1029/2010JB007502.

Hawthorne, J. C., and A. M. Rubin (2013), Tidal modulation and back-propagating fronts in slow slip events simulated with a velocity weakening to -strengthening friction law, *J. Geophys. Res., 118*, 1216–1239, doi:10.1002/jgrb.50107.

Heard, H.C., and N.L. Carter, (1968), Experimentally induced natural intergranular flow in quartz and quartzite, *Am. J. Sci*., *266,* 1-42.

Hill, D. P. (2010), Surface-wave potential for triggering tectonic (nonvolcanic) tremor, *Bull. Seismol. Soc. Am., 100*, 1859–1878, doi:10.1785/ 0120090362.

Hirth, G. and D.L. Kohlstedt, (1995), Experimental constraints on the dynamics of the partially molten upper mantle: Deformation in the diffusion creep regime, *J. Geophy. Res., 100*, 1981–2001, doi: 10.1029/94JB02128.

Hirth, G., C. Teyssier, and W. J. Dunlap, (2001), An evaluation of quartzite flow laws based on comparisons between experimentally and naturally deformed rocks, *Int. J. Earth Sci., 90*, 77-87.

Hirth, G. and N. M. Beeler, (2015), The role of fluid pressure on frictional behavior at the base of the seismogenic zone, *Geology*, *43*, 223-226.

Hubbert, M.K., and W.W. Rubey, (1959), Role of fluid pressure in mechanics of overthrust faulting, *Bull. Geol. Soc. Am*., *70*, 115-160.

Hubbert, M.K., and W.W. Rubey, (1960), Role of fluid pressure in mechanics of overthrust faulting- A reply, *Bull. Geol. Soc. Am*., *71*, 617-628.

Johnson, K. (1987), Contact Mechanics (paperback), 452 pp, Cambridge University, New York.

Karato, S., (2012), Deformation of earth materials: An introduction to the rheology of solid Earth, Cambridge University Press

King, G.C.P. and S. Wesnousky, (2007), Scaling of fault parameters for continental strike-slip earthquakes, Bull. Seismol. Soc. Am., 97, 1833-1840.

Kirby, S. H. (1980), Tectonic stresses in the lithosphere: Constraints provided by the experimental deformation of rocks, J. Geophys. Res., 85, 6353 – 6363.

Lachenbruch, A. H., and J. H. Sass, (1973), Thermo-mechanical aspects of the San Andreas fault system in *Proceedings of the Conference on the Tectonic Problems of the San Andreas Fault System*, eds R.L. Kovach and A. Nur, pp. 192-205, Stanford University Press, Stanford, Calif..

Lambert, A., H. Kao, G. Rogers, and N. Courtier (2009), Correlation of tremor activity with tidal stress in the northern Cascadia subduction zone, *J. Geophys. Res., 114*, B00A08, doi:10.1029/2008JB006038.

Liu, Y., and Rice, J. R. (2005), Aseismic slip transients emerge spontaneously in 3D rate and state modeling of subduction earthquake sequences, *J. Geophys. Res*., *110*, B08307, doi:10.1029/2004JB003424.

Mandl, G., (1988), *Mechanics of tectonic faulting*, Elsevier, Amsterdam, 407p

Mandl, G., (2000), *Faulting in brittle rocks- An Introduction to the Mechanics of tectonic faults*, Springer Verlag, Berlin, 434p

Marone, C., C. B. Raleigh, and C. H. Scholz (1990), Frictional behavior and constitutive modeling of simulated fault gouge, *J. Geophys. Res., 95*, 7007 – 7025.

Masuda, T., T. Hiraga, H. Ikei, H. Kanda, Y. Kugimiya, and M. Akizuki, Plastic deformation of quartz at room temperature: a Vickers nano-indentation test, *Geophys. Res. Lett.*, 27, 2773-2776, 2000.

McCaffrey, R., A. I. Qamar, R. W. King, R. Wells, G. Khazaradze, C. A. Williams, C. W. Stevens, J. J. Vollick, and P. C. Zwick (2007). Fault locking, block rotation and crustal deformation in the Pacific Northwest, *Geophys. J. Int*. **169** 1315–1340, doi:10.1111/j.1365-246X.2007.03371.

Mei, S., Suzuki, A. M., Kohlstedt, D. L., Dixon, N. A., and Durham, W. B., (2010). Experimental constraints on the strength of the lithospheric mantle, J. Geophys. Res., 115, B08204, doi:10.1029/2009JB006873

Miyazawa, M., and E. E. Brodsky (2008), Deep low-frequency tremor that correlates with passing surface waves, *J. Geophys. Res., 113*, B01307, doi:10.1029/2006JB004890.

MOORE, D.E., and RYMER, M.J. (2012), Correlation of clayey gouge in a surface exposure of serpentinite in the San Andreas Fault with gouge from the San Andreas Fault Observatory at Depth (SAFOD). J. of Struct. Geol., 38, 51-60, doi:10.1016/j.jsg.2011. 11.014

Morrow, C., and J. Byerlee (1989), Experimental studies of compaction and dilatancy during frictional sliding on faults containing gouge, *J. Struct. Geol., 11*, 815 – 825.

Morrow, C., B. Radney and J. Byerlee, (1992), Frictional strength and the effective pressure law of montmorillonite and illite clays, in Fault mechanics and transport properties of rocks, ed Evans, Wong, , p 69-88.

Nadeau, R. M., and A. Guilhem (2009), Nonvolcanic tremor evolution and the San Simeon and Parkfield, *Science, 325*, 191–194, doi:10.1126/ science.1174155.

Nakata, R., N. Suda, and H. Tsuruoka (2008), Non-volcanic tremor resulting from the combined effect of Earth tides and slow slip events, *Nat. Geosci., 1*, 676–678, doi:10.1038/ngeo288.

Nur, A., and J.D. Byerlee, (1971), An exact effective stress law for elastic deformation of rock with fluids, J. Geophys. Res., 76, 6414-6419.

Obara, K. (2002), Nonvolcanic deep tremor associated with subduction in southwest Japan, *Science, 296*, 1679–1681, doi:10.1126/science.1070378.

Peacock, S. M. (2009), Thermal and metamorphic environment of subduction zone episodic tremor and slip, *J. Geophys. Res*., 114, B00A07, doi:10.1029/2008JB005978.

Peacock S. M., N. I. Christensen, M. G. Bostock, and P. Audet (2011), High pore pressures and porosity at 35 km depth in the Cascadia subduction zone, *Geology*, 39(5), 471–474, doi:10.1130/G31649.1.

Peng, Z., J. E. Vidale, A. G. Wech, R. M. Nadeau, and K. C. Creager (2009), Remote triggering of tremor along the San Andreas Fault in central California, *J. Geophys. Res., 11*4, B00A06,doi:10.1029/ 2008JB006049.

J. D. Platt, J. W. Rudnicki, and J. R. Rice (2014), Stability and Localization of Rapid Shear in Fluid-Saturated Fault Gouge, 2. Localized zone width and strength evolution, Journal of Geophysical Research, 119, 4334-4359.

Proctor, B., and G. Hirth, (2015), Role of pore fluid pressure on transient strength changes and fabric development during serpentine dehydration at mantle wedge conditions, Earth and Planetary Science Letters, 421, 1–12.

Rice, J.R. (1992), Fault Stress States, Pore Pressure Distributions, and the Weakness of the San Andreas Fault, in Fault Mechanics and Transport Properties of Rocks, ed B. Evans and T.-f. Wong, Academic Press, 475-503.

Rice, J. R., and M. P. Cleary, (1976), Some basic stress diffusion solutions for fluid-saturated elastic porous media with compressible constituents*, Rev. Geophys.*, *14,* 227-241.

Rice, J.R., J. W. Rudnicki, and J.D. Platt, (2014), Stability and Localization of Rapid Shear in Fluid-Saturated Fault Gouge, 1. Linearized stability analysis, J. Geophys. Res. Solid Earth, 119, 4311–4333.

Robin, Y. P., (1973), Note on effective pressure, J. Geophys. Res., 78(14), 2434–2437, doi:10.1029/JB078i014p02434.

Roeloffs, E. A., P.G. Silver, W.A. McCausland, (2009). Transient Strain During and Between Northern Cascadia Episodic Tremor and Slip Events From Plate Boundary Observatory Borehole Strainmeters, *Eos Trans. AGU* **90**(22) Joint. Assem. Suppl., Abstract G12A-02.

Roeloffs, E. A. and W.A. McCausland, (2010), Constraints on aseaismic slip during and between northern Cascadia episodic tremor and slip events from PBO borehole strainmeters, *Seismol. Res. Lett.* **81** 337.

Schmalzle, G. M., R. McCaffrey, and K. C. Creager (2014), Central Cascadia subduction zone creep, Geochem. Geophys. Geosyst., 15, 1515–1532, doi:10.1002/2013GC005172.

Scholz, C.H., (1990), *The Mechanics of Earthquakes and Faulting*, Cambridge: Cambridge Univ. Press.

Scholz, C. H., and J. T. Engelder (1976), The role of asperity indentation and ploughing in rock friction, I, Asperity creep and stick-slip, *Int. J. Rock Mech. Sci. Geomech. Abstr., 13*, 149 – 154.

Segall, P., and Bradley, A. M., (2012), Slow-slip evolves into megathrust earthquakes in 2D numerical simulations, *Geophys. Res. Lett.*, 39, doi:10.1029/2012GL052811.

Segall, P., A.M. Rubin, A.M., Bradley, and J.R. Rice, (2010), Dilatant strengthening as a mechanism for slow slip events, *J. Geophys. Res*., 115, B12305, doi:10.1029/2010JB007449.

Shelly, D. R., and J. L. Hardebeck (2010), Precise tremor source locations and amplitude variations along the lower-crustal central San Andreas Fault, *Geophys. Res. Lett., 37*, L14301, doi:10.1029/2010GL043672.

Shelly, D. R., G. C. Beroza, S. Ide, and S. Nakamula (2006), Low frequency earthquakes in Shikoku, Japan, and their relationship to episodic tremor and slip, *Nature, 442*, 188–191, doi:10.1038/nature04931.

Shelly, D. R., Z. Peng, D. P. Hill, and C. Aiken (2011), Triggered creep as a possible mechanism for delayed dynamic triggering of tremor and earthquakes, *Nature. Geosci., 4*, 384–388, doi:10.1038/ngeo1141.

Shelly, D. R., and K. M. Johnson (2011), Tremor reveals stress shadowing, deep postseismic creep, and depth-dependent slip recurrence on the lower-crustal San Andreas fault near Parkfield, *Geophys. Res. Lett., 381,* L13312, doi:10.1029/2011GL047863.

Sibson, R.H., (1996), Structural permeability of fluid-driven fault-fracture meshes, J. Struct. Geol., 18 1031-1042.

Skempton, A.W., (1960), Effective stress in soils, concrete and rocks, *Conf. on Pore Pressure and Suction in Soils*, London, Butterworths.

Sleep, N. H. (1997), Application of a unified rate and state friction theory to the mechanics of fault zones with strain localization, *J. Geophys. Res*., *102*(B2), 2875–2895.

Sleep, N.H. (2006), Frictional dilatancy, *Geochemistry Geophysics Geosystems*, 7, Q10008, doi:10.1029/2006GC001374

Sleep, N. H., E. Richardson, and C. Marone (2000), Physics of strain localization in synthetic fault gouge, J. Geophys. Res., 105(B11), 25,875–25,890.

Stesky, R. M. (1978), Mechanisms of high temperature frictional sliding in Westerly granite, Can. J. Earth Sci., 15, 361 – 375.

Terzaghi, K., (1936), The shearing resistance of saturated soils, Proc. 1st Int. Conf. Soil Mech., 1, 54-56.

Terzaghi, K., (1943), *Theoretical soil mechanics*, 503pps, John Wiley & Sons, Inc., New York.

Thomas, A. M., R. M. Nadeau, and R. Burgmann (2009), Tremor-tide correlations and near-lithostatic pore pressure on the deep San Andreas fault, *Nature, 462*, 2009| doi:10.1038/nature08654.

Thomas, A.M., R. Burgmann, D.R. Shelly, N.M. Beeler and M.L. Rudolph, (2012), Tidal sensitivity of low frequency earthquakes near Parkfield, CA: implications for fault mechanics within the brittle-ductile transition, *J. Geophys. Res., 117*, B05301, doi:10.1029/2011JB009036.

Wang, K., I. Wada, and J. He (2011). Thermal and Petrologic Environments of ETS, *Eos Trans. AGU* **90**(52) Fall Meet. Suppl., Abstract T22B-01.

Wang, K. and J. He (1994), Mechanics of low-stress forearcs: Nankai and Cascadia, *J. Geophys. Res.*, 104, 15191-15205.

Wang, Y., D.W. Forsyth, C.J. Rau, N. Carriero, B. Schmandt, J.B. Gaherty and B. Savage, (2013), Fossil slabs attached to unsubducted fragments of the Farallon plate, Proc Natl Acad Sci., 110, 5342–5346.

Wark, D.A. and E.B. Watson, (1998), Grain-scale permeabilities of texturally-equilibrated monomineralic rocks. Earth Planet. Sci. Lett. 164, 591-605.

