

Peer Review File

Deformation evolves from shear to extensile in rocks due to energy optimization



Open Access This file is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. In the cases where the authors are anonymous, such as is the case for the reports of anonymous peer reviewers, author attribution should be to 'Anonymous Referee' followed by a clear attribution to the source work. The images or other third party material in this file are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Web links to the author's journal account have been redacted from the decision letters as indicated to maintain confidentiality

9th Jun 23

Dear Dr McBeck,

Please allow me to sincerely apologise for the long delay in sending a decision on your manuscript titled "Deformation evolves from shear to extensile in rocks due to energy optimization". It has now been seen by 2 reviewers, and we include their comments at the end of this message. They find your work of interest, but some important points are raised. We are interested in the possibility of publishing your study in Communications Earth & Environment, but would like to consider your responses to these concerns and assess a revised manuscript before we make a final decision on publication.

We therefore invite you to revise and resubmit your manuscript, along with a point-by-point response that takes into account the points raised. Please highlight all changes in the manuscript text file.

In particular, please ensure that the revised manuscript meets the following editorial thresholds:

- * Present additional evidence to support your conclusion of a transition from shear to tensile fractures in your experiments including a thorough and robust treatment of uncertainties.
- * Provide sufficient details of your methods and analytical approach to ensure that your work could be repeated.
- * Place your findings clearly in the context of relevant previous studies, both in terms of rock deformation experiments and synchrotron work.

We are committed to providing a fair and constructive peer-review process. Please don't hesitate to contact us if you wish to discuss the revision in more detail.

Please use the following link to submit your revised manuscript, point-by-point response to the referees' comments (which should be in a separate document to any cover letter) and the completed checklist:

[Link redacted]

** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

We hope to receive your revised paper within six weeks; please let us know if you aren't able to submit it within this time so that we can discuss how best to proceed. If we don't hear from you, and the revision process takes significantly longer, we may close your file. In this event, we will still be happy to reconsider your paper at a later date, as long as nothing similar has been accepted for publication at Communications Earth & Environment or published elsewhere in the meantime.

We understand that due to the current global situation, the time required for revision may be longer than usual. We would appreciate it if you could keep us informed about an estimated timescale for resubmission, to facilitate our planning. Of course, if you are unable to estimate, we are happy to accommodate necessary extensions nevertheless.

Please do not hesitate to contact us if you have any questions or would like to discuss these revisions further. We look forward to seeing the revised manuscript and thank you for the opportunity to review your work.

Best regards,

Joe Aslin

Senior Editor,
Communications Earth & Environment
<https://www.nature.com/commsenv/>
Twitter: @CommsEarth

EDITORIAL POLICIES AND FORMATTING

We ask that you ensure your manuscript complies with our editorial policies. Please ensure that the following formatting requirements are met, and any checklist relevant to your research is completed and uploaded as a Related Manuscript file type with the revised article.

Editorial Policy: [Policy requirements](https://www.nature.com/documents/nr-editorial-policy-checklist.pdf) (Download the link to your computer as a PDF.)

Furthermore, please align your manuscript with our format requirements, which are summarized on the following checklist:

[Communications Earth & Environment formatting checklist](https://www.nature.com/documents/commsj-phys-style-formatting-checklist-article.pdf)

and also in our style and formatting guide [Communications Earth & Environment formatting guide](https://www.nature.com/documents/commsj-phys-style-formatting-guide-accept.pdf).

*** DATA: Communications Earth & Environment endorses the principles of the Enabling FAIR data project (<http://www.copdess.org/enabling-fair-data-project/>). We ask authors to make the data that support their conclusions available in permanent, publically accessible data repositories. (Please contact the editor if you are unable to make your data available).

All Communications Earth & Environment manuscripts must include a section titled "Data Availability" at the end of the Methods section or main text (if no Methods). More information on this policy, is available at <http://www.nature.com/authors/policies/data/data-availability-statements-data-citations.pdf>.

In particular, the Data availability statement should include:

- Unique identifiers (such as DOIs and hyperlinks for datasets in public repositories)
- Accession codes where appropriate
- If applicable, a statement regarding data available with restrictions
- If a dataset has a Digital Object Identifier (DOI) as its unique identifier, we strongly encourage including this in the Reference list and citing the dataset in the Data Availability Statement.

DATA SOURCES: All new data associated with the paper should be placed in a persistent repository where they can be freely and enduringly accessed. We recommend submitting the data to discipline-specific, community-recognized repositories, where possible and a list of recommended repositories is provided at [a](#)

<http://www.nature.com/sdata/policies/repositories>><http://www.nature.com/sdata/policies/repositories>.

If a community resource is unavailable, data can be submitted to generalist repositories such as figshare or Dryad Digital Repository. Please provide a unique identifier for the data (for example a DOI or a permanent URL) in the data availability statement, if possible. If the repository does not provide identifiers, we encourage authors to supply the search terms that will return the data. For data that have been obtained from publically available sources, please provide a URL and the specific data product name in the data availability statement. Data with a DOI should be further cited in the methods reference section.

Please refer to our data policies at http://www.nature.com/authors/policies/availability.html.

REVIEWER COMMENTS:

Reviewer #1 (Remarks to the Author):

Please find my full review report attached.

Reviewer #2 (Remarks to the Author):

Please, find my report as PDF attachment below.

**Review of *Deformation evolves from shear to tensile in rocks due to energy optimization* by
McBeck et al.**

This paper investigates the evolution of microcrack orientation in a suite of triaxial compression experiments on a crystalline rock that systematically change the confining pressure, finding that the microcrack orientations transition from shear to tensile approaching system-sized failure. They explain these results in terms of energy optimisation using numerical models of frictional sliding and fracture growth that optimise the total mechanical efficiency of the system. I enjoyed reading this paper. It is clearly written and engaging, the methods are sound and state-of-the-art, and the conclusions are novel with mostly good evidence to support them. However, my main concern is the observations on which the paper's claims are based, specifically the observation of a transition in the geometry of the microcracks from shear to tensile, contradicts the majority of previous observations of damage development in crystalline rocks that show a transition from mainly tensile to mainly shear geometry. More evidence of the shear fractures identified in the x-ray images would strengthen the paper's claims, along with a more detailed discussion of why the geometry might be evolving in this way. I also suggest a few minor revisions. Details below.

Best, Alexis Cartwright-Taylor

What are the major claims of the paper?

This paper claims that the geometry of fracture networks evolves from shear-dominated to tensile-dominated by optimising the change in external work done during fracturing. They also claim that a new physical interpretation of the internal friction coefficient may include this work optimisation.

Are the claims novel, and will they be of interest to others in the community? If not, please identify the major papers that compromise novelty.

This paper provides new evidence for a specific way in which the geometry of fractures evolves to optimise the energy expended during deformation. The physical interpretation of the internal friction coefficient in rocks remains an outstanding question in the field, and this paper provides a new view on this. These findings will be of interest to scientists working on understanding the evolution of fracture networks in Earth materials, and to materials scientists more generally.

Will the paper influence thinking in the field in terms of either conceptual understanding or technological capability?

Yes, I believe that the paper is likely to influence conceptual understanding of how fracture networks evolve.

Are the claims convincing? If not, what further evidence is needed?

The claims are mostly convincing. However, the observation of a geometrical transition from shear to tensile microcracks contradicts the majority of previous observations of damage development in crystalline rocks that show a transition from mainly tensile to mainly shear geometry, e.g., ref 19 (Stanchits et al., 2006 - Aue granite $P_c = 60$ and 40 MPa); Stanchits and Dresen (2010 - Aue granite $P_c = 10$ MPa); ref 20 (Graham et al., 2010 - Aue granite $P_c = 20$ MPa); ref 4 (Cartwright-Taylor et al., 2020 - Ailsa Craig granite $P_c = 15$ MPa, both intact and heat-treated); Petružálek et al. (2018 - Westerly granite $P_c = 0$ MPa); Satoh et al. (1990 - Oshima granite $P_c = 40$ MPa); King et al. (2021 - Alzo granite $P_c = 5$ - 40 MPa). Some older studies do report dominance of shear events, e.g., Chang and Lee (2004 - Hwangdeung Granite and Yeosan Marble $P_c = 0$ - 20 MPa) and Zang et al. (1998 - Aue granite $P_c = 0$ MPa), although Zang et al. (1998) showed shear dominance overall but no indication

of how the source mechanisms evolve over time. But the complexity of fracture geometry evolution is highlighted in Shah and Labutz (1995 - Charcoal granite $P_c = 0$ MPa), who reported that most events had a shear component but the majority of source planes were vertically oriented, and in Lei et al. (1992), who found that the dominant source mechanism was different between granites of different grain-size (coarse-grained Inada granite $P_c = 50$ MPa showed dominant shear fracturing throughout while fine-grained Oshima granite $P_c = 40$ MPa showed tensile transitioning to shear, although the confining pressures for the two rocks were different, and Satoh et al., 1990 report Oshima granite as medium-grained).