Watson, E.B., and J.M. Brennan, (1987), Fluids in the lithosphere, 1. Experimentally-determined wetting characteristics of CO2-H20 fluids and their implications for fluid transport, host-rock physical properties, and fluid inclusion formation, *Earth and Planetary Science Letters*, *85*, 497-515.

Wech, A. G., and K. C. Creager, (2008), Automated detection and location of Cascadia tremor, *Geophys. Res. Lett.*, *35*, L20302, doi:10.1029/2008GL035458.

Worthington, C., T. E. Tullis and N. M. Beeler, (1997), Stress dilatancy-relationships during frictional sliding, *Eos Trans. AGU,* 78, F475.

Yoshioka, S., K. Wang, and S. Mazzotti (2005). Interseismic locking of the plate interface in the northern Cascadia subduction zone, inferred from inversion of GPS data, *Earth and Planetary Science Letters* **231** 239–247.

Zhu, W., C. David, and T.-f. Wong, (1995), Network modeling of permeability evolution during cementation and hot isostatic pressing, *J. Geophys. Res.*, *100*, 15451-15464.

983 **Table 1.** Symbols in order of appearance

symbol	Definition	1st appearance
σ_n^e	effective normal stress	(1a)
σ_n	applied normal stress	(1a)
p	pore pressure	(1a)
τ	applied shear stress	text section 1
μ	friction coefficient	text section 1
V_p/V_s	ratio of p to s wave speed	text section 1
σ^e	effective stress (general)	(1b)
σ	applied stress (general)	(1b)
α	effective pressure coefficient (general)	(1b)
α_f	effective pressure coefficient for friction	text section 2
N	applied normal force	(2a)
N_c	contact scale normal force	(2a)
A	Area	(2a)
A_c	contact area	(2a)
σ_3	least principal stress	text section 3
σ_1	greatest principal stress	text section 3
σ_1^c	contact scale greatest principal stress	text section 3
σ_3^c	contact scale least principal stress	Figure 3b
σ_Δ	differential stress	text section 3
σ_y	yield stress	text section 3
σ_m^c	contact scale mean stress	Figure 3b
σ_m	mean stress	Figure 3a
σ_c	contact scale normal stress	text section 3
S	applied shear force	text section 3

S_c	contact scale shear force	text section 3
τ_c	contact scale shear stress	text section 3
ϕ	friction angle	Figure 3c
σ_1^e	effective greatest principal stress	text section 3
σ_3^e	effective least principal stress	text section 3
χ	constant specific to the stress component of interest	text section 3
δ_n	fault normal displacement	text section 4
ε_n	normal strain	text section 4
γ	shear strain	text section 4
δ_s	fault shear displacement	text section 4
$\dot{\varepsilon}_n$	normal strain rate	text section 4
V	slip velocity	text section 4
$\dot{\varepsilon}_n^c$	contact scale normal strain rate	text section 4
$\dot{\gamma}$	shear strain rate	text section 4
σ_Δ^{LTP}	differential stress from low temperature plasticity	text section 5
σ_Δ^{DC}	differential stress from dislocation creep	text section 5
$\sigma_\Delta^{friction}$	differential stress from friction	text section 6
σ_Δ^{flow}	differential stress from flow	text section 6
V_L	loading velocity, plate motion rate	text section 6
w	fault zone width	text section 6
$\dot{\varepsilon}_0$	reference strain rate	(A1)
σ_0	reference differential stress	(A1)
Q	activation energy	(A1)

R	gas constant	(A1)
T	temperature in °K	(A1)
σ_p	Peierls stress	(A2)

984

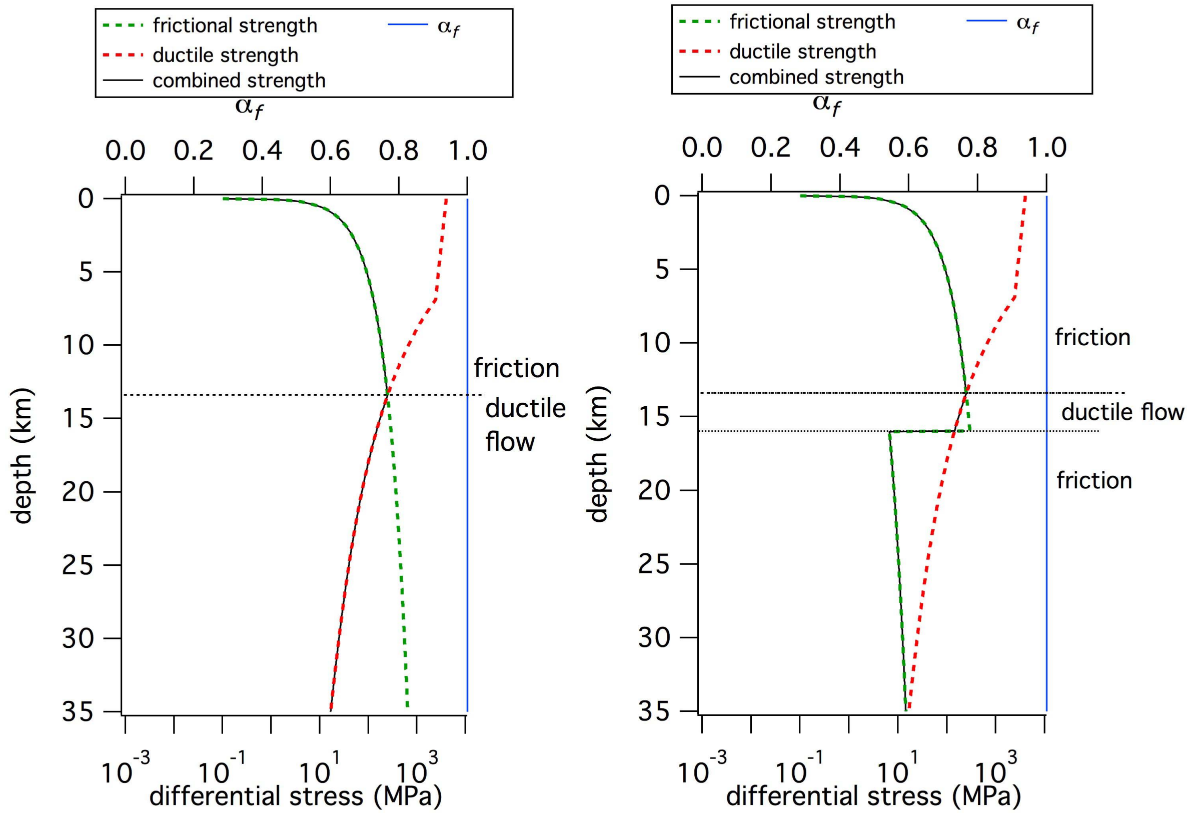

frictional strength
ductile strength
combined strength
α_f
0.0 0.2 0.4 0.6 0.8 1.0
depth (km)
0 5 10 15 20 25 30 35
friction
ductile
flow
10^{-3} 10^{-1} 10^{1} 10^{3}
differential stress (MPa)
frictional strength
ductile strength
combined strength
α_f
0.0 0.2 0.4 0.6 0.8 1.0
depth (km)
0 5 10 15 20 25 30 35
friction
ductile flow
friction
10^{-3} 10^{-1} 10^{1} 10^{3}
differential stress (MPa)

Figure 1. Crustal strength profiles**.** Differential strength (black solid) with depth from friction and creep for quartz after *Goetze and Evans* [1979] for a strain rate of 1 x 10^{-12}/s with $\sigma_e = \sigma_n - p$. The horizontal axis is plotted on a logarithmic scale to better illustrate the small deep stress levels. Overburden is 28 MPa/km, μ = 0.6, and the average of the greatest and least principal stresses is equal to the overburden. The assumed temperature gradient is from *Lachenbruch and Sass* [1973]. Friction is shown in dashed green and ductile strength in dashed red; the lower of the two (black line) corresponds to the failure strength at any given depth. The upper-crustal ductile strength at depths above ~7 km follows a relation for low temperature plasticity [*Mei et al.*, 2010] that well represents low temperature data from *Evans* [1984]. At depths below 7 km the flow strength follows the dislocation creep flow law as constrained by the laboratory data of *Hirth et al.* [2001]. The parameters used in these flow laws are listed in Tables 1 and 2 in the Appendix. The brittle-ductile transition, the intersection of frictional and flow strengths, is at ~13 km depth. Shown on the top axis is the effective pressure coefficient α_f, assumed to be depth and temperature independent. a) For hydrostatic pore pressure at all depths (10 MPa/km). b) Same as in a) except below 16 km depth where the pore pressure is 27.6 MPa/km.

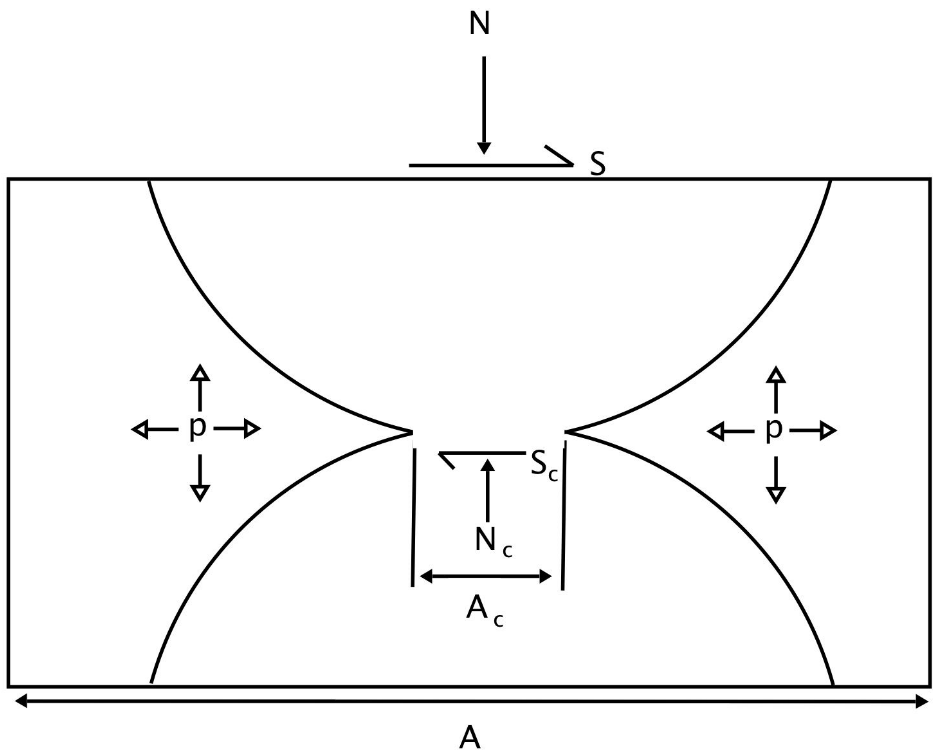
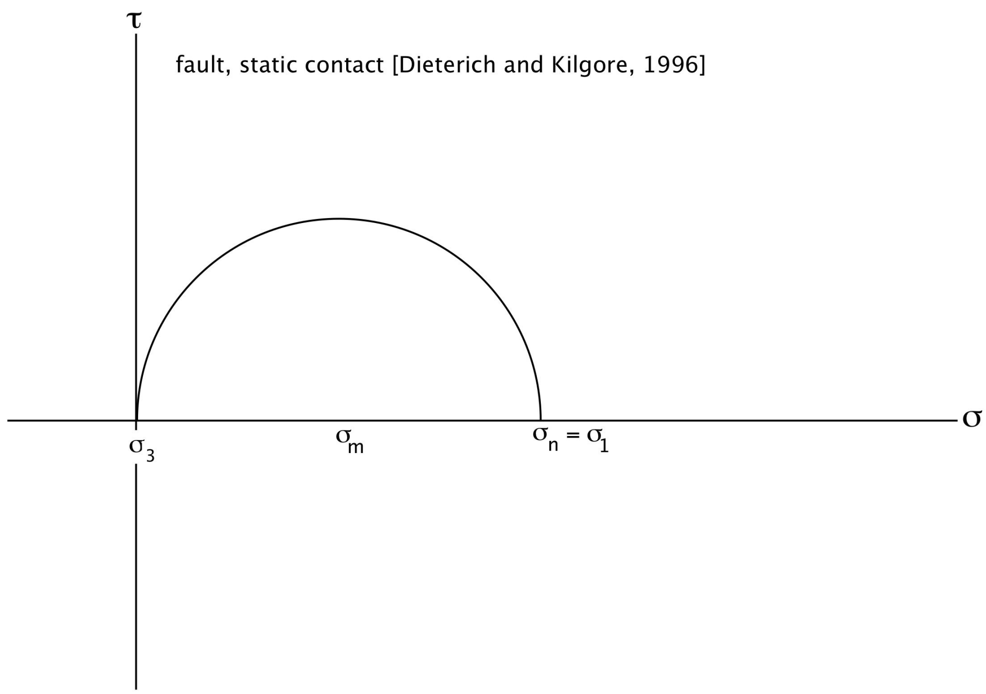
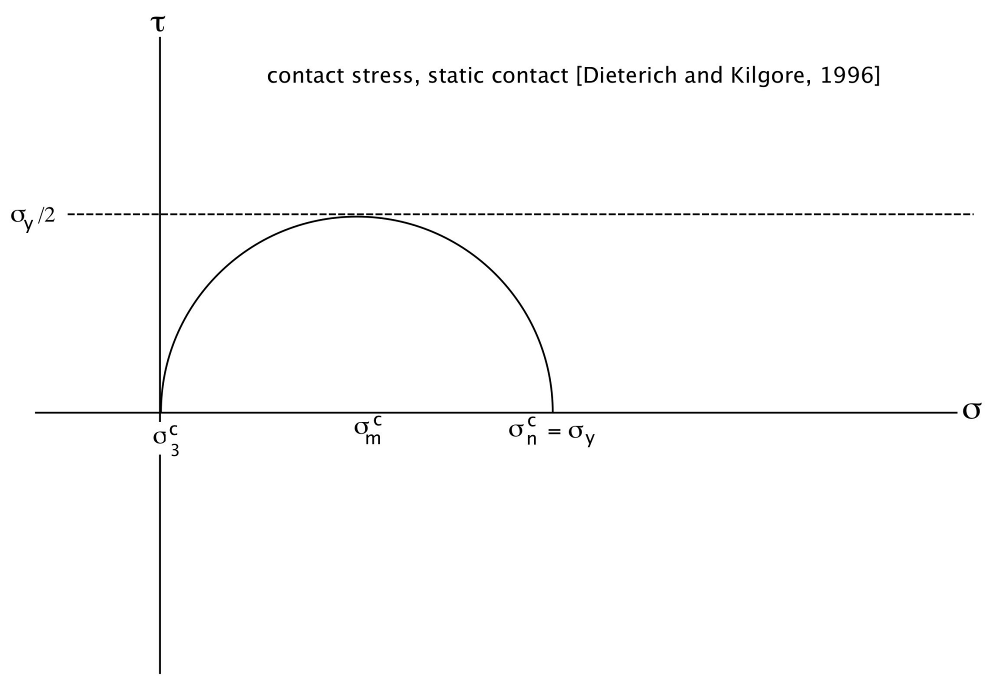