The paper would therefore be strengthened by some further evidence:

1. Include in figure 2, similar images of the identified fractures around 85-90% of the failure stress showing the fracture network when shear cracks dominate the volume proportion.
2. A figure of example greyscale slices at 85-90% of the failure stress when there is a greater volume proportion of shear cracks and at >99% of the failure stress when tensile cracks dominate, showing (i) how the orientation of the microcracks relate spatially to the mineral grains, the orientations of grain boundaries and pre-existing cracks, and (ii) identification of the initiation methods of the microcracks (e.g., intergranular cracks at grain boundaries vs intra- or trans-granular cracks) and whether they differ systematically for each fracture type. This would provide evidence to understand what in the microstructure might be influencing the microcrack geometry, such as why certain crack orientations are more mechanically efficient and to what extent the heterogeneity of the microstructure plays a role. What is the potential for the segmentation method to identify unfractured grain boundaries as fractures due to contrasts between the attenuation intensity of different mineral grains or the presence of phase contrast artefacts?

The modelling results make a convincing argument for an evolution of fracture geometry from shear to tensile. However, the definition of shear cracks ($17-32^\circ$) in the experimental data seems to bake in the theory of internal friction, which is presented as a finding in the abstract. Such low angles are likely to still have a large volumetric strain component and could even be tensile fractures that have experienced rotation. An alternative might be to use equivalent strain tensors for ideal extensile (0 degrees) and pure shear (45 degrees), and use that to make the discrimination, i.e. independent of Coulomb theory.

The most efficient starting geometry in the models is $c=35^\circ$ (line 306), which is larger than the 32° limit on the shear fractures defined in the experiments. Some evidence of the distribution of experimental fracture angles within the $17-32^\circ$ limit would be useful. This might show whether the distribution is weighted to one end or the other. The discussion in lines 303-315 does not say much about fractures with angles down to 17° .

Also, if the shear to tensile evolution is the most mechanically efficient, then the work optimisation explanation does not explain previous observations of tensile to shear evolution cited above. Furthermore, the models do not fully explain the common observation of en-echelon tensile microcracks localising along a plane with an angle consistent with the Mohr-Coulomb failure criterion but without the presence of shear microcracks from which to evolve.

The paper would be strengthened by discussion of these points, and some further evidence:

3. How are the experimentally-observed largest six fractures at $\sigma_1 = 45$ MPa arranged spatially within the sample? Are they clustered together or distributed randomly throughout the sample volume? And are they shear or tensile, or both?
4. Did the initial crack geometries in the models include microcrack clusters? Enhanced compaction of a region containing a cluster of tensile microcracks might promote shear motion elsewhere in the sample, accommodated by tensile cracks, as suggested in ref. 21.
5. Are the pre-existing through-going faults represented as healed in the second stage of growth? Or with different local strength parameters? How was slip only along the through-going fracture prevented?
6. How does the evolution of modelled fractures compare with the evolution of fractures observed in the greyscale images? Do tensile fractures evolve from shear fractures in the greyscale, or are the tensile microcracks initiated in other ways? How does the reality of the experimental fractures evolving in a heterogeneous microstructure compare with the model results?

Are the claims appropriately discussed in the context of previous literature?

Yes, mostly. However, the paragraph ending with the sentence (lines 394-399) *'This result differs from the evolution inferred from theory and some laboratory observations of triaxial compression with low and moderate confining stresses that suggest a transition from extensile-dominated deformation under lower differential stress early in loading to shear-dominated deformation under higher differential stress later in loading, and following the peak stress (Figure 1c-d)'* would benefit from citing such studies (for example those cited above) and a following paragraph that expands on this observed contradiction, discussing the discrepancy between this study and the literature.

The explanation given in the introduction (lines 63-65 *'This apparent difference in the dominance of shear and tensile failure may occur because the opening of tensile fractures is relatively quiet acoustically, and thus less detectable than slip on shear fractures'*) does not explain why more tensile, and therefore quieter, events would be detected earlier on if shear fracturing was dominant. The correlation between AE event amplitude and the source mechanism of microcracking is unclear (see the discussion in Cheng, Y., & Wong, L. N. Y. (2018). *Microscopic characterization of tensile and shear fracturing in progressive failure in marble. Journal of Geophysical Research: Solid Earth*, 123, 204–225. <https://doi.org/10.1002/2017JB014581>, and references included therein).

How can the differences in these observations be explained? Are there any systematic differences in the definitions of crack type between this study and previous studies based on angle? For example, some of the angles defined as shear in this study may be defined as mixed-mode or tensile in AE studies. Quantification of the orientation of microcracks observed in x-ray data is estimated by, for example, a best-fitting ellipse, which averages out a crack that may have tensile and shear components. Do the authors see mixed-mode behaviour in the nucleation/propagation of the defined shear fractures? Do the authors' conclusions change if the proportion of number of shear vs tensile fractures is used rather than the volume proportion? Are there any systematic differences in the volumes of shear fractures compared with the volumes of tensile fractures in the experiments that might account for different observations between the proportion of tensile/shear AE events vs the tensile/shear volume proportion? For example, a large number of tensile cracks hosting a smaller crack volume vs a smaller number of shear fractures hosting a larger crack volume. Could differences in sample size be a factor? Could the heat treatment influence the volume proportion by introducing grain boundary microcracks from which shear sliding could initiate?

In addition to the above, further discussion points that would strengthen the paper include:

1. The limitation of the current study, whereby the experiments presented show only damage up to peak stress, with no information about what happens during the strain softening period after peak stress. This is expected given the very short strain softening period in granite vs the time required for imaging. However, a second evolution from tensile to shear immediately preceding failure during strain softening (as seen in the studies mentioned above) may be missed here. This may go part of the way to explaining the differences between AE source mechanisms the results of this study mentioned above.
2. How the microstructure influences the geometry of the microcracks, as shown in the greyscale images (suggested additional figure in point 2 above), relating this to previous observations of microcrack development in crystalline rocks.
3. There is nice summary of the model limitations in the methods, which could also include an important additional limitation that the models do not take into account variations in local strength at the grain scale due to heterogeneity in the microstructure. The models show that higher material strength promotes the development of tensile cracks at higher loads, and local strength of individual grains is likely higher than the overall strength of the bulk rock and will vary according to their mineralogy.

Have they been fair in their treatment of previous literature?

Yes

Have they provided sufficient methodological detail that the experiments could be reproduced?

Yes

Is the statistical analysis of the data sound?

Yes

Minor comments

1. The paper would benefit from a brief summary of the main conclusions of the paper either at the end of the Discussion or in a separate Conclusions section if the structure of a *Communications Earth & Environment* paper allows it.
2. The title could give the impression that this is a general result. The differences with the literature may originate, at least in part, due to the angle ranges used to define tensile and shear. Here Mohr-Coulomb theory is used to make this distinction, rather than local strain tensors or focal mechanisms that others may have used. Hybrid or mixed-mode mechanisms likely play an important role but are not considered here.
3. Please define σ_D as the differential stress step at first use, and state somewhere that σ_{D_F} is also the peak differential stress.
4. Please state in the methods how the volumes of the tensile and shear cracks are defined (e.g., as the total volume of the voxels containing the crack).
5. Was a filter applied to remove small objects in the segmentation that might bias the orientations of the microcracks? If so, please include in the size of the filter in the methods.
6. How were the orientations of the fracture calculated in Avizo? Are they the eigenvectors of the best-fit ellipsoids calculated from the 3D moments of inertia, or some other measure?