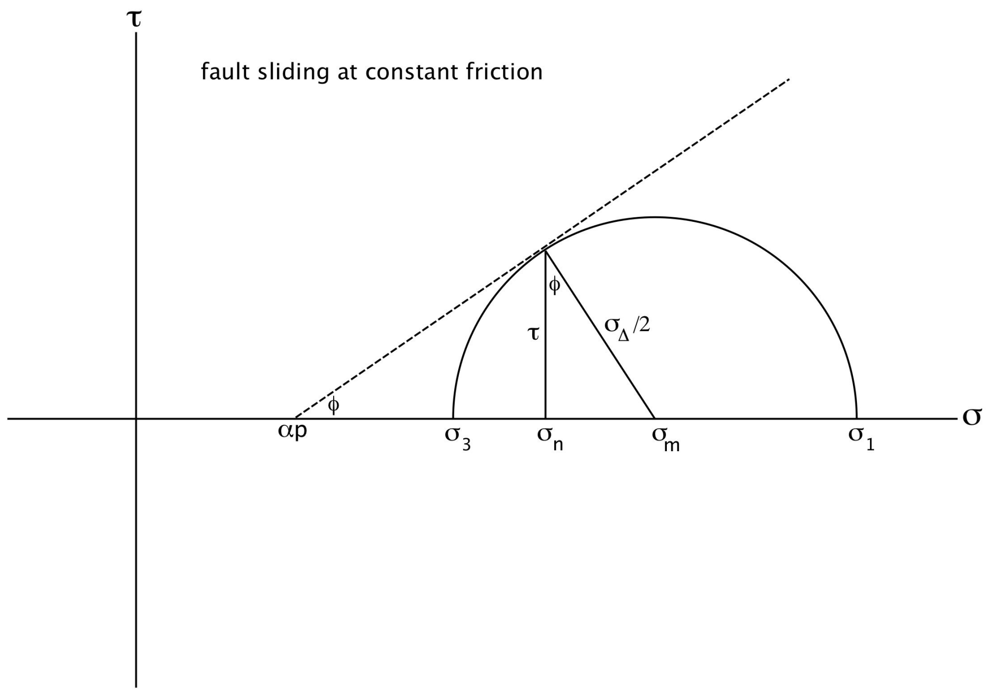
Figure 2. Schematic diagram of the force balance at a representative asperity contact area on a frictional sliding surface in the presence of pressurized fluid [after *Skempton*, 1960]. See text for discussion.


τ
fault, static contact [Dieterich and Kilgore, 1996]
σ
σ3
σm
σn = σ1

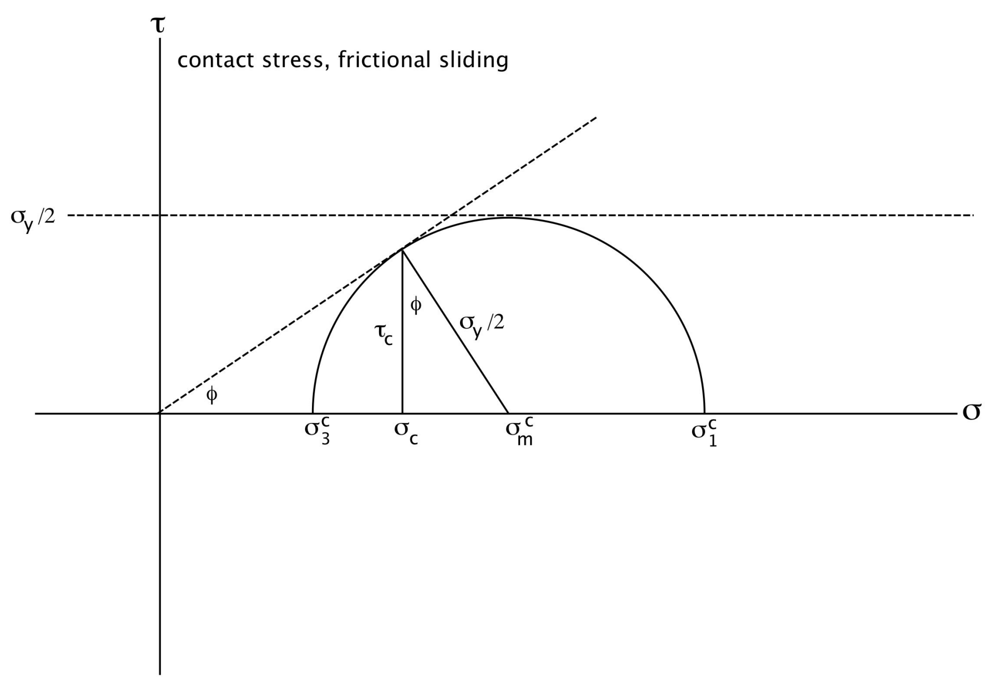
a)


τ
contact stress, static contact [Dieterich and Kilgore, 1996]
σy /2
σ
σc3
σcm
σcn = σy

b)

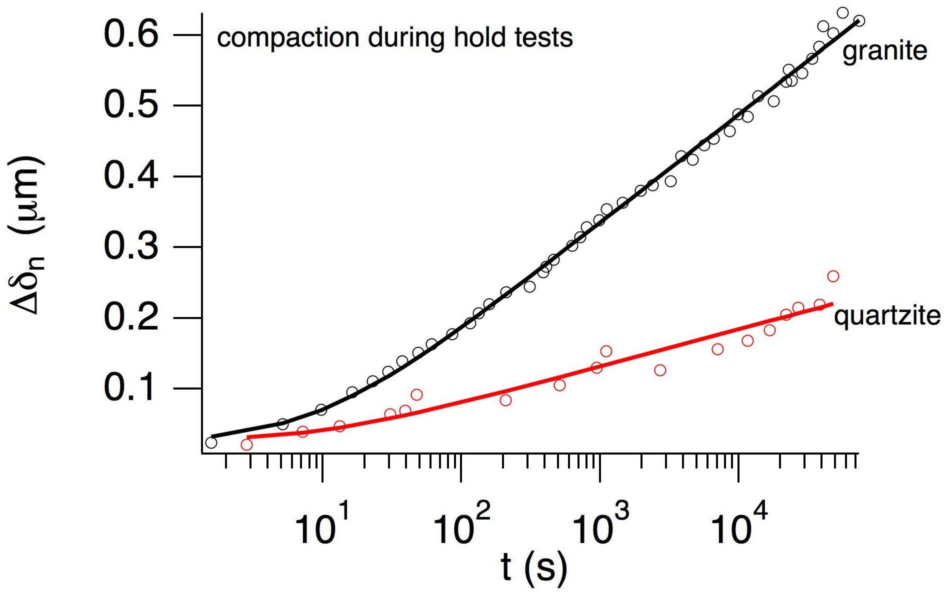

τ
fault sliding at constant friction
ϕ
τ
σΔ/2
ϕ
σ
αp
σ3
σn
σm
σ1

c)

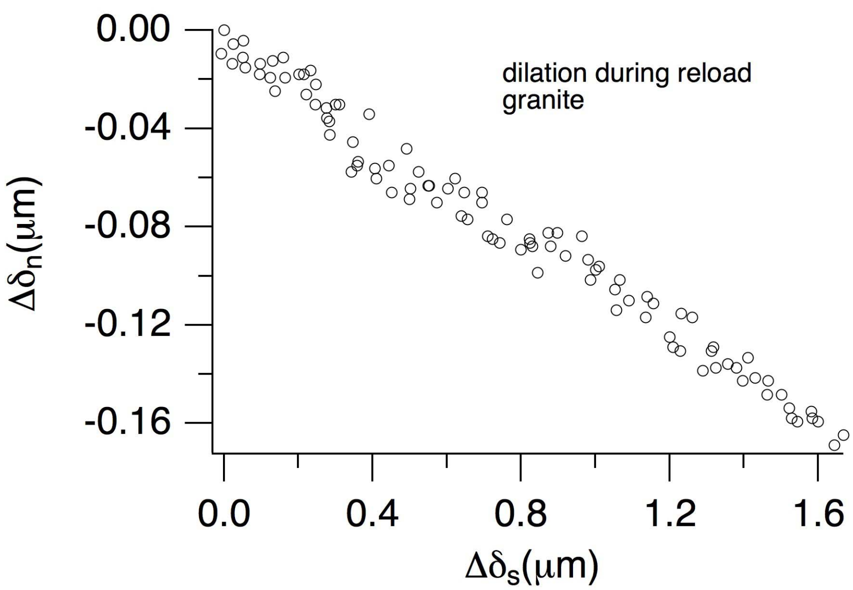

τ
contact stress, frictional sliding
σy/2
ϕ
τc
σy/2
ϕ
σ
σ3c
σc
σmc
σ1c

d)

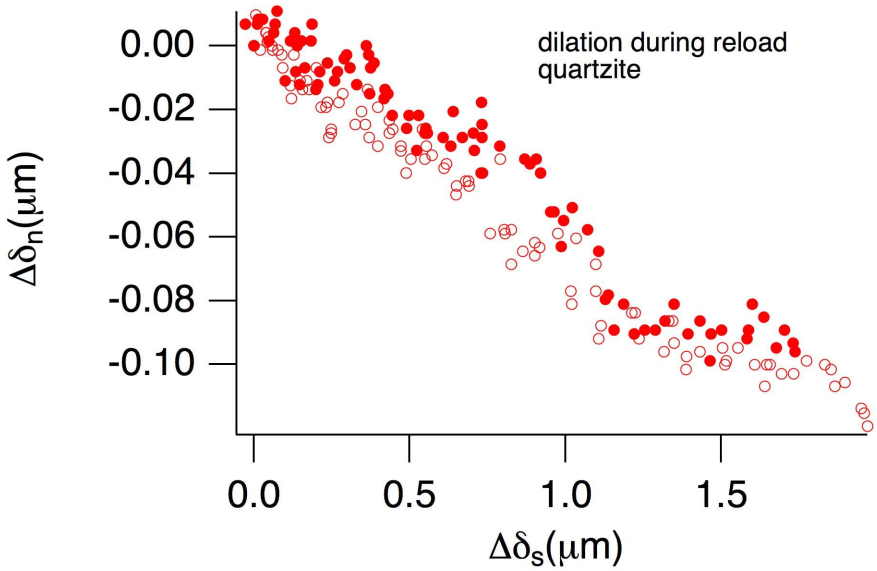
Figure 3. Mohr diagrams of stress. a) Uniaxial stress. True static stress conditions where there is no shear stress resolved on to the fault and no confining stress as in the laboratory experiments of *Dieterich and Kilgore* [1996]. b) Contact stresses for the case shown in a) assuming the contact stress is limited by yielding. c) Frictional sliding. A fault optimally oriented for slip. d) Contact stresses for the case shown in c) assuming stress is limited by yielding.


compaction during hold tests
granite
quartzite
0.6
0.5
0.4
0.3
0.2
0.1
$\Delta\delta_n$ (μm)
10^1
10^2
10^3
10^4
t (s)

a)


dilation during reload
granite
0.00
-0.04
-0.08
-0.12
-0.16
$\Delta\delta_n$(μm)
0.0
0.4
0.8
1.2
1.6
$\Delta\delta_s$(μm)

b)



c)

Figure 4. Relation between dilatancy and compaction during frictional sliding from experiments of *Worthington et al.* [1997]. Compaction corresponds to positive changes in fault normal displacement $\Delta\delta_n$. a) Data showing time dependent compaction during a hold test for bare surfaces of granite (black) and quartzite (red). b) Shear dilatancy during reloading following a hold test for bare granite at room temperature and 25 MPa normal stress. c) Shear dilatancy following two hold tests for bare quartzite at room temperature and 25 MPa normal stress.