7. Lines 61-63: *'However, some observations from acoustic emissions indicate that shear is the dominant form of deformation, rather than extension, at higher differential stresses and thus immediately preceding macroscopic failure [ref. 3 p.113, 19-20]'*. These references observed a transition from tensile-dominated to shear-dominated close to and after peak stress – please clarify this in the sentence above. It is also not clear how this sentence is relevant to the subsequent sentence.
8. Lines 132-135: *'In summary, Griffith analyses of elliptical cavities and laboratory observations indicate that extensile deformation may dominate rock deformation at least at the onset of loading in triaxial compression, if not throughout loading for confining stresses less than 10 MPa (e.g., Figure 1c).'* This seems like a critical statement that needs a bit more evidence for its generality. Also, observations of tensile-dominated deformation in the early stages of loading are not restricted to confining pressures <10 MPa (see above citations).
9. Lines 143-144: *'However, such idealized wing cracks are rarely observed in experimental data'* – yes, but maybe they capture the net effect of the superposition of tensile and shear of the population as a whole.
10. Lines 254-256: Please explain more clearly what is meant by the sentence: *'In particular, when the normalized σ_1 is greater than 99%, the proportion of the differential stress in which r is greater than one ranges from 75-100% of the total differential stress for each experiment.'*
11. Lines 373-387 (caption for Figure 5): I suggest including a sentence here explaining that for higher strength parameters (d and h), a higher axial load is needed to initiate fracturing (as also stated in the methods).
12. Lines 421-430: Here, the microstructural observations discussed in previous literature would benefit from being related to a figure showing example greyscale slices, as suggested above, to support the claim of consistency with the paper's experimental observations.
13. Lines 437-438: The phrase *'the development of a through-going fracture includes the buckling of narrow columns separated by arrays of fractures aligned subparallel to the σ_1 direction'* could also refer to the rotation of such columns as observed in ref. 21. See also lines 57-58.
14. Lines 445-446: *'Even if the results of the present study only apply for confining stresses lower than 20 MPa, and effective stresses lower than 10 MPa.'* Do you not mean confining pressures between 10 and 20 MPa, given the literature supports the opposite conclusion for <10MPa?
15. Lines 543-544: *'If the observational bias was strong, we would expect larger anisotropies for the shear fractures than the tensile fractures.'* If the tensile cracks are longer in general than the shear cracks, that could explain the overall similar or larger shape anisotropy for the tensile cracks than the shear cracks, so the detection may still favour tensile cracks. Is there any evidence for tensile cracks being longer or shorter overall than the shear cracks?
16. Supplementary Information: There are several conclusions presented in the supp. info. that would be strengthen the main text if included there. These are: (i) confining pressure rather than effective stress may be the primary control on the type of deformation, although the experiment with the lowest confining pressure also does not contain pore fluid to the presence or not of fluid may also exert a control, (ii) the damaged rocks experience larger changes, and a faster increase, in the proportion of stress steps with a greater volume of tensile microcracks than the intact rocks, and (iii) a higher fracture volume is associated with higher accumulated

axial strain and lower differential stress at failure. Perhaps these could be included in the summary of conclusions suggested in point 1.

17. Supplementary Information: It is surprising that for the experiments at $P_c = 10$ MPa and $P_p = 5$ MPa that the intact sample had more fracture volume prior to loading than the damaged sample. Did the porosity measured by sample imbibition also differ in this way? Is there an explanation for the difference? Maybe in the mineral assemblages or, sample preparation, or natural variability between samples?

N.B. All numbered references refer to the corresponding numbers in the manuscript. Un-numbered references not given in full above are as follows:

Chang and Lee (2004), Estimation of cracking and damage mechanisms in rock under triaxial compression by moment tensor analysis of acoustic emission, *International Journal of Rock Mechanics and Mining Sciences*, 41(7), 1069-1086.

King et al. (2021), Source mechanisms of laboratory earthquakes during fault nucleation and formation, *Journal of Geophysical Research: Solid Earth* 126, e2020JB021059.

Lei et al. (1992), Fractal Structure of the Hypocenter Distributions and Focal Mechanism Solutions of Acoustic Emission in Two Granites of Different Grain Sizes, *J. Phys. Earth*, 40, 617-634.

Petružálek et al. (2018), Acoustic Emission in a Laboratory: Mechanism of Microearthquakes Using Alternative Source Models, *Journal of Geophysical Research: SolidEarth*, 123, 4965–4982.

Satoh et al. (1990), Fault Development in Oshima Granite under Triaxial Compression Inferred from Hypocenter Distribution and Focal Mechanism of Acoustic Emission, *Tohoku Geophys. Journ.* 33(3,4), 241-250.

Shah and Labutz (1995), Damage mechanisms in stressed rock from acoustic emission, *Journal of Geophysical Research* 100(B8), 15527-15539.

Stanchits and Dresen, (2010) Advanced acoustic emission analysis of brittle and porous rock fracturing, *EPJ Web of Conferences* 6, 22010.

Zang et al. (1998), Source analysis of acoustic emissions in Aue granite cores under symmetric and asymmetric compressive loads, *Geophysical Journal International*, Volume 135, Issue 3, December 1998, Pages 1113–1130.

Review of manuscript “Deformation evolves from shear to extensile in rocks due to energy optimization” by McBeck et al.

This interesting manuscript reports results from triaxial compression experiments on a rock-mechanics classic – Westerly granite – observed with in-situ synchrotron micro-CT and applies numerical methods to examine how fracture propagation modes and fracture networks evolve during shortening and to test if the observed fracture orientation evolution can be explained with energy optimisation criteria. This work should be of interest to researchers concerned with brittle failure of solids, namely rock physicists, structural geologists, geophysicists, and mechanical as well as civil engineers. Overall, the results are novel and interesting, and the methodology seems fine. However, I believe that the paper requires some improvements before publication, mainly concerned with a better description of certain aspects of the methodology and raw data, adding information on errors and uncertainties, and better consideration of the broad body of existing experimental work on Westerly granite specifically, which has largely been ignored. In addition, I feel that the numerical experiments could have been tied much better to the synchrotron experiments and would have benefited from careful calibration against existing rock-physics experiments on Westerly granite. Below, I elaborate on these recommendations and add some minor comments for the authors’ and the editor’s consideration.

Many thanks for the opportunity to study this interesting manuscript.

Best wishes

Christoph Schrank

1. The authors consider fracture orientation alone as criterion for shear versus extensional deformation. As Pollard (2000) points out: “the modes (of fracture propagation) may be distinguished on the basis of the displacement discontinuity measured between the two surfaces of a fracture in the near-tip region (Lawn and Wilshaw, 1975, pp. 51±56).” He further points out: “Furthermore, one cannot know from the geometry if the average deformation across the structure was an extension or a contraction. For example, internal fluid pressure in excess of the least compressive stress would cause a contraction in the adjacent host rock as the fracture opened.” In other words, using fracture orientation as a proxy for the mode of propagation, or bulk strain around the fracture, can lead to erroneous fracture classification.

For example, in a heterogeneous aggregate made of anisotropic minerals and/or minerals with different elastic (and inelastic) moduli, the local stress and displacement fields can be very complex and “overwrite” the stresses applied at the far-field boundaries (e.g., Raghmi et al., 2019). Grain boundaries generally crack first and play an important role for the local crack path evolution (e.g., Nasser et al., 2009a). It is possible that a polycrystalline aggregate produces purely extensional grain boundary cracks that are not parallel to the maximum principal stress.

The authors used 3D DVC techniques in previous work, which – given suitable resolution – could be used for a proper kinematic classification of fractures in the present experiments. I understand that doing so would involve substantial work but it would be informative. At least, the limitations of using fracture orientation alone as kinematic proxy should be discussed. Moreover, it would be interesting to see post-mortem micrographs of the deformed specimens to study how crack paths

interact with the microstructure of the material (proportion of inter- versus transgranular cracks, etc., as has been done for Westerly granite before (Nasseri et al., 2009a; Nasseri et al., 2009b; Nasseri et al., 2007)).

2. In general, experimental errors/uncertainties are not reported. The uncertainties for strain, stress, fracture volume, and fracture orientation should be stated. Volume errors associated with fracture segmentation in particular should be reported and discussed critically. As illustrated in Fig. 1, it is unlikely that the boundaries of segmented fractures coincide with the voxel grid of the tomograph. Therefore, segmented boundary voxels of individual fractures are subject to significant uncertainty due to this inherent discretisation error (which occurs regardless of the segmentation method if binary thresholding without refinement of the voxel grid is used, as is the case here). If the fracture width is on the order of a few voxels, the related volume uncertainty can become very large (see Fig. 1). Most fractures observed in the presented experiments probably consist of only tens of voxels (see manuscript Fig. 2) and are thus subject to substantial volume uncertainty. Indeed, the authors show such an effect in their own work implicitly (McBeck et al., 2021) where they compare tomograms of fracture networks with coarse (discretisation error is very large) and fine resolution (discretisation error is smaller in relative terms). In their coarse tomograms – which have the same resolution as the data presented here – fractures made of < 75 voxels overestimate the true volume by easily a factor of 2. Moreover, the statistical properties of these small fractures will also be highly sensitive to the segmentation (thresholding) approach.

The choice of threshold will also have a large impact on the connectivity of “skinny” fractures. This is why it is important to examine how robust the data are in relationship to volume uncertainty (and threshold choice, see also (McBeck et al., 2021)). It is not straightforward to quantify the uncertainty of fracture volume due to the discretisation error. However, a very conservative upper-bound estimate can be obtained with the simple erosion-dilation method described by Liu and Regenauer-Lieb (2011). Considering that fracture-volume ratios are the key result of this study, I strongly recommend that volume errors be analysed and discussed. At early experimental stages, they certainly will play an important role.