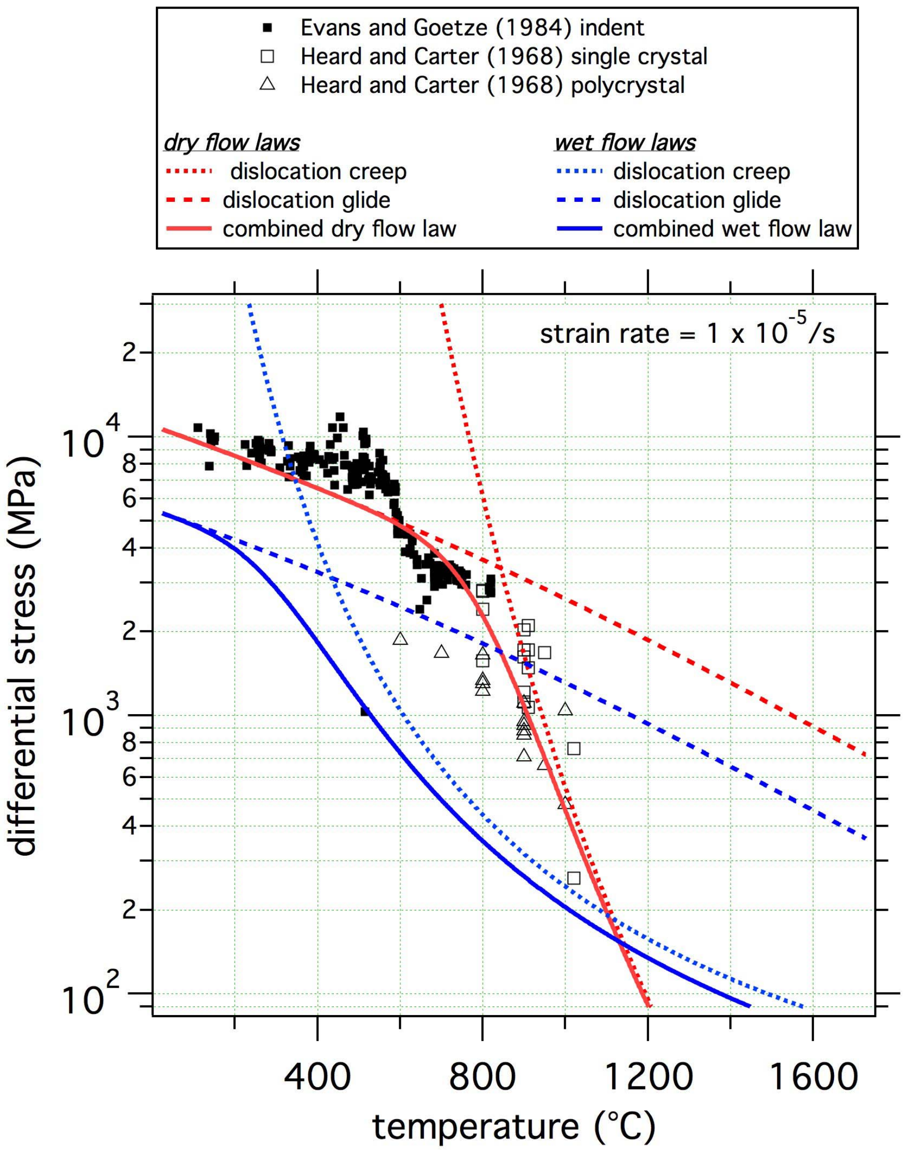
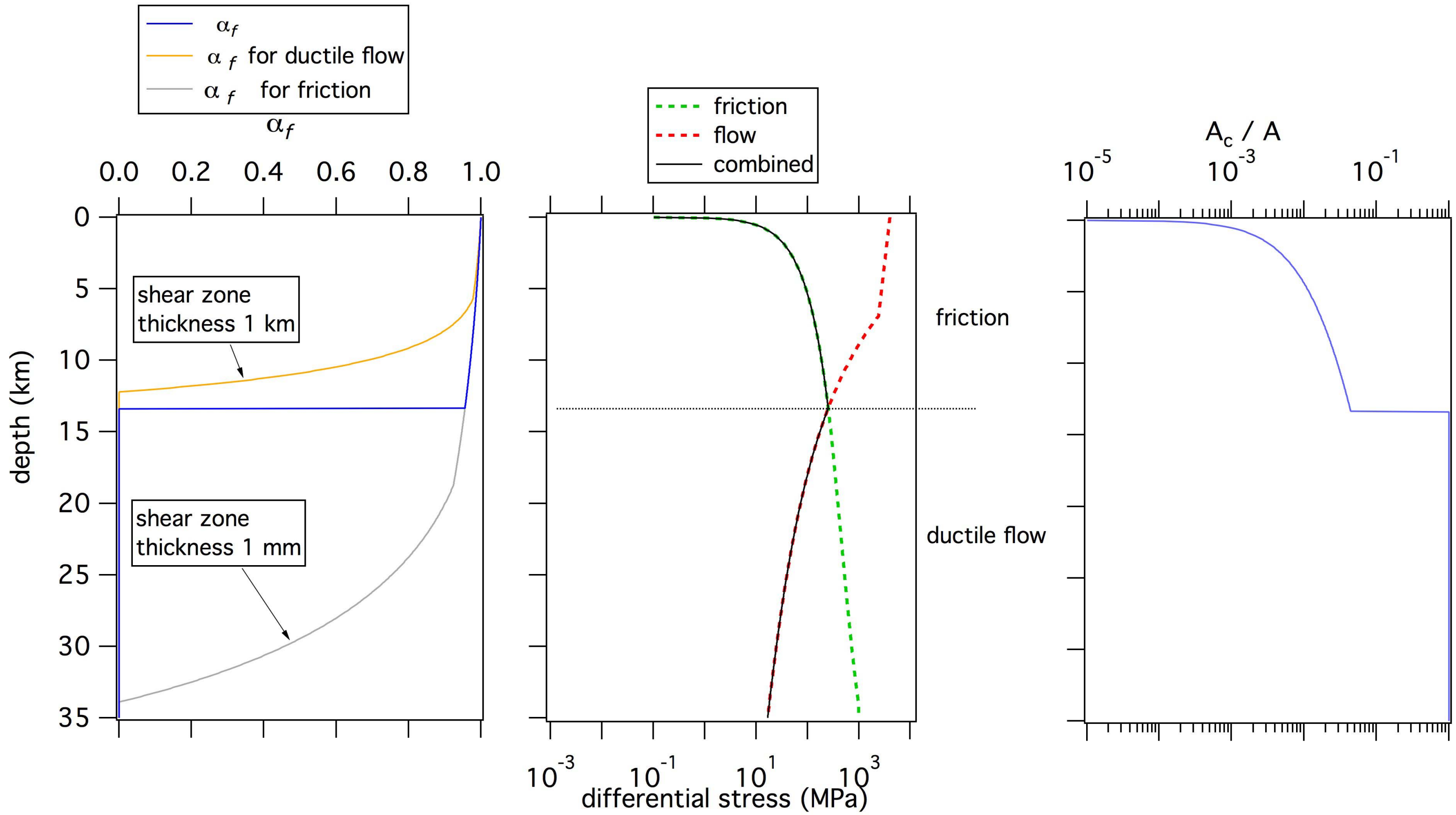


Figure 5. Laboratory data and contact scale flow laws. a) Data from *Evans* [1984] for dry indentation of quartz from room temperature to around 800°C and triaxial deformation to ~1000°C from *Heard and Carter* [1968]. Shown for reference in red are flow laws for low temperature plasticity from *Mei et al.* [2010] and dislocation glide of the standard form [*Hirth et al.*, 2003] using parameters listed in Tables 1 and 2 in the Appendix, assuming a strain rate of 1 x 10^{-5}. Also shown are the same flow laws at the same strain rate but for wet conditions (blue).



α_f
α_f for ductile flow
α_f for friction
α_f
0.0 0.2 0.4 0.6 0.8 1.0
0 5 10 15 20 25 30 35
depth (km)
shear zone thickness 1 km
shear zone thickness 1 mm
friction
flow
combined
friction
ductile flow
10^{-3} 10^{-1} 10^{1} 10^{3}
differential stress (MPa)
A_c / A
10^{-5} 10^{-3} 10^{-1}

1038

Figure 6. Center panel shows shear strength (black solid) of an optimally oriented strike-slip fault (29.5° from σ_1) using the geothermal gradient of *Lachenbruch and Sass* [1973] (~30°/km), $(\sigma_1 + \sigma_3)/2$ of 28 MPa/km, pore pressure of 10 MPa/km, $\mu = 0.6$, wet quartz yield stress for low temperature plasticity using *Mei et al.'s* [2010] flow law, *Evans'* [1984] indentation data, and dislocation creep from *Hirth et al.* [2001] at strain rate of 1 x 10^{-12}/s. Left panel shows α_f calculated from (4) (blue solid) using the same pore pressure, mean stress and flow laws at the contact scale, resulting from two possible normal strain rates (yellow and grey). Which effective pressure coefficient is used depends on which macroscopic shear resistance is lower, the brittle or ductile strength. The effective pressure coefficient associated with a 1-mm-thick shear zone and a contact normal strain rate of 1 x 10^{-7}/s is shown in grey. This is the active shear zone above the BDT. Below the BDT the shear zone is 1 km thick with a contact normal strain rate of 1 x 10^{-13}/s and an effective pressure coefficient shown in yellow. In the center panel is frictional strength shown in green and flow in red. There are almost no differences between the stresses shown here and those in the reference calculation in **Figure 1a**. Right panel is fractional contact area.