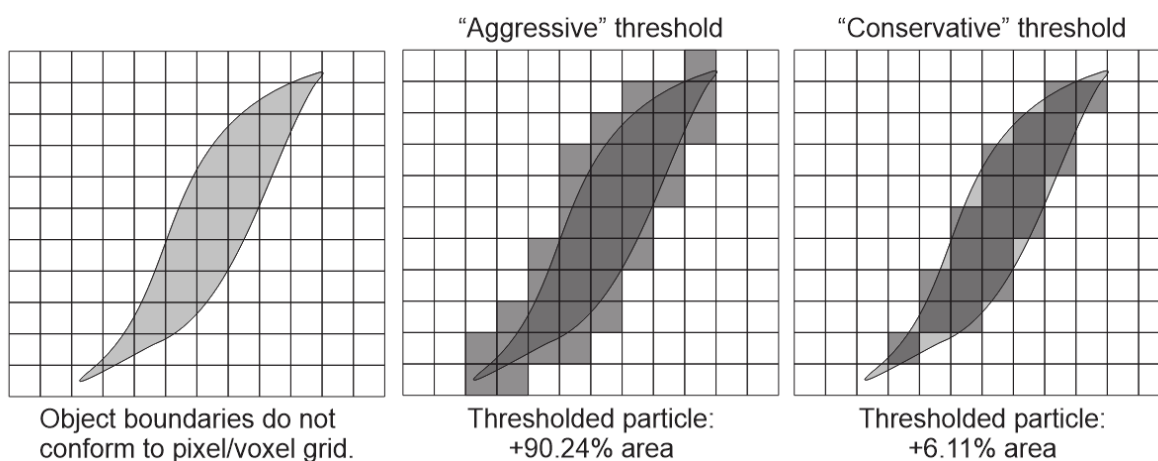


Fig. 1: Left: A “natural” fracture (light grey) superposed onto a pixel grid. Centre: Binary thresholding with a method that segments all pixels overlapping with the fracture results in an object with grossly overestimated area (volume in 3D). Right: More conservative thresholding, segmenting only pixels with > 50% fracture content, produces a smaller error on the fracture area.

3. I noticed the link to the tomograms at the end of manuscript. The dataset is > 1.3 GB, and I did not have time to download it. My question: are these the raw tomograms before segmentation, filtering, etc.? It is important to show at least some representative examples to enable the reader to assess the quality of the unsegmented tomograms. The authors should also give documented examples of the various processing steps, including the relevant parameters, applied before segmentation. Otherwise, the study is not reproducible. The thresholds for segmentation should also be reported. Were they constant for a given experiment or did the authors recompute the threshold for each individual experiment? This information should be provided. In their method section, the authors state that “(the) thresholding technique [...] produces segmented scans with porosity similar to values measured with sample imbibition [60].” Misleadingly, they cite a paper examining a porous sandstone which contains mainly interconnected pores. For such a material, sample imbibition is expected to be representative of total porosity. However, this is certainly not the case in the examined Westerly granite where, at least in early stages of the experiments, porosity is low and not interconnected. Some empirical analyses on the effect of threshold choice would be useful.

4. Almost all curves for the evolution of fracture volumina (Fig. 3, Figs. S2-S4) show very significant oscillations. Why does the volume ratio of extension over shear fractures oscillate by a factor of 2 several times over small stress increments? Total fracture volume also oscillates, most strongly in the experiment with $\sigma_2 = 20$ MPa and $P_f = 10$ MPa (Fig. S2). The physical origin of these oscillations needs to be discussed. Are these processing artifacts? Do fractures really open, close, and open again?

5. Why do the authors choose highly simplified geometries in their numerical models, such as 6 perfectly aligned fractures of equal length in an orientation consistent with the Mohr-Coulomb angles, when their code can handle 100's of fractures (Fattaruso et al., 2022). Why not make use of the actual tomograms? One could have chosen slices through an experimental cylinder at a suitable (shear-dominated) time before and after peak stress (extension-dominated) and discretised those for the numerical experiments to then compare the modelled and observed orientation and frequency statistics of the fracture network. In the authors' stage 1, about half of the fractures have orientations expected to represent mode-II/III but there is a large proportion of fractures representing mode-I. This geometry differs substantially from 6 aligned shear cracks of equal length. In lines 303 – 304, it is stated that “The fracture geometry that produces the largest ΔW [...] shares key aspects with the experimental fracture networks. At this early stage of fracture development, the most efficient fracture network is not composed of fractures with sub-vertical orientations, but instead includes fractures oriented at $c=35^\circ$ from the σ_1 direction, consistent with the experiments.” Which early stage of the experiment in terms of normalised differential stress? How can 6 perfectly aligned model “shear fractures” be deemed consistent geometrically with the natural sample that has both vertical and inclined fractures?

6. Information on how the geometric characteristics of fractures are measured is missing (except for volume, which is defined as the sum of number of voxels multiplied by voxel volume). For example, is the orientation measured relative as that of the plane spanned by largest and second-largest eigenvector of the 3D best-fit ellipsoid to a fracture? What are fracture length and width?

7. The Discussion opens with the statement: “Our experimental and numerical results indicate that shear is the dominant form of deformation preceding about 90-95% of the differential stress at macroscopic failure (Figure 3, Figure S4).” Voxel size is 6.5 μm . Therefore, fracture with apertures smaller than the voxel size cannot be resolved. Fracture aperture analysis of natural rocks shows clearly that micro-fractures with widths less than instrument resolution are common (Hooker et al., 2014). Because of the resolution limit, the experimental data shown here do not reveal insights into fracture processes at the onset of cracking. Experiments on Westerly granite demonstrate that grain-boundary cracks form first, with aperture well below the resolution of the tomograph used here (Schrack et al., 2012; Tapponnier and Brace, 1976). This limitation should be stated and discussed. Fracture statistics of unheated and heat-treated Westerly granite are readily available in the literature (e.g., Nasser et al., 2007, and references therein).

8. Mechanical properties (elastic moduli, UCS, fracture toughness, etc.) for Westerly granite, both heat-treated and undamaged, are also widely available in the literature (see, for example, the papers by Nasser et al. and co-workers in the reference list below). Why have these data not been used in the numerical models? I also note that none of this body of work is discussed. The authors use very low values for tensile and shear strength and argue: “The rock near the tips of growing fractures is heavily damaged, and thus the strength of this rock is much lower than the bulk strength measured on intact granite samples. The Methods section and the supplementary information (Table S3) further justify the choice of these three strength parameters.” In contrast to the assertion that further justification (other than “rock near the tips of growing fractures is heavily damaged”) can be found elsewhere, this is not the case. What results do the models give if higher strength values are used? These must be somehow unfavourable, otherwise there would be no reason to use very low strength values. It would be interesting to know. A very useful approach for justification would be to calibrate the numerical models such that they can reproduce the existing experiments on (unheated and heated) Westerly granite by Nasser et al. and many others, and then apply those to the experiments presented here. These authors all studied fracture toughness, which lends itself very nicely for integration with GROW.

9. In Fig. 5, the authors report numerical results for the ratio of extension over shear fractures. First, why do these plots not show the normalised differential stress on the x-axis, as in the synchrotron experiments? This would permit a much better comparison. Second, the authors do not explain how fracture orientation and length are measured in the 2D numerical experiments. In the few model snapshots shown in Fig. 5, there one usually observes one interconnected fracture with several vertical and inclined branches (rather than multiple single fractures). So how can the authors determine the plotted ratio for such models? Ideally, the measurement strategy should be the same as that for the segmented tomograms (albeit in 2D). Whatever it is, it needs to be documented.

Detailed comments/questions

Lines 49 – 50: A pre-existing common kinematic (“strain”) term in the literature is “extension fracture” (Marrett and Peacock, 1999) rather than “extensile fracture”, as proposed here. That said, I personally tend to agree with the measured critical response of Pollard (2000) to the paper of Marrett and Peacock (1999), who prefers the nomenclature related to mode of propagation.

Lines 127 – 131: I am not sure if alternating step-overs would form in nature. The initial Griffith cracks might be randomly distributed, certainly grow at different times, and would have varying initial lengths. Once two cracks link up, one obtains a larger stepped crack. Such a crack would propagate more easily than its smaller unlinked siblings because of the inverse relationship between crack length and critical stress for propagation. In addition, due its stepped geometry, the crack-tip stress fields would become asymmetric. In short, there could be some dynamic effects that lead to a preferred step-over direction.

Line 167: The term “unparalleled” is rather sensational.

Line 169: The expression “partitioning of the local strain tensors” appears a bit awkward semantically. Better perhaps: local deformation is partitioned. An individual strain tensor quantifies locally homogeneous deformation and is not in itself “partitioned”.

Line 173: This line states “previous work has not been able to directly compare the contribution of shear and extensile deformation to the system”. However, the authors employed DVC in their own previous paper and report on contraction/dilation/shear strain. Why is that not “a direct comparison of shear and extensile deformation”?

Line 188: The adverb “explicitly” is not needed.

Line 197: I recommend replacing “unparalleled”.

Lines 231 – 232: What is the width of the median filter?

Lines 251 – 254: This has all been said at the beginning of the paragraph and should thus be removed.

Line 264: Replace “during loading” with “before peak stress”. After peak stress, loading continues for a while, after all.

Line 288: The parameter “g” for grid geometry is not well explained, here or in the supplement. Why are numbers used rather than clear statements such as “left-dipping”, “right-dipping”, “alternating dip”, etc. What would $g = 3$ or 5 denote?

Line 334 ff: The authors name material parameters without sufficient explanation. Are these the mechanical properties of the intact matrix rock or the pre-existing fractures themselves? It needs to be stated in the main text.

Line 356 ff: The authors state: “The high temporal resolution provided by the models shows that the fracture networks experience some temporary phases in which the ratio of the extensile/shear fracture length decreases.” This is also the case in the synchrotron data. Why?

Line 426: The authors do not present any observations of randomly oriented cracks at early deformation stages. Hence, the statement “consistent with our observations” seems incorrect. As argued above, their tomograms are too coarse to resolve pre-existing and early cracks.

Line 427 ff: The authors state: “The evolution observed in the present analysis agrees with microstructural observations that suggest that after about 75% of the maximum differential stress, new transgranular fractures develop at high angles to interfaces of different minerals and generally subparallel to the σ_1 direction [22].” This is not how I interpret the data in Fig. 3. In the

interval 75 to 85%, only 1 experiment shows, more or less, a monotonic increase in the proportion of extensile fractures. The authors also do not report any micrographs and hence cannot ascertain if the new fractures are transgranular cracks or not.

Lines 469 ff: “Constraining the geometry of first few fractures that fail during triaxial compression using energy optimization indicates that these fractures will not be aligned sub-parallel to the σ_1 direction, but instead at 30-35° from the σ_1 direction (Figure 4), consistent with the higher proportion of shear fractures observed early in loading in the experiments (Figure 3).” I disagree with this statement because the numerical model does not model fracture nucleation (= the “first few fractures that fail”) and thus does not show that energy optimisation would produce this geometry. In real rocks, the available evidence shows that the “first few” fractures are entirely controlled by the complex heterogeneous microstructure of natural rock.

Lines 571: “Consequently, we follow the approach of decades of previous work that have used the orientation of fractures relative to the global directions of the principal stresses to compare the relative proportions of extensile and shear deformation”. Just because something has been done for a long time, does not mean that it is correct/applicable. As stated above, the mode of propagation is strictly defined via the displacement continuity at the fracture tip – a purely kinematic quantity that can, in principle, be derived from the velocity fields obtained with DIC. The definition is not associated with stress-tensor orientation.

Line 596: You probably mean σ_2 here, rather than σ_1 .

Line 650ff: “The two-dimensional system prevents the development of polymodal faulting [69].” That reminds me: it would be useful to plot the 3D fracture orientation distribution. Do you obtain radially symmetrical fracture orientations? This adds additional complexity to interpreting the data in terms of Mohr-Coulomb plasticity.

Line 651ff: “The low strain simplification of the linear elastic models prevents simulating the full loading of a triaxial compression experiment, and thus requires simulating snapshots of the experiment at particular stress steps.” From the model snapshots shown in the paper, it appears that these models contain fractures that propagated through the whole model. At the fracture tips of shear fractures, the low-strain assumptions most certainly breaks down. The stresses there will be greatly overestimated (see Schrank et al. (2017)) and thus lead to unrealistic propagation timeframes.

Fig. S9: There are dots with error-bar-like features. It is not explained what they represent (supposedly some characteristic moments of some frequency distribution).

References

- Fattaruso, L., Cooke, M., McBeck, J., 2022. The Influence of Fracture Growth and Coalescence on the Energy Budget Leading to Failure. *Frontiers in Earth Science* 10.
- Hooker, J.N., Laubach, S.E., Marrett, R., 2014. A universal power-law scaling exponent for fracture apertures in sandstones. *GSA Bulletin* 126, 1340-1362.
- Liu, J., Regenauer-Lieb, K., 2011. Application of percolation theory to microtomography of structured media: Percolation threshold, critical exponents, and upscaling. *Physical Review E* 83, 016106.

Marrett, R., Peacock, D.C.P., 1999. Strain and stress. *Journal of Structural Geology* 21, 1057-1063.

McBeck, J.A., Cordonnier, B., Renard, F., 2021. The influence of spatial resolution and noise on fracture network properties calculated from X-ray microtomography data. *International Journal of Rock Mechanics and Mining Sciences* 147, 104922.

Nasseri, M., Tatone, B., Grasselli, G., Young, R., 2009a. Fracture Toughness and Fracture Roughness Interrelationship in Thermally treated Westerly Granite. *Pure and Applied Geophysics* 166, 801-822.

Nasseri, M.H.B., Schubnel, A., Benson, P.M., Young, R.P., 2009b. Common evolution of mechanical and transport properties in thermally cracked westerly granite at elevated hydrostatic pressure. *Pure and Applied Geophysics* 166, 927-948.

Nasseri, M.H.B., Schubnel, A., Young, R.P., 2007. Coupled evolutions of fracture toughness and elastic wave velocities at high crack density in thermally treated Westerly granite. *International Journal of Rock Mechanics and Mining Sciences* 44, 601-616.

Pollard, D.D., 2000. Strain and stress: Discussion. *Journal of Structural Geology* 22, 1359-1367.

Raghami, E., Schrank, C., Kruhl, J.H., 2019. 3D modelling of the effect of thermal-elastic stress on grain-boundary opening in quartz grain aggregates. *Tectonophysics*, 228242.

Schrank, C.E., Fousseis, F., Karrech, A., Regenauer-Lieb, K., 2012. Thermal-elastic stresses and the criticality of the continental crust. *Geochem. Geophys. Geosyst.* 13, Q09005.

Schrank, C.E., Karrech, A., Boutelier, D.A., Regenauer-Lieb, K., 2017. A comparative study of Maxwell viscoelasticity at large strains and rotations. *Geophysical Journal International*.

Tapponnier, P., Brace, W.F., 1976. Development of stress-induced microcracks in Westerly Granite. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts* 13, 103-112.

Dear Dr. Cartwright-Taylor, and Dr. Schrank,

Thank you for these helpful comments. We have modified the manuscript following your suggestions.

We now show an analysis (new Figure S9) in which we systematically varied the gray level threshold used to perform the segmentation of the tomograms at two representative scans (early in loading and immediately preceding failure) in each experiment. We find that varying this threshold does not change the key results of the analysis. Please see the response to comment #34 below for details. We also include two new figures that show the fracture geometries and microstructures of several experiments to better illustrate how the extensile and shear fractures develop (Figure 7, Figure 8). Please see the response to comments #1 and #2 for details.

We now include more details about the methods and analysis. For details, please see the responses to comments #2, 18, 19, 20, 33, 35, and 38.

We now summarize the key differences that could explain the varying results of this work and previous studies. Please see the responses to comments #5 and #10 for details.

We highlight where we have made changes in the manuscript using red colored font and the comment function of Word, so the changes following reviewer's comments #3 and #40 for example, are indicated with C#03 and C#40. Below, our responses to the reviewer's comments are in blue colored font.

Sincerely,
Jess McBeck, on behalf of the authors

REVIEWER COMMENTS:

Review of Deformation evolves from shear to tensile in rocks due to energy optimization by McBeck et al.

This paper investigates the evolution of microcrack orientation in a suite of triaxial compression experiments on a crystalline rock that systematically change the confining pressure, finding that the microcrack orientations transition from shear to tensile approaching system-sized failure. They explain these results in terms of energy optimisation using numerical models of frictional sliding and fracture growth that optimise the total mechanical efficiency of the system. I enjoyed reading this paper. It is clearly written and engaging, the methods are sound and state-of-the-art, and the conclusions are novel with mostly good evidence to support them. However, my main concern is the observations on which the paper's claims are based, specifically the observation of a transition in the geometry of the microcracks from shear to tensile, contradicts the majority of previous observations of damage development in crystalline rocks that show a transition from mainly tensile to mainly shear geometry. More evidence of the shear fractures identified in the x-ray images would strengthen the paper's claims, along with a more detailed discussion of why the geometry might be evolving in this way. I also suggest a few minor revisions. Details below.

Best, Alexis Cartwright-Taylor

We have provided more evidence to support the observations of a transition from shear to extensile deformation, including new **Figures 7 and 8**. Please see the responses to comments #1 and #2 for details below. We also summarize why we observe this evolution, and others have not previously. Please see the responses to comments #10 and #5 below.

What are the major claims of the paper?

This paper claims that the geometry of fracture networks evolves from shear-dominated to tensile-dominated by optimising the change in external work done during fracturing. They also claim that a new physical interpretation of the internal friction coefficient may include this work optimisation.

Are the claims novel, and will they be of interest to others in the community? If not, please identify the major papers that compromise novelty.