1054

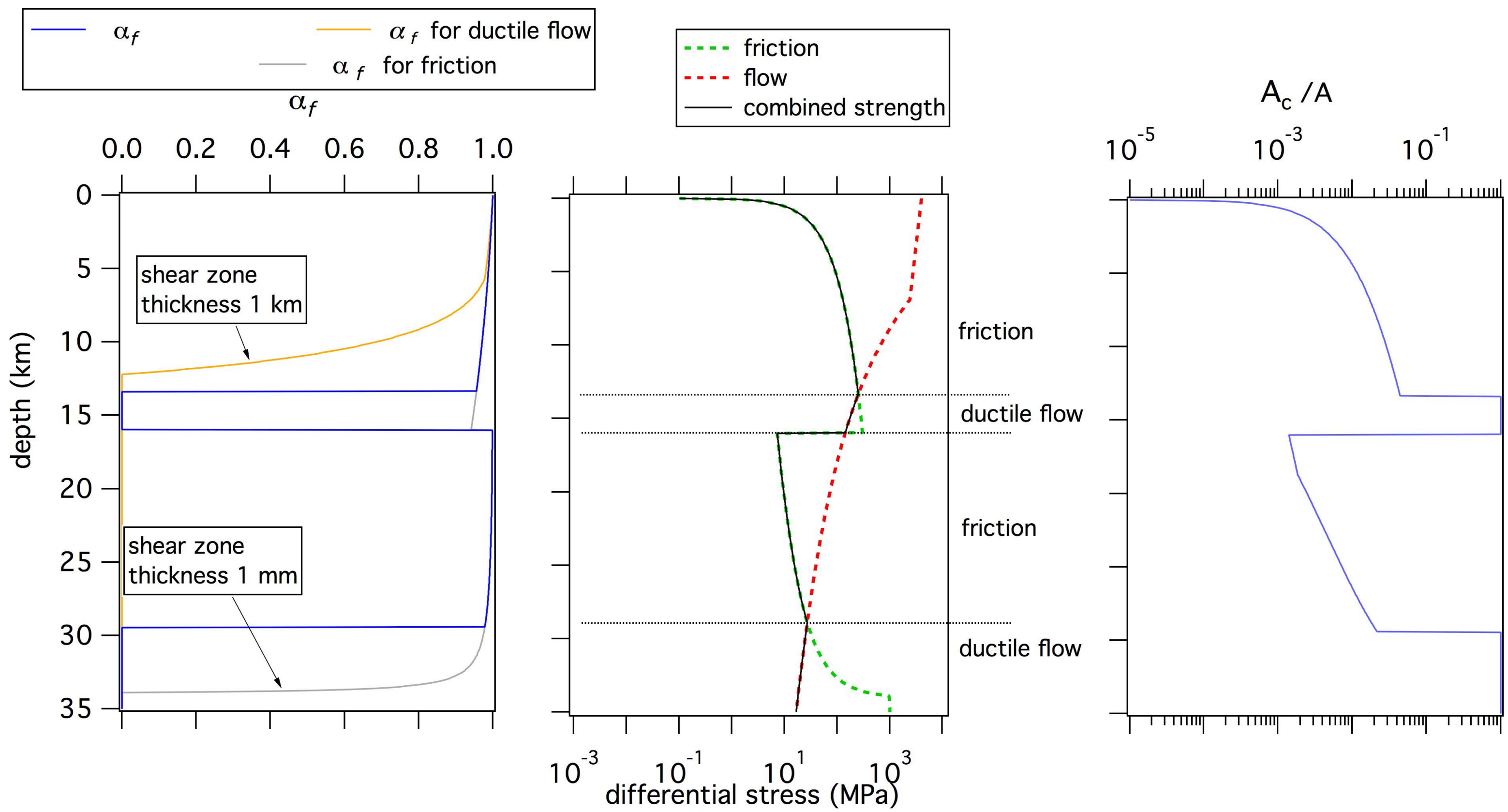
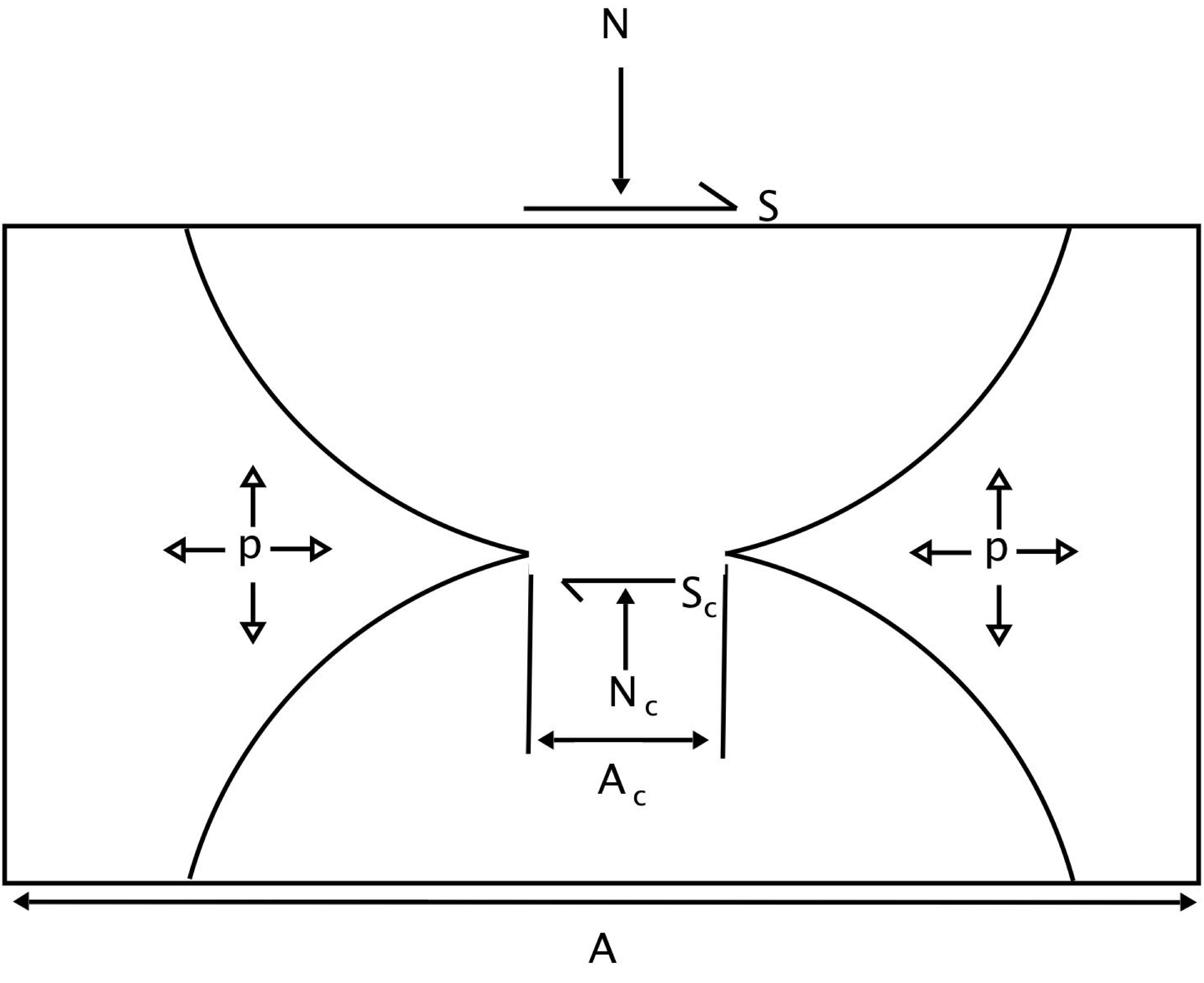

α_f
α_f for ductile flow
α_f for friction
α_f
0.0 0.2 0.4 0.6 0.8 1.0
depth (km)
0 5 10 15 20 25 30 35
shear zone thickness 1 km
shear zone thickness 1 mm
friction
flow
combined strength
friction
ductile flow
friction
ductile flow
10^{-3} 10^{-1} 10^{1} 10^{3}
differential stress (MPa)
A_c /A
10^{-5} 10^{-3} 10^{-1}

Figure 7. Calculation of the effective pressure coefficient (left panel), differential stress (center panel), and fractional contact area (right panel) using equation (4) for the same conditions as shown in Figures 1b and 6, above 16 km depth. There are three effective pressure coefficients shown. In yellow is the coefficient associated with a 1 km shear zone, and in grey is that for a 1 mm shear zone. In blue is the coefficient associated with the active thickness of the shear zone, which in this calculation varies with depth. There are 3 transitions between localized and distributed shear, the shallowest is at around 13 km. Below 16 km the pore pressure gradient is elevated to 27.6 MPa/km, within 0.4 MPa /km of lithostatic. This produces a transition back to brittle, localized deformation, a dramatic decrease in strength, and an increase in the effective pressure coefficient. Localized shear persists to nearly 30 km depth


frictional strength
ductile strength
combined strength
α_f
α_f
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
35
depth (km)
friction
ductile
ductile
flow
10^{-3}
10^{-1}
10^{1}
10^{3}
differential stress (MPa)

frictional strength
ductile strength
combined strength
α_f
α_f
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
35
depth (km)
friction
ductile flow
friction
10^{-3}
10^{-1}
10^{1}
10^{3}
differential stress (MPa)



N
S
p
p
S_c
N_c
A_c
A

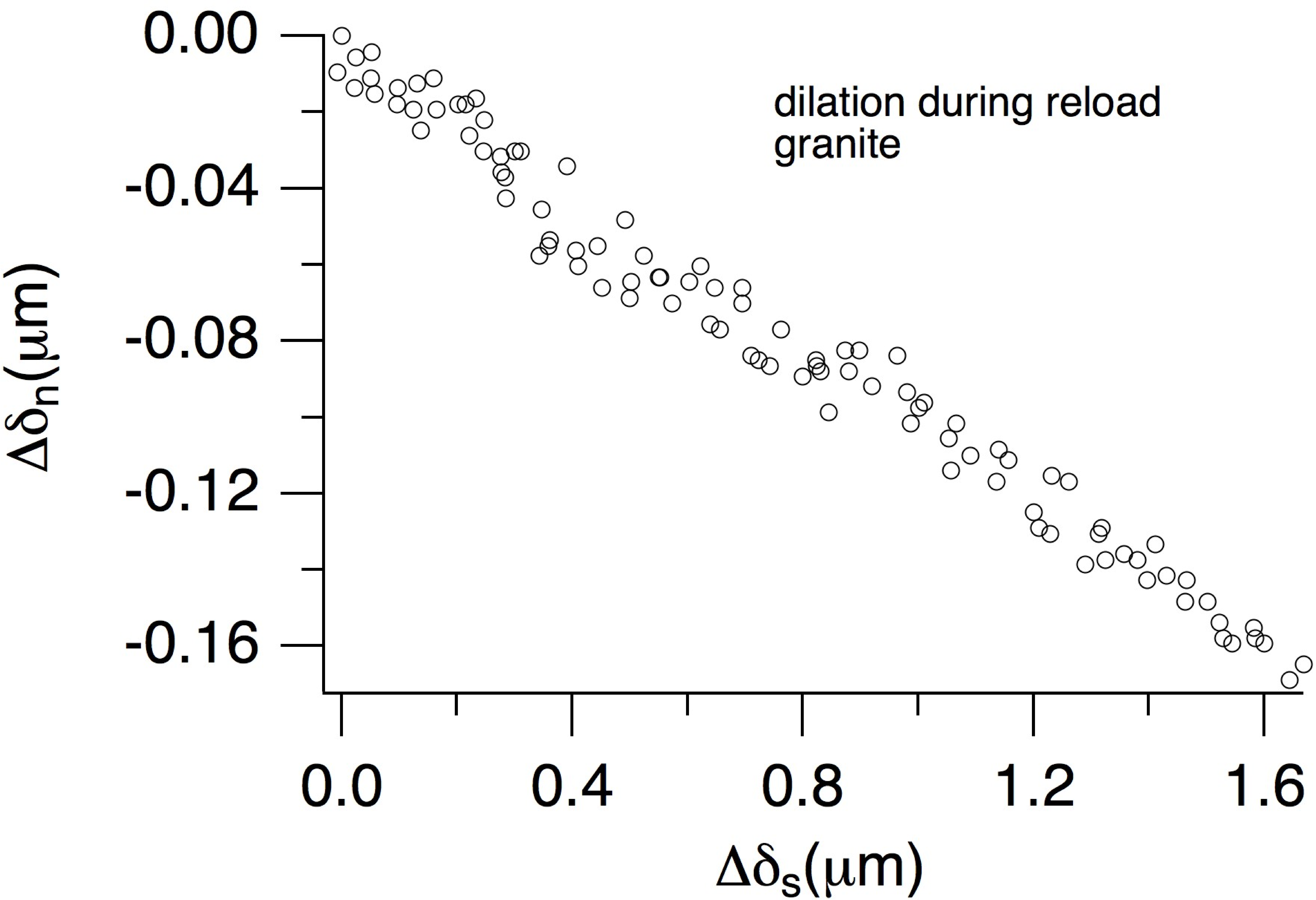

τ
fault, static contact [Dieterich and Kilgore, 1996]
σ
σ3
σm
σn = σ1


τ
contact stress, static contact [Dieterich and Kilgore, 1996]
σy /2
σ
σ3c
σmc
σnc = σy


τ
fault sliding at constant friction
φ
φ
τ
σΔ/2
σ
αp
σ3
σn
σm
σ1

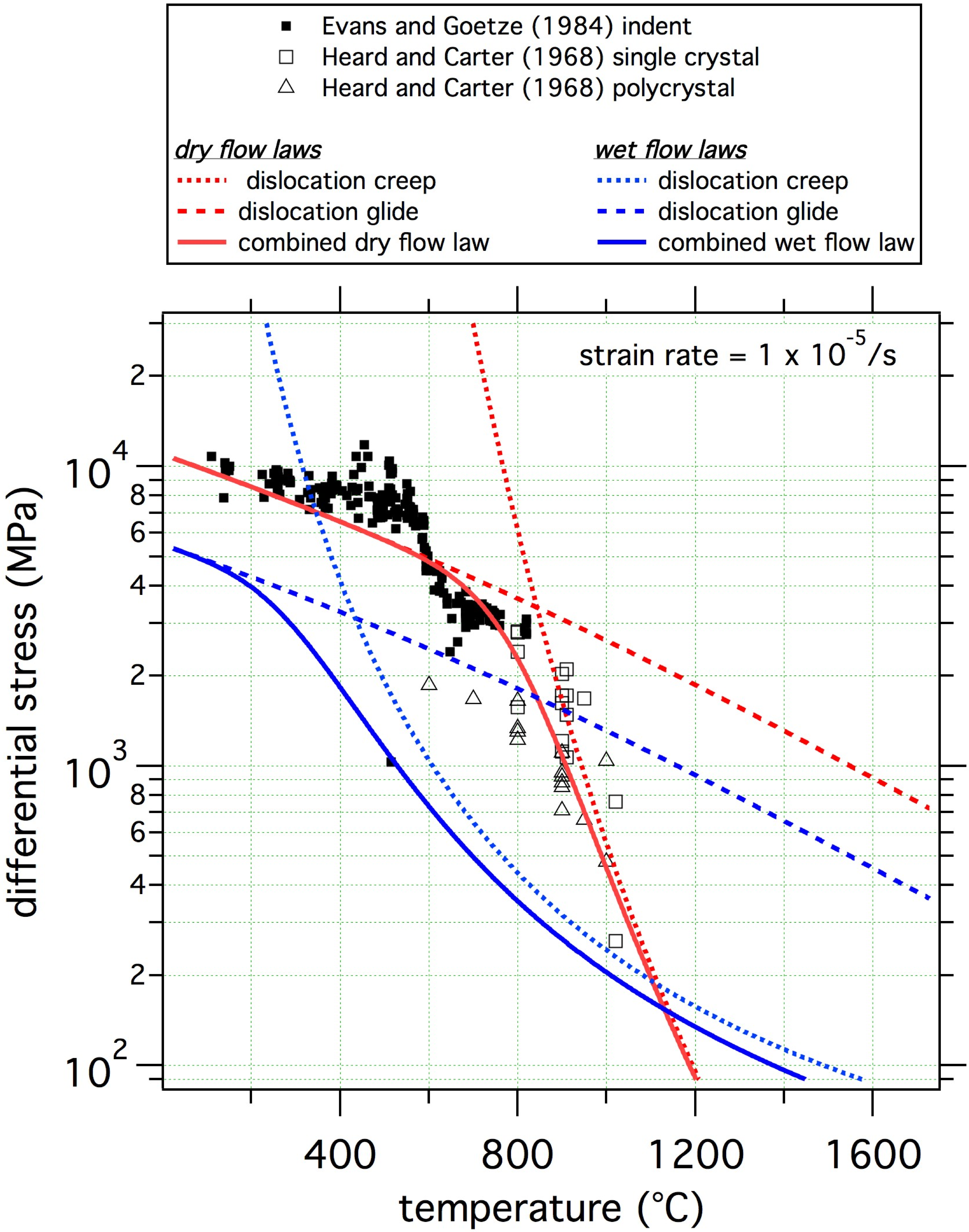
τ
contact stress, frictional sliding
σy /2
ϕ
σy /2
τc
ϕ
σ
σ3c
σc
σmc
σ1c

compaction during hold tests
granite
quartzite
0.6
0.5
0.4
0.3
0.2
0.1
$\Delta\delta_n$ (µm)
10^1
10^2
10^3
10^4
t (s)