This paper provides new evidence for a specific way in which the geometry of fractures evolves to optimise the energy expended during deformation. The physical interpretation of the internal friction coefficient in rocks remains an outstanding question in the field, and this paper provides a new view on this. These findings will be of interest to scientists working on understanding the evolution of fracture networks in Earth materials, and to materials scientists more generally.

Will the paper influence thinking in the field in terms of either conceptual understanding or technological capability?

Yes, I believe that the paper is likely to influence conceptual understanding of how fracture networks evolve.

Are the claims convincing? If not, what further evidence is needed?

The claims are mostly convincing. However, the observation of a geometrical transition from shear to tensile microcracks contradicts the majority of previous observations of damage development in crystalline rocks that show a transition from mainly tensile to mainly shear geometry, e.g., ref 19 (Stanchits et al., 2006 - Aue granite $P_c = 60$ and 40 MPa); Stanchits and Dresen (2010 - Aue granite $P_c = 10$ MPa); ref 20 (Graham et al., 2010 - Aue granite $P_c = 20$ MPa); ref 4 (Cartwright-Taylor et al., 2020 - Ailsa Craig granite $P_c = 15$ MPa, both intact and heat-treated); Petružálek et al. (2018 - Westerly granite $P_c = 0$ MPa); Satoh et al. (1990 - Oshima granite $P_c = 40$ MPa); King et al. (2021 - Alzo granite $P_c = 5-40$ MPa). Some older studies do report dominance of shear events, e.g., Chang and Lee (2004 - Hwangdeung Granite and Yeosan Marble $P_c = 0-20$ MPa) and Zang et al. (1998 - Aue granite $P_c = 0$ MPa), although Zang et al. (1998) showed shear dominance overall but no indication of how the source mechanisms evolve over time. But the complexity of fracture geometry evolution is highlighted in Shah and Labutz (1995 - Charcoal granite $P_c = 0$ MPa), who reported that most events had a shear component but the majority of source planes were vertically oriented, and in Lei et al. (1992), who found that the dominant source mechanism was different between granites of different grain-size (coarse-grained Inada granite $P_c = 50$ MPa showed dominant shear fracturing throughout while fine-grained Oshima granite $P_c = 40$ MPa showed tensile transitioning to shear, although the confining pressures for the two rocks were different, and Satoh et al., 1990 report Oshima granite as medium-grained). The paper would therefore be strengthened by some further evidence:

1. Include in figure 2, similar images of the identified fractures around 85-90% of the failure stress showing the fracture network when shear cracks dominate the volume proportion.

In a new figure (**Figure 7**), we now include grayscale images of the mineral structure and fracture networks at 97% of the failure stress in order to highlight the fracture geometry when the system is dominated by shear.

2. A figure of example greyscale slices at 85-90% of the failure stress when there is a greater volume proportion of shear cracks and at >99% of the failure stress when tensile cracks dominate, showing (i) how the orientation of the microcracks relate spatially to the mineral grains, the orientations of grain boundaries and pre-existing cracks, and (ii) identification of the initiation methods of the microcracks (e.g.,

intergranular cracks at grain boundaries vs intra- or trans-granular cracks) and whether they differ systematically for each fracture type. This would provide evidence to understand what in the microstructure might be influencing the microcrack geometry, such as why certain crack orientations are more mechanically efficient and to what extent the heterogeneity of the microstructure plays a role. What is the potential for the segmentation method to identify unfractured grain boundaries as fractures due to contrasts between the attenuation intensity of different mineral grains or the presence of phase contrast artefacts?

We now include a new figure in the manuscript (**Figure 8**) that shows the fracture geometries at >99% of the differential stress at failure. This figure shows that many fractures nucleate at the edges of mineral grains, and that shear fractures often link extensile fractures.

We also include another new figure that shows the fracture geometries at 97% of the failure stress, when shear deformation dominates (**Figure 7**). In future work, we plan to quantify the volume of fractures that develop at grain boundaries, and within the same mineral (transgranular vs. intragranular). However, this quantification is outside the scope of the present work.

The microstructure influences where the first fractures nucleate, but this nucleation stage is not captured in the numerical models. Even though the numerical models do not include the heterogeneity that arises from the interlocking crystalline mineral structure of the granite, they do capture the key transition from shear to extensile deformation observed in the experiments. This suggests that the microstructure of this granite may not exert a dominant control on the transition that we investigate in this study. Furthermore, the earliest stage of the experiments that the models capture is when the differential stress is 45 MPa and when the longest fractures are about 0.3 mm. This is an important point that we now mention in the manuscript (C#02).

To account for the potential bias that unfractured grain boundaries may be identified as fractures due to phase contrast imaging, we remove from the analysis fractures with shape anisotropies (one minus the shortest/longest eigenvalue) below a particular threshold. We now mention this point in the manuscript (C#02). Visual inspection of the tomograms indicate that this segmentation procedure predominantly selects microfractures. Moreover, we did not observe many grain boundaries that were identified as microfractures with this segmentation technique.

3. The modelling results make a convincing argument for an evolution of fracture geometry from shear to tensile. However, the definition of shear cracks ($17\text{--}32^\circ$) in the experimental data seems to bake in the theory of internal friction, which is presented as a finding in the abstract. Such low angles are likely to still have a large volumetric strain component and could even be tensile fractures that have experienced rotation. An alternative might be to use equivalent strain tensors for ideal extensile (0 degrees) and pure shear (45 degrees), and use that to make the discrimination, i.e. independent of Coulomb theory.

We agree that mixed-mode fractures may play an important role in fracture development, and now mention this point in the manuscript (C#03, C#11, C#16).

Indeed, we use the internal friction coefficient and the corresponding Coulomb theory to classify the shear and extensile fractures in the experiments. However, the key result about the link between energy optimization and the internal friction coefficient arises from the numerical models, and not from the experimental observations. We highlight this point in the manuscript: “The agreement between the orientation of

these fractures [in the numerical models] relative to the σ_1 direction and the orientation predicted by the Coulomb criterion for intact rock with internal friction coefficients of 0.5-1.5 (17-32°) [3] suggests that a new physical interpretation of the internal friction coefficient may include work optimization. In particular, fracture development that optimizes the global mechanical energy is consistent with the orientation of failure planes within nominally intact rock determined from the internal friction coefficient" (C#03).

4. The most efficient starting geometry in the models is $c=35^\circ$ (line 306), which is larger than the 32° limit on the shear fractures defined in the experiments. Some evidence of the distribution of experimental fracture angles within the 17-32° limit would be useful. This might show whether the distribution is weighted to one end or the other. The discussion in lines 303-315 does not say much about fractures with angles down to 17° .

Indeed, we set a limit of 32° to classify the shear fractures in the experiments, following Mohr-Coulomb theory. However, the efficiency of models with fractures oriented at 30° differs by <10% from the efficiency of models with fractures oriented at 35° . Consequently, work optimization predicts that a fracture network including fractures oriented at both 30° and 35° are similarly likely to develop, with only a slight preference for 35° , consistent with the experimental observations. We now include this key point in the manuscript (C#04).

5. Also, if the shear to tensile evolution is the most mechanically efficient, then the work optimisation explanation does not explain previous observations of tensile to shear evolution cited above. Furthermore, the models do not fully explain the common observation of en-echelon tensile microcracks localising along a plane with an angle consistent with the Mohr-Coulomb failure criterion but without the presence of shear microcracks from which to evolve.

Indeed, the numerical models do not explain why previous work generally observed a transition from extensile to shear deformation. There are many factors that influence the difference between the results shown here and previous work, such as differences in the way that this and previous work classify shear and extensile deformation (using fracture orientation vs. acoustic emissions, and the number or length of fractures vs. the volume of fractures) and the limitations inherent in these detection methods. For example, previous work did not, and could not, compare the volume of fractures that are dominated by shear and extensile deformation. We view the reliance of past work on counting the numbers of fracture or numbers of acoustic emissions as a shortcoming that the present analysis addresses. As described in the manuscript, "comparing the numbers of acoustic emissions classified as extensile or shear may be misleading because the frequency-magnitude distribution of acoustic emissions in many experiments is similar to the Gutenberg-Richter relation for earthquakes [3], and so a population of acoustic emissions at a given time includes a wide range of event magnitudes. Consequently, analyses that count the number of acoustic emissions represent two or more acoustic emissions, which may differ by orders of magnitude in energy content, as equivalent contributions to the overall deformation." (C#05).

In addition to the differences that arise from the varying ability of methods to detect different forms of deformation, we suggest that the main difference between this work and others is that the X-ray tomography technique enables detecting the small shear fractures that develop earlier in loading, and that previous techniques could not identify. As described in comment #12 below, previous work also focused on the deformation mechanisms following peak stress, whereas we could not extract this

portion of the experiment with these loading conditions. Whereas previous work identified a transition from extensile to shear deformation following peak stress, as the extensile fractures coalesced into a system-spanning shear fracture. Here, we observe a transition from shear to extensile deformation preceding the peak stress. We now address this point more clearly in the manuscript (C#05, C#10).