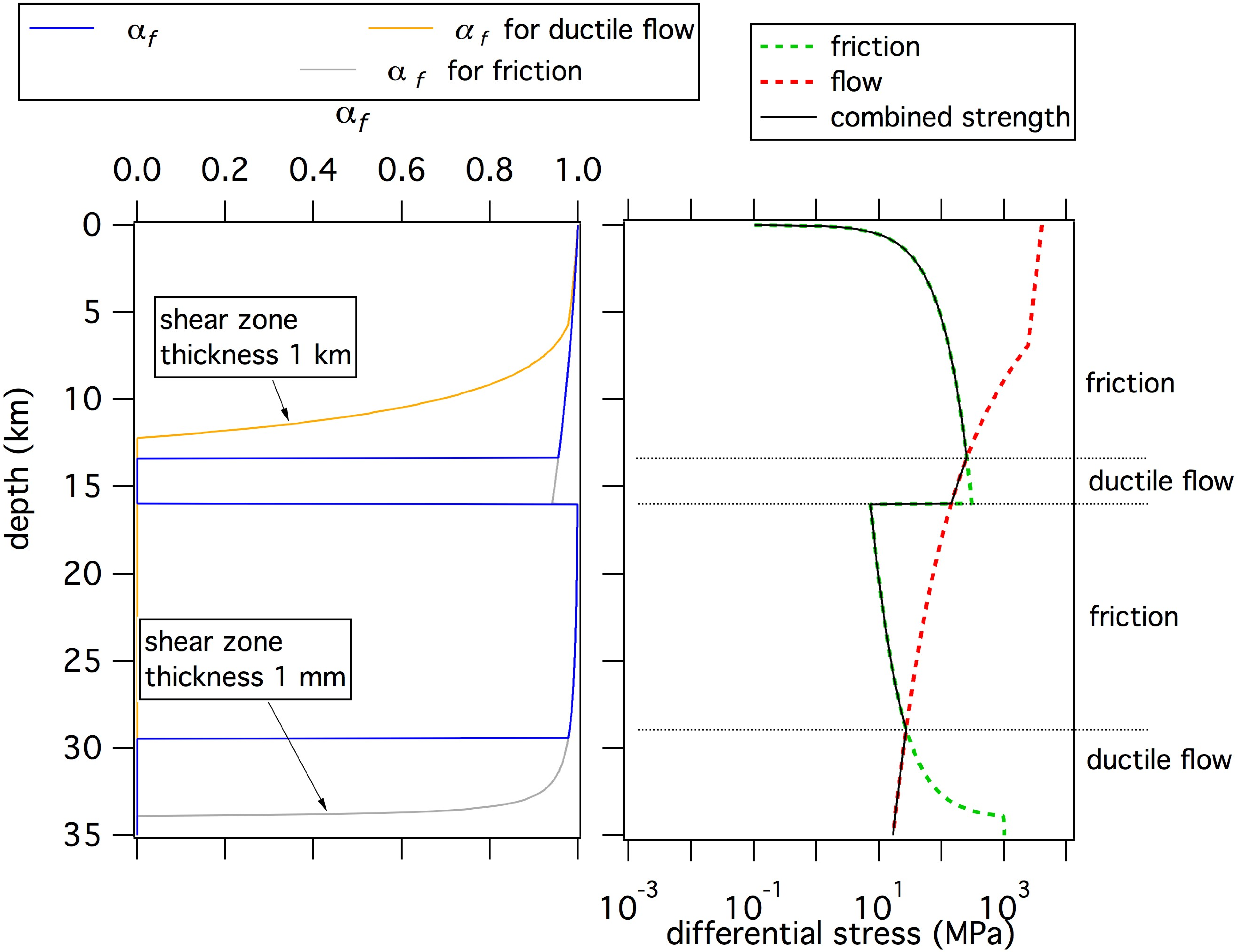
0.00
-0.04
-0.08
-0.12
-0.16
$\Delta\delta_n(\mu m)$
dilation during reload
granite
0.0
0.4
0.8
1.2
1.6
$\Delta\delta_s(\mu m)$

dilation during reload
quartzite
0.00
-0.02
-0.04
-0.06
-0.08
-0.10
0.0
0.5
1.0
1.5
$\Delta\delta_n(\mu m)$
$\Delta\delta_s(\mu m)$



Evans and Goetze (1984) indent
Heard and Carter (1968) single crystal
Heard and Carter (1968) polycrystal
dry flow laws
dislocation creep
dislocation glide
combined dry flow law
wet flow laws
dislocation creep
dislocation glide
combined wet flow law
strain rate = 1 x 10^-5/s
differential stress (MPa)
temperature (°C)
400
800
1200
1600
10^2
10^3
10^4
2
4
6
8

α_f
α_f for ductile flow
α_f for friction
α_f
0.0 0.2 0.4 0.6 0.8 1.0
depth (km)
0 5 10 15 20 25 30 35
shear zone thickness 1 km
shear zone thickness 1 mm
friction
flow
combined
friction
ductile flow
10^{-3} 10^{-1} 10^{1} 10^{3}
differential stress (MPa)
A_c / A
10^{-5} 10^{-3} 10^{-1}



α_f
α_f for ductile flow
α_f for friction
friction
flow
combined strength
α_f
0.0 0.2 0.4 0.6 0.8 1.0
depth (km)
0 5 10 15 20 25 30 35
shear zone thickness 1 km
shear zone thickness 1 mm
friction
ductile flow
friction
ductile flow
10^{-3} 10^{-1} 10^{1} 10^{3}
differential stress (MPa)

[Journal of Geophysical Research: Solid Earth]

Supporting Information for

Effective stress, friction and deep crustal faulting

N. M. Beeler, US Geological Survey, Cascades Observatory, Vancouver, Washington

Greg Hirth, Brown University, Providence, Rhode Island

Amanda Thomas, University of Oregon, Eugene, Oregon

Roland Bürgmann, University of California, Berkeley, California

Contents of this file

Introduction

The supplements to this paper contain analysis of a previously published model (Supplement 1. Effective pressure coefficient from Skempton [1960]), analysis and models of prior experiments (Supplement 2. Prior experiments on effective stress, Bishop and Skinner [1977]), a rudimentary model for effective stress (Supplement 3. Effective stress for friction with cohesion) and description of a model of sliding contact (Supplement 4. Stresses associated with sliding contact).

1. Effective pressure coefficient from *Skempton* [1960]

Despite the physical basis (**Figure 2**) and its appearance in the earthquake fault mechanics literature [*Scholz*, 1990], effective stress relations for faulting of the type described by equations (2), (3) and (4), are disputed [*Hubbert and Rubey*, 1959, 1960; *Skempton*, 1960; *Bishop and Skinner*, 1977; *Mandl*, 1988, 2000]. Unlike our conclusion $\alpha_f \leq 1$, that results from assuming the contact stresses are limited by the material yield

[*Bowden and Tabor*, 1950; *Terzahgi*, 1936], *Skempton* [1960] concludes $\alpha_f = 1$ while making exactly the same assumption of yield-limited stress. The difference lies in the contribution of pore pressure to the contact-scale stress state. Skempton assumes in addition that pore pressure on the grain or contact scale acts as a local confining stress whereas we make no such assumption. A simplified version of Skempton's derivation follows, using his notation, which differs from that in the present paper. The equivalent expressions using our notation are provided in parentheses and **Table S1** lists the equivalences.

Under dry conditions in the absence of applied shear force, the contact normal stress σ_S, the ratio of the contact normal force P_S to contact area A_S, is Nk, where N is a factor depending on the contact geometry and the stress-strain characteristics of the material, and k is an intrinsic material strength. Under wet conditions the contact normal stress is larger than under dry conditions by the pore pressure u, namely,

$$\sigma_s = Nk + u \qquad \left(\sigma_n^c = \sigma_y + p\right). \tag{S1}$$

The assumption is that pore pressure acts as a confining stress at the contact scale. The macroscopic contact normal stress is the contact stress times the area ratio, $A_S/A = a$, so

$$a\sigma_s = a\left(Nk + u\right) \qquad \left(\frac{A_c}{A}\sigma_n^c = \frac{A_c}{A}\left(\sigma_y + p\right)\right),$$

or

$$a\left(\sigma_s - u\right) = aNk \qquad \left(\frac{A_c}{A}\left(\sigma_n^c - p\right) = \frac{A_c}{A}\sigma_y\right). \tag{S2}$$

Normalizing the force balance (Figure 2), expressed in equation (2a),

$$P = P_S + \left(A - A_S\right)u \qquad \left(N = N_c + \left(A - A_c\right)p\right),$$

by total area A, fault normal stress is

$$\sigma = a\sigma_S + u - au \qquad \left(\sigma_n = \frac{A_c}{A}\sigma_n^c + p - \frac{A_c}{A}p\right), \tag{S3}$$

and

$$a = \frac{\sigma - u}{\sigma_S - u} \qquad \left(\frac{A_c}{A} = \frac{\sigma_n - p}{\sigma_n^c - p} \right). \tag{S4}$$

Substituting (A6) into (A4) is

$$a = \frac{\sigma - u}{Nk} \qquad \left(\frac{A_c}{A} = \frac{\sigma_n - p}{\sigma_y} \right). \tag{S5}$$

Accordingly, the normal stress at a representative contact is the sum of pore pressure and a term related to the yield strength (S3) whereas in (2) and (3) the contact normal stress is independent of pore pressure (because it is only assumed that the contact is at its yield stress). The result is that in Skempton's analysis the area ratio A_c/A is exactly proportional to σ_n - p via the material yield strength (S5) whereas in (2c) the proportionality is to σ_n - α_f p. Thus, in Skempton's treatment α_f is always exactly 1. During distributed deformation of soils and aggregates at low ambient applied stress, and small contact area, pore pressure may act to confine the individual particles. However, particle confinement is less likely to be an appropriate assumption in the deep crust as contact area becomes large, particularly for slip on a localized fault surface rather than bulk shear. Because we are interested in effective stress at conditions appropriate for fault slip near the BDT (high temperature, high confining pressure, lower porosity) where solid-liquid area should not differ greatly from total area minus contact area, we have used equation (4) in this study as a trial relationship to calculate effective stress at depth.

2. Prior experiments on effective stress, *Bishop and Skinner* [1977]

Bishop and Skinner [1977] conducted triaxial deformation experiments on soils and aggregates at room temperature and at nominal effective stresses (σ - p) on the order of a few tenths of an MPa specifically to determine if equation (2b) is the appropriate effective stress relation for friction. The experiments were at nominal effective stresses (σ - p) on the order of a few tenths of an MPa. The experiments were inspired by rewriting the effective stress relation (2c) using confining stress σ_3 rather than normal stress as the independent variable,

$$\sigma_3^e = (\sigma_3 - p) + \frac{A_c}{A} p, \tag{S6}$$

Note that as described in section 2.1, the model (4) can be rewritten as in (S6), in this case substituting σ_3 for σ and $\chi = 0.5(1/\sin\phi - 1)$. The experimental approach was to vary the pore pressure and confining stress from around 1 MPa up to 27 MPa holding their difference constant at a low value of 0.36 MPa. If equation (S6) is appropriate, and the fractional contact area is on the order 0.01, the imposed changes in pore pressure of 26 MPa result in a change in effective stress of 0.26, which is first order relative to the nominal difference (σ_3 - p). If on the other hand effective stress were simply Terzaghi's equation $\sigma_3^e = (\sigma_3 - p)$ there would be no change in effective stress associated with the imposed stress changes, and therefore no changes in strength. *Bishop and Skinner* [1977] were able to resolve changes in differential stress of 0.5%. In experiments on quartz sand, crushed marble, and silt, no changes in strength associated with the changes in stress state were observed. For these materials the predicted changes in differential stress from equation (4) were near the resolution of the measurements. For example a simulation with (4) for the conditions of Bishop and Skinner's experiments and an estimated yield stress of 4.9 GPa for wet room temperature deformation predicts fractional contact area of a few hundredths of a percent and small changes in differential stresses that are essentially at the resolution limit of the apparatus (**Figure S1a**).

The Bishop and Skinner experimental approach is an important method for distinguishing among effective stress models as implied by their other principal set of experiments on aggregates of lead shot. The lead experiments use the same test procedure described immediately above. Because the yield strength of lead is much smaller than for quartz, contact areas are expected to be a few percent, about 100 times larger than quartz at the same applied stress. However the lead tests are complicated by showing a very small friction coefficient of 0.1 but higher absolute strength than quartz sand. To account for the difference a 'cohesion' term can be added to the contact model. The modification and implied contact scale stress state are described in detail in Supplemental section 3 below. The modified model predicts first order changes in differential stress for the confining stress excursions between 1 and 27 MPa imposed in the experiments (**Figure S1b**). In contrast, no resolvable changes in strength were observed in the experiments. These are the only experiments to explicitly address the physical basis of effective stress.

Nevertheless, there are critical differences between the faulting model (4) and the experiments of *Bishop and Skinner* [1977]. Unfortunately because (4) is for localized fault slip it performs poorly in simulations of distributed deformation within aggregates at the low normal stresses accessible in soil mechanics tests. The model deficiency arises when the solid-liquid surface area, the area of solid that is in contact with the fluid throughout the fault zone, is large (i.e. much larger than ($1-A_c/A$)) [see *Hirth and Kohlstedt*, 1995; *Karato*, 2012]. Large solid-liquid area also means large relative to the area of any planar fault surface within the sample, as is the likely condition at low stress. Moreover, for cohesionless aggregates such as soils, if the deformation is distributed, then solid-liquid area within the shear zone may always be large even at high contact area.

To assess contributions from solid-liquid area to effective stress, consider contact area, solid-liquid area and sample external area in a geometrically simple example, a cubic-packed array of identical spheres of initial radius r_0 as it is deformed isotropically. The array is equi-dimensioned with initial length L_0. The number of spheres in the array is $N=(L_0/(2r_0))^3$. At zero strain assume point contacts (zero contact area) between the spheres. As the array is deformed assume constant solid volume and that while each sphere is truncated by six grain contacts, each of those contacts is identical with contact area that increases while the radius *r* of each grain increases uniformly. The solid liquid area associated with each grain is

$$A_{sl}^{g} = 4\pi r^{2} - 12\pi r\delta, \tag{S7}$$

where δ is the height of the missing portion of the sphere due to being truncated by a contact (truncation). Each truncation has an associated missing area $2\pi r\delta$ and there are six of them. The total solid-liquid area is the product of N and A_{sl}^{g}. The external area of the array is $12L_0(r - \delta)^2/r_0$. The assumption of constant solid volume can be applied on the scale of the unit cell that contains a single sphere, resulting in the requirement

$$r^{3}\delta - \frac{9}{2}\delta r + \frac{3}{2}\delta^{3} - r_0^{3} = 0. \tag{S8}$$

The area of a grain contact is determined by the amount of deformation. Using δ as a measure of the deformation, the contact radius is

$$r_c = \sqrt{2\delta r - \delta^2}. \tag{S9}$$

The area ratio associated with localized slip within such an array is the ratio of a single contact to the area of a face of the unit cell about a single grain: $A_c/A = \pi r_c^2/(4r^2)$. The porosity remains connected until the contacts intersect. This percolation threshold occurs when the contact radius is equal to $r - \delta$. The associated area ratio is $A_c/A = \pi/4$. We calculate the solid-liquid and contact area within the array as δ is varied from zero to the value associated with the percolation threshold, δ_{pt}. That threshold is reached when

$$\delta_{pt} = r\left(1 - \cos\frac{\pi}{4}\right). \tag{S10}$$

Assuming a grain radius r_0=0.5 mm, as in *Bishop and Skinner*'s [1977] lead shot experiments, and length L_0=25.4 mm, the undeformed external area of the sample is 3.9 x 10^3 mm^2. The solid-liquid area of the undeformed sample (zero contact area) is the number of grains times the surface area of a single grain, resulting in 5.1 x 10^4 mm^2, greatly exceeding the external area. This disparity between the aggregate's external surface area and its internal solid-liquid area is maintained as the array is deformed from zero strain all the way to the strain necessary to reach the percolation threshold (S10) (**Figure S2**). Thus the area over which fluid pressure is transmitted to the grains of the aggregate exceeds the area over which the external stresses are applied, regardless of the porosity, over the entire range of conditions where effective stress operates. Therefore, fluid pressure is likely to be fully efficient in reducing effective stress, as in equation (1a), during distributed deformation of soils and aggregates. In particular at contact areas of a few percent, as in the *Bishop and Skinner* [1977] lead shot experiments, the solid-liquid area is more than 10 times larger than external sample area (**Figure S2**).

There are some other significant differences between our model and the *Bishop and Skinner* [1997] experiments. In the experiments the lead shot has significantly higher shear strength than quartz sand. While this can be accounted for by adding cohesion to the model (see Supplemental section 3 below), it is unexpected and the physical basis is unclear. This material should have no shear strength at zero confining stress. Among the possible explanations is that a portion of shear strength of lead shot is due to plasticity rather than frictional sliding or true cohesion. Because lead undergoes dislocation creep at room temperature and the differential stress of plastically deforming materials is insensitive to changes in pore pressure, this is an important consideration. Unfortunately resolving these outstanding issues requires additional experiments and is beyond the scope of the present study. In the meantime, accounting for solid-liquid area appears to explain the *Bishop and Skinner* [1977] experiments.

3. Effective stress for friction with cohesion

The shear strength of lead shot in the experiments of *Bishop and Skinner* [1977] has a small pressure dependence, consistent with a friction coefficient of 0.1. But at the fixed value of σ_3 - p of 0.363 MPa the absolute strength is large at 1.1 MPa. These observations require a frictional strength relation with significant 'cohesion', c,

$$\tau = c + \mu\sigma_n^e \; . \qquad \text{(S11a)}$$

The macroscopic stress state is shown in **Figure S3a**. The fault normal stress is

$$\sigma_n = \frac{c + \sigma_3 \tan\theta - \mu\alpha p}{\tan\theta - \mu} \qquad \text{(S11b)}$$

where the angle $\theta = 45 + \phi/2$ is defined in the Mohr construction. The contact-scale force balance (**Figure 2**) requires that $\tau = \tau_c A_c/A$, and $\sigma_n^e = \sigma_c A_c/A$, just as for the cohesionless implementation, further requiring the macroscopic and contact scale frictional resistances are $\mu' = \mu + c/\sigma_n^e$. The contact-scale stresses and stress orientation are fixed by the material yield strength and by assuming no contact scale cohesion (**Figure S3b**). A simulation of steady-state friction for the *Bishop and Skinner* [1977] lead shot experiments with μ = 0.1, c = 0.45 MPa and σ_y = 93 MPa indicates changes in differential stress of ~0.1 MPa (**Figure S3b**) that were not observed in the experiments.

4. Stresses associated with sliding contact

Here we present a 2D example of the stresses associated with a sliding contact to contrast with the simple average stress analysis in the main body of this paper. The contact solution is from *Johnson* [1987], section 7.1, (p 202-209), a) cylinder sliding perpendicular to its axis along a flat surface, Johnson's equations (7.8). Here this is considered to be a possible solution for the stress distribution about a 'steady-state' representative contact on a frictional surface during sliding. The geometry is shown in **Figure S4**. The solution descends from Hertz's original analysis from uniaxial loading of spheres normal to their contact, equivalently uniaxial loading of a sphere on a flat. The Hertzian contact normal stress distribution applies here as well:

$$\sigma_n = p_0\sqrt{1-\left(\frac{x}{a}\right)^2}, \tag{S12a}$$

where a is the 1/2 length of the contact and p_0 is the normal stress at the contact center (the maximum normal stress). The contact center is at x=0 and the contact extends from -a to +a. The shear stress τ at the contact results from assuming a contact scale friction coefficient μ, requiring that $\tau=\mu\sigma_n$, and

$$\tau = \mu p_0\sqrt{1-\left(\frac{x}{a}\right)^2}. \tag{S12b}$$

This assumption of micro-scale friction is consistent with the assumptions in the text of this paper that result in there being a constant micro-scale friction coefficient. However, while in this *Johnson* [1987] solution the shear and normal stresses are symmetric about the contact, stress in the plane of the contact is not

$$\sigma_s = p_0\left[\sqrt{1-\left(\frac{x}{a}\right)^2} + 2\mu\frac{x}{a}\right]. \qquad \text{(S12c)}$$

An example of these stresses (S12a) - (S12c) is shown in **Figure S5** for a case where the average normal and shear stresses are 3.2 and 1.9 GPa, respectively. The average contact normal stress is related to the maximum contact normal stress as in Hertz's original solution

$$\overline{\sigma_n} = \frac{p_0}{a}\int_{-a}^{a}\sqrt{1-\left(\frac{x}{a}\right)^2}\,dx = 0.785 p_0\,. \qquad \text{(S12d)}$$

The differential stress at the contact is

$$\sigma_\Delta = 2\sqrt{\left[\frac{(\sigma_s-\sigma_n)}{2}\right]^2 + \tau^2}\,. \qquad \text{(S12e)}$$

When evaluated the differential stress turns out to be constant within the contact (as shown in **Figure S5**), $\sigma_\Delta = 2\mu p_0$, equivalently, expressed as a constant fraction of the average normal stress, $\sigma_\Delta = \frac{2\mu\overline{\sigma_n}}{0.785}$.

This is an interesting result for comparison with the average stress model in the text of this paper where the contact differential stress is calculated from an assumed friction coefficient and yield stress resulting in average contact shear and normal stresses. In this Johnson model the shear and normal stresses are spatially varying at the contacts and the stress state is asymmetric about the contact due to the requirement of ongoing slip. Nonetheless, the differential stress at the contact that leads to yielding there is neither asymmetric nor spatially varying.

4.1 Comparison with the average stress model at yield

In the average representation of contact stresses in the text of this paper, two constants are assumed, a macroscopic friction coefficient that due to the force balance dictates an equivalent microscopic friction, μ, and a yield stress σ_y. According to the assumptions, these values completely specify the stress state at the contact as shown in **Figure 3d**. The contact shear and normal stresses are related by the friction coefficient $\sigma_c = \tau_c / \mu$ and the contact normal stress is

$$\sigma_c = \frac{\sigma_y}{2\mu}\cos\left(\tan^{-1}\mu\right). \tag{S13}$$

Equating the contact normal stress (S12e) with the average contact normal stress in the *Johnson* [1987] model above (S13e), $\sigma_\Delta = \frac{2\mu\overline{\sigma_n}}{0.785}$, the differential stress in the *Johnson* model is

$$\sigma_\Delta = \frac{\sigma_y \cos\left(\tan^{-1}\mu\right)}{0.785}. \tag{S14}$$

For $\mu = 0.6$ as assumed for the average contact stress model in the text of this paper, $\sigma_\Delta = 1.0929\sigma_y$. Thus, for conditions of yielding in the average contact stress model, the predicted differential stress at the contact of the *Johnson* model differs by only 9%. Nonetheless, this is example is of the limited application to the deep crust as it is entirely elastic. The relation between contact scale stress state and the macroscopic shear resistance during sliding remains largely unexplored, particularly at elevated temperature. Some additional considerations for elastic friction models are found in *Boitnott et al*. [1992] and references therein.