Regarding the experimental observation of an array of extensile fractures coalescing into a shear fracture, this stage of fracture development is likely not possible to capture with these numerical models. Even if the models were designed with an array of vertical fractures at the expected orientations, these models do not enable significant rotation of the intact material between the vertical fractures, or the vertical buckling of columns that are suggested to occur within such an array.

The paper would be strengthened by discussion of these points, and some further evidence:

6. How are the experimentally-observed largest six fractures at $\sigma_1 = 45$ MPa arranged spatially within the sample? Are they clustered together or distributed randomly throughout the sample volume? And are they shear or tensile, or both?

Because the numerical models are two-dimensional, and do not contain heterogeneities such as minerals of varying mechanical properties, we do not expect that the numerical models can capture all of the aspects of fracture network development observed in the experiments. For this reason, we use a parametric study with a simple fracture network to explore the impact of particular aspects of the fracture network arrangement (e.g., orientation and alignment) on the transition of deformation rather than a numerical simulation of one particular experiment. This simple fracture network geometry allows us to test the ability of energy optimization to predict key aspects of the fracture geometry observed in the experiments, such as the orientation of the fractures. If the transition from shear to tensile deformation is a pervasive process, then the arrangement of fractures should represent the salient aspects of the experimental network without relying on a specific fracture arrangement observed in the experiments to account for the observations. We have added further justification of this approach in the manuscript (C#06).

We set the length of the six fractures included in the models using the experimental data. Then, the work optimization analysis suggested that the most efficient orientations of these fractures are within the range expected for shear, at 35° . As described in the manuscript, this result is generally consistent with the observed preference for shear fractures in the experiments early in loading. Following this helpful suggestion, we have also checked the orientation of the six most volumetric (largest) fractures when the differential stress is 45 MPa in all of the experiments. Consistent with our previous, more general observations, we find that all of the six or five of the six (for experiment WG10) largest fractures have orientations within the range expected for shear ($>17^\circ$). We have added this point to the manuscript (C#06).

We are also interested in the spatial arrangement and localization of the fracture network. To quantify the localization of these largest fractures, we calculated the mean distance between these fractures for each experiment, and the minimum distance between the centroids of these fractures. The ratio of the mean distance and minimum distance to the distance from the center of the rock core to the upper and lower boundary of the core (parallel to the maximum compression direction) ranges from 0.68 to 0.91 (for the mean), and 0.08-0.22 (for the minimum). This suggests that the mean distance between the centroids of the fractures nearly approaches the half-length of the core, while the nearest neighbors of the fractures is closer to 10% of the half-length. It is difficult to classify whether these statistics

indicate that the fracture network is localized or not as the term localization remains rather subjective. Regardless of this ambiguity, because the models are two-dimensional, we do not expect that the localization of the fractures in the numerical models will match the localization observed in the experiments.

7. Did the initial crack geometries in the models include microcrack clusters? Enhanced compaction of a region containing a cluster of tensile microcracks might promote shear motion elsewhere in the sample, accommodated by tensile cracks, as suggested in ref. 21.

The initial crack geometries of the models only include six fractures of 0.3 mm length, and infinitely small width. We show examples of this geometry in Figure S7.

8. Are the pre-existing through-going faults represented as healed in the second stage of growth? Or with different local strength parameters? How was slip only along the through-going fracture prevented?

The through-going faults have the same material properties as the smaller faults in the numerical models. Slip occurs on both the through-going faults and the smaller faults. We now specify this point in the manuscript (C#08).

9. How does the evolution of modelled fractures compare with the evolution of fractures observed in the greyscale images? Do tensile fractures evolve from shear fractures in the greyscale, or are the tensile microcracks initiated in other ways? How does the reality of the experimental fractures evolving in a heterogeneous microstructure compare with the model results?

Because the numerical models are two-dimensional, and do not contain heterogeneities such as minerals of varying mechanical properties, we do not expect that the numerical models can capture the details of fracture nucleation. The key experimental observation that the models seek to simulate is the observed transition from shear to tensile deformation. We now state this point explicitly in the manuscript (C#09).

The new figures **Figure 7** and **8** of the slices of the grayscale images highlight that many extensile fractures initiate at the boundaries between minerals, but then propagate along orientations expected for shear fractures, as well as for extensile fractures. Please see the responses to comments #1 and #2 for details.

Are the claims appropriately discussed in the context of previous literature?

10. Yes, mostly. However, the paragraph ending with the sentence (lines 394-399) 'This result differs from the evolution inferred from theory and some laboratory observations of triaxial compression with low and moderate confining stresses that suggest a transition from extensile-dominated deformation under lower differential stress early in loading to shear-dominated deformation under higher differential stress later in loading, and following the peak stress (Figure 1c-d)' would benefit from citing such studies (for example those cited above) and a following paragraph that expands on this observed contradiction, discussing the discrepancy between this study and the literature.

We have now added several sentences to summarize the key reasons why we may observe this discrepancy. In particular: "The main difference between this previous work and the present analysis is that we are able to detect the small shear fractures early in loading, and compare the volume of the extensile and shear fractures, rather

than the number of acoustic emissions, for example [e.g., 19-20]. Moreover, we do not observe fracture development following the peak stress, when others have inferred or directly observed the coalescence of arrays of extensile fractures into system-spanning shear fractures [e.g., 18, 21].” (C#10). In addition, please see the response to comment #5 above.

11. The explanation given in the introduction (lines 63-65 ‘This apparent difference in the dominance of shear and tensile failure may occur because the opening of tensile fractures is relatively quiet acoustically, and thus less detectable than slip on shear fractures’) does not explain why more tensile, and therefore quieter, events would be detected earlier on if shear fracturing was dominant. The correlation between AE event amplitude and the source mechanism of microcracking is unclear (see the discussion in Cheng, Y., & Wong, L. N. Y. (2018). Microscopic characterization of tensile and shear fracturing in progressive failure in marble. *Journal of Geophysical Research: Solid Earth*, 123, 204–225. <https://doi.org/10.1002/2017JB014581>, and references included therein).

How can the differences in these observations be explained? Are there any systematic differences in the definitions of crack type between this study and previous studies based on angle? For example, some of the angles defined as shear in this study may be defined as mixed-mode or tensile in AE studies. Quantification of the orientation of microcracks observed in x-ray data is estimated by, for example, a best-fitting ellipse, which averages out a crack that may have tensile and shear components. Do the authors see mixed-mode behaviour in the nucleation/propagation of the defined shear fractures? Do the authors’ conclusions change if the proportion of number of shear vs tensile fractures is used rather than the volume proportion? Are there any systematic differences in the volumes of shear fractures compared with the volumes of tensile fractures in the experiments that might account for different observations between the proportion of tensile/shear AE events vs the tensile/shear volume proportion? For example, a large number of tensile cracks hosting a smaller crack volume vs a smaller number of shear fractures hosting a larger crack volume. Could differences in sample size be a factor? Could the heat treatment influence the volume proportion by introducing grain boundary microcracks from which shear sliding could initiate?

Following these suggestions, we have reworded the sentence in the introduction about potential differences in the ability to detect extensile and shear events with acoustic emissions (C#11).

Regarding the difference between this work and others in detecting various proportions of extensile and shear deformation, please see the response to comments #5 and #10 above.

Regarding the point about differences between the number of extensile and shear fractures relative to their volume, we view the reliance of past work on counting the numbers of fracture or numbers of acoustic emissions as a shortcoming that the present analysis addresses. As described in the manuscript, “comparing the numbers of acoustic emissions classified as extensile or shear may be misleading because the frequency-magnitude distribution of acoustic emissions in many experiments is similar to the Gutenberg-Richter relation for earthquakes [3], and so a population of acoustic emissions at a given time includes a wide range of event magnitudes. Consequently, analyses that count the number of acoustic emissions represent two or more acoustic emissions, which may differ by orders of magnitude in energy content, as equivalent contributions to the overall deformation.” (C#05).

Because we observe this evolution from shear to extensile in both the intact and damaged rocks, the shear fractures do not only initiate from the cracks introduced through heat treatment. The new figures of the grayscale images of the mineral structure (**Figures 7 and 8**) highlight how many fractures nucleate from the edges of grains.

We agree that mixed-mode fractures may play an important role, and now mention this point in the manuscript (C#11, C#16).

In addition to the above, further discussion points that would strengthen the paper include:

12. The limitation of the current study, whereby the experiments presented show only damage up to peak stress, with no information about what happens during the strain softening period after peak stress. This is expected given the very short strain softening period in granite vs the time required for imaging. However, a second evolution from tensile to shear immediately preceding failure during strain softening (as seen in the studies mentioned above) may be missed here. This may go part of the way to explaining the differences between AE source mechanisms the results of this study mentioned above.