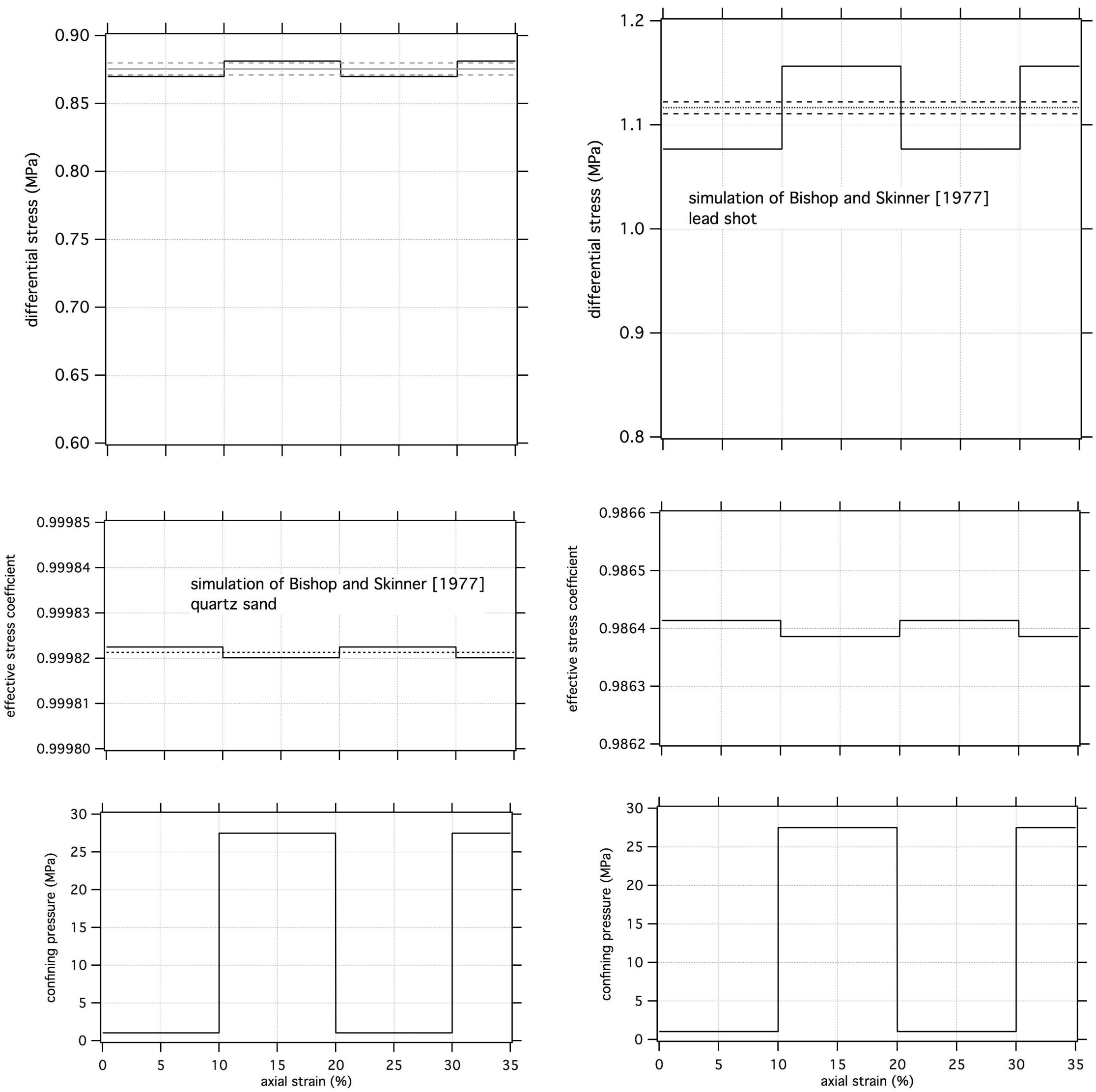


Figure S1. Simulation of *Bishop and Skinner*'s [1977] soil mechanics experiments. a) Simulation for quartz sand that assumes the yield stress is 4.9 GPa and the friction coefficient is 0.65. The lower plot shows the imposed variation in confining stress. The pore pressure changes in tandem with confining stress so that their difference is constant. The dotted and dashed lines on the upper plot are, respectively, the mean differential stress and the limits of resolution on differential stress in the experiments (+/- 0.5%). b) Simulation for lead shot that assumes the yield stress is 93 MPa, the friction coefficient is 0.1 and cohesion is 0.45 MPa.

.

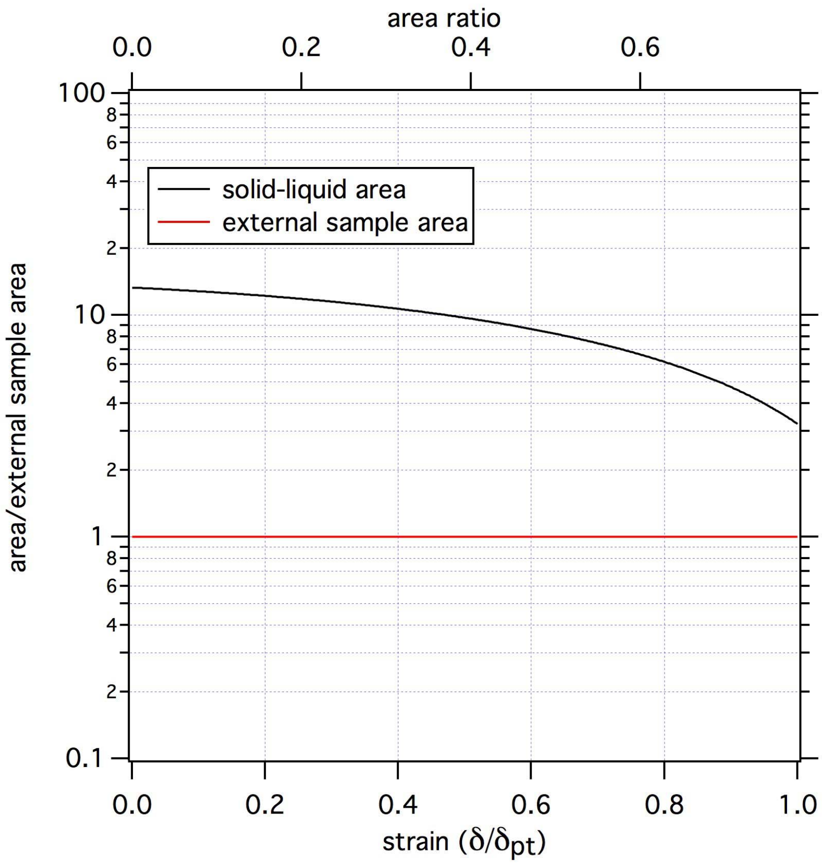


Figure S2. Solid-liquid area and external sample area for a 2.54 x 2.54 cm cubic-packed array of 1 mm diameter identical spheres as the array is deformed isotropically. The horizontal axis is a measure of strain where the deformation necessary to reach the percolation threshold δ_{pt} is the reference length.

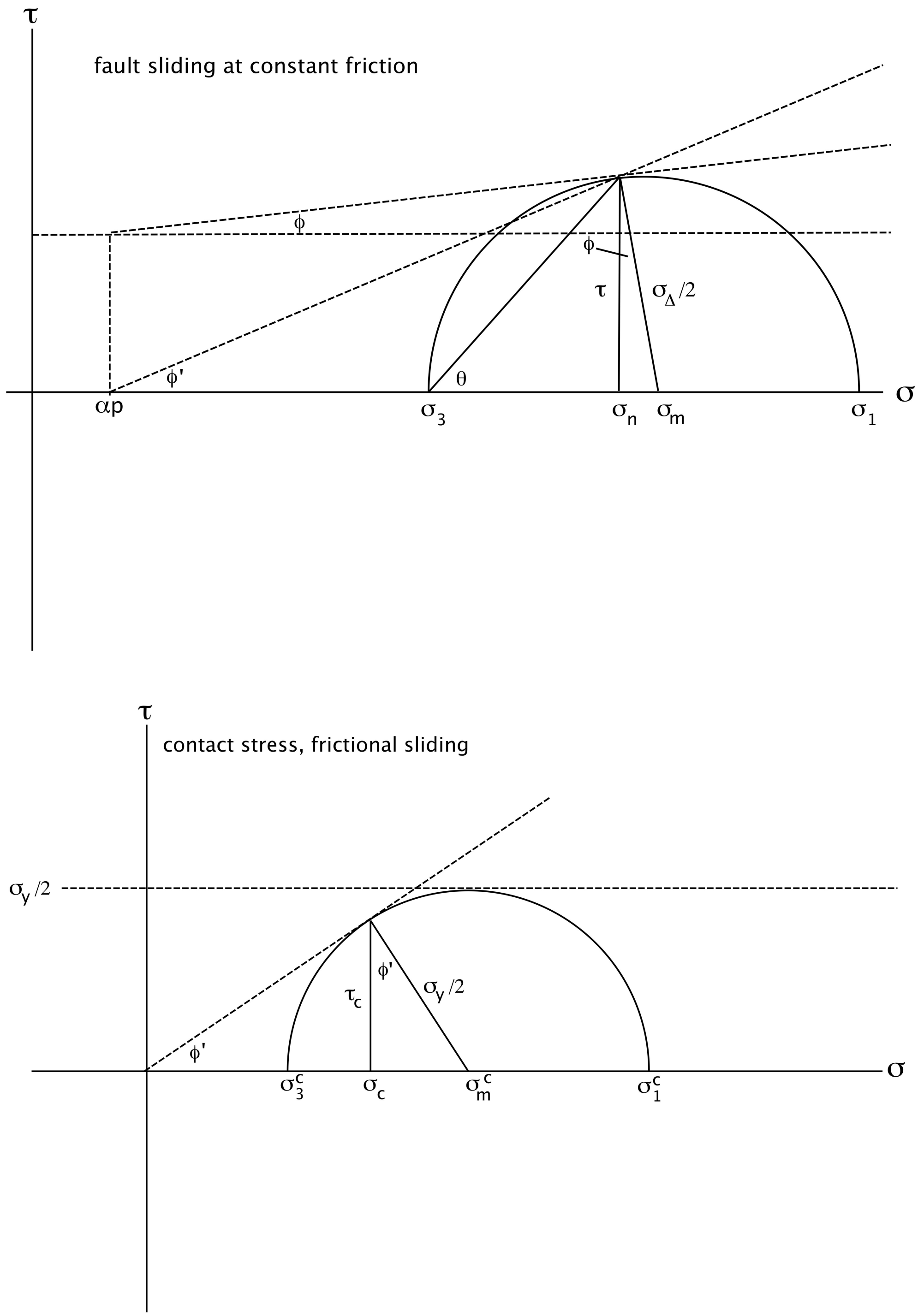


Figure S3. Mohr diagrams of stress. a) A fault with cohesion optimally oriented for slip. b) Contact stresses for the case shown in a) assuming stress is limited by yielding.

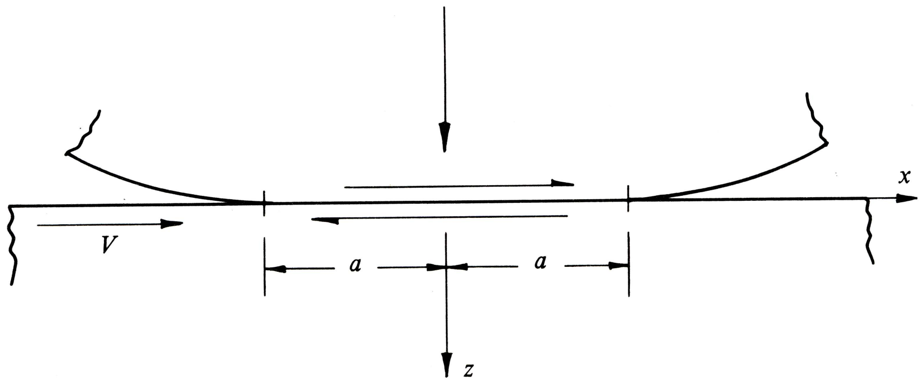


Figure S4. Geometry of the *Johnson* [1987] sliding contact solution described in Supplement section 4. The contact is between an infinite length cylinder sliding normal to its axis on a flat surface. The slip direction is *x*, the fault normal direction is *z* and the contact half-length is *a*. (Figure is modified after *Johnson* [1987], Figure 7.1)

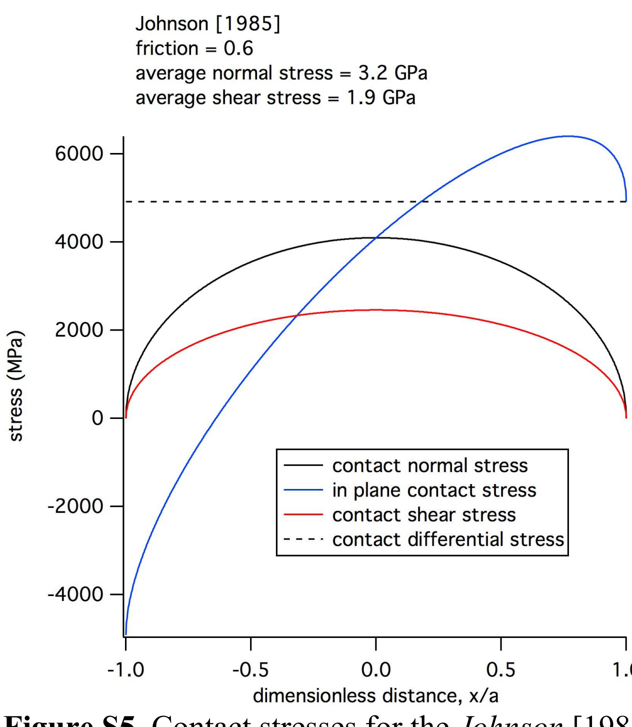


Figure S5. Contact stresses for the *Johnson* [1987] sliding contact solution described in Supplement section 4.

Skempton's notation	definition	this paper's notation
P	normal force	N
P_s	normal force at contact	N_c
u	pore pressure	p
A_s	contact area	A_c
A	total area	A
$a=A_s/A$	area ratio	A_c/A
σ_s	contact normal force	σ_n^c
Nk	yield strength	σ_y

Table S1..