As described in the response to comment #5, we agree with this point, and now we highlight it better in the manuscript (C#05, C#12).

13. How the microstructure influences the geometry of the microcracks, as shown in the greyscale images (suggested additional figure in point 2 above), relating this to previous observations of microcrack development in crystalline rocks.

We have included these new figures; please see the response to comments #1 and #2 above for details.

14. There is nice summary of the model limitations in the methods, which could also include an important additional limitation that the models do not take into account variations in local strength at the grain scale due to heterogeneity in the microstructure. The models show that higher material strength promotes the development of tensile cracks at higher loads, and local strength of individual grains is likely higher than the overall strength of the bulk rock and will vary according to their mineralogy.

We have added this information to the manuscript. Please see the response to comment #2 above for details.

Have they been fair in their treatment of previous literature?

Yes

Have they provided sufficient methodological detail that the experiments could be reproduced?

Yes

Is the statistical analysis of the data sound?

Yes

Minor comments

15. The paper would benefit from a brief summary of the main conclusions of the paper either at the end of the Discussion or in a separate Conclusions section if the structure of a Communications Earth & Environment paper allows it.

We agree that this could benefit the manuscript. However, this journal does not allow a separate conclusions section. The word limit of this journal prevents adding a summary paragraph unfortunately.

16. The title could give the impression that this is a general result. The differences with the literature may originate, at least in part, due to the angle ranges used to define tensile and shear. Here Mohr-Coulomb theory is used to make this distinction, rather than local strain tensors or focal mechanisms that others may have used. Hybrid or mixed-mode mechanisms likely play an important role but are not considered here.

We agree that mixed-mode fractures may play an important role, and now mention this point in the manuscript (C#16). We prefer to maintain the same title because it is the most concise summary of the key finding of the work.

17. Please define σ_D as the differential stress step at first use, and state somewhere that σ_{D_F} is also the peak differential stress.

Corrected as suggested (C#17).

18. Please state in the methods how the volumes of the tensile and shear cracks are defined (e.g., as the total volume of the voxels containing the crack).

We have now included this important point (C#18).

19. Was a filter applied to remove small objects in the segmentation that might bias the orientations of the microcracks? If so, please include in the size of the filter in the methods.

Indeed, in the analysis we only include fractures and pores with volumes greater than 20 voxels. We now include this important point in the Methods (C#19).

20. How were the orientations of the fracture calculated in Avizo? Are they the eigenvectors of the best-fit ellipsoids calculated from the 3D moments of inertia, or some other measure?

Indeed, they were calculated using the eigenvectors of the best-fit ellipsoids. We now include this point in the Methods section (C#20).

21. Lines 61-63: 'However, some observations from acoustic emissions indicate that shear is the dominant form of deformation, rather than extension, at higher differential stresses and thus immediately preceding macroscopic failure [ref. 3 p.113, 19-20]'. These references observed a transition from tensile-dominated to shear-dominated close to and after peak stress – please clarify this in the sentence above. It is also not clear how this sentence is relevant to the subsequent sentence.

We have reworded this sentence for clarity (C#21). This sentence highlights that observations of AEs find a dominance of shear events near failure, and then the subsequent sentence highlights that limitations in detecting acoustic emissions could have produced this result.

22. Lines 132-135: 'In summary, Griffith analyses of elliptical cavities and laboratory observations indicate that extensile deformation may dominate rock deformation at least at the onset of loading in triaxial compression, if not throughout loading for confining stresses less than 10 MPa (e.g., Figure 1c).' This seems like a critical

statement that needs a bit more evidence for its generality. Also, observations of tensile-dominated deformation in the early stages of loading are not restricted to confining pressures <10 MPa (see above citations).

This sentence was meant to summarize the literature described in the previous section. We have now included the citations here (C#22). We agree that previous work indicates that extensile-dominated deformation early in loading may occur in experiments with confining stresses >10 MPa. We have reworded this sentence for clarity (C#22).

23. Lines 143-144: 'However, such idealized wing cracks are rarely observed in experimental data' – yes, but maybe they capture the net effect of the superposition of tensile and shear of the population as a whole.

We have added this good point to the manuscript (C#23).

24. Lines 254-256: Please explain more clearly what is meant by the sentence: 'In particular, when the normalized σ_1 is greater than 99%, the proportion of the differential stress in which is greater than one ranges from 75-100% of the total differential stress for each experiment.'

We have reworded the sentence for clarity (C#24). This sentence is meant to quantify the previous sentence, which more generally states: "the volume of extensile fractures exceeds the volume of shear fractures for the majority of the accumulated time, in terms of stress, when the normalized σ_1 is greater than 99%".

25. Lines 373-387 (caption for Figure 5): I suggest including a sentence here explaining that for higher strength parameters (d and h), a higher axial load is needed to initiate fracturing (as also stated in the methods).

Corrected as suggested.

26. Lines 421-430: Here, the microstructural observations discussed in previous literature would benefit from being related to a figure showing example greyscale slices, as suggested above, to support the claim of consistency with the paper's experimental observations.

We have included these new figures. Please see the responses to comments #1 and #2 above.

27. Lines 437-438: The phrase 'the development of a through-going fracture includes the buckling of narrow columns separated by arrays of fractures aligned subparallel to the 1 direction' could also refer to the rotation of such columns as observed in ref. 21. See also lines 57-58.

We have now included this point in the manuscript (C#27).

28. Lines 445-446: 'Even if the results of the present study only apply for confining stresses lower than 20 MPa, and effective stresses lower than 10 MPa.' Do you not mean confining pressures between 10 and 20 MPa, given the literature supports the opposite conclusion for <10MPa?

Here, we wish to emphasize the experimental conditions of the present work, which include confining stresses lower than 20 MPa, and effective stresses lower than 10 MPa.

29. Lines 543-544: 'If the observational bias was strong, we would expect larger anisotropies for the shear fractures than the tensile fractures.' If the tensile cracks are longer in general than the shear cracks, that could explain the overall similar or larger shape anisotropy for the tensile cracks than the shear cracks, so the detection may still favour tensile cracks. Is there any evidence for tensile cracks being longer or shorter overall than the shear cracks?

This is a good point. Following this suggestion, we compared the aperture of the fractures classified as extensile and shear throughout each experiment. If the extensile fractures are thinner than the shear fractures, then we may observe less of them. However, comparing the aperture of the extensile and shear fractures indicates that these fractures have similar apertures throughout loading in each experiment. In particular, the ratio of the median fracture aperture of the extensile and shear fractures when the differential stress is <50%, 50-75%, and >75% ranges from 0.6-1.2, 0.5-1.13, and, 0.6-1.05, respectively, for all the experiments except experiment WG10. The mean of this ratio for these time ranges across all of the experiments except WG10 is 0.94, 0.83, and 0.89. While experiment WG10 generally has thicker extensile fractures than shear fractures, for all of the other experiments, the thickness of the extensile fractures is similar to the thickness of the shear fractures. Consequently, the thinness of the fractures did not influence their detection. We now mention this point in the manuscript (C#29).

30. Supplementary Information: There are several conclusions presented in the supp. info. that would be strengthen the main text if included there. These are: (i) confining pressure rather than effective stress may be the primary control on the type of deformation, although the experiment with the lowest confining pressure also does not contain pore fluid to the presence or not of fluid may also exert a control, (ii) the damaged rocks experience larger changes, and a faster increase, in the proportion of stress steps with a greater volume of tensile microcracks than the intact rocks, and (iii) a higher fracture volume is associated with higher accumulated axial strain and lower differential stress at failure. Perhaps these could be included in the summary of conclusions suggested in point 1.

The word limit restrictions prevent the addition of these points in the main manuscript.

31. Supplementary Information: It is surprising that for the experiments at $P_c = 10$ MPa and $P_p = 5$ MPa that the intact sample had more fracture volume prior to loading than the damaged sample. Did the porosity measured by sample imbibition also differ in this way? Is there an explanation for the difference? Maybe in the mineral assemblages or, sample preparation, or natural variability between samples?

We agree, and were also initially surprised by this result. We did not measure the porosity of these rock cores with sample imbibition. The sentence in manuscript that refers to imbibition refers to sandstone samples, and we have now modified the sentence for clarity (C#31). Part of the explanation of why the intact sample (WG19) has less fracture volume than the damaged sample (WG06) in the first scan could be because the spatial resolution of the tomogram limits the detection of some of the thinnest fractures. The similarity of the differential stress at failure for these two experiments suggests that indeed they did have similar fracture volumes. However, the differential stress at failure for the damaged sample is lower than the intact sample. This suggests that the damaged sample had a slightly higher fracture volume at the onset of loading than the intact sample, but the spatial resolution of the tomograms obscured this difference. The almost equal fracture volume of these

samples in the final scans of both experiments supports this idea. The other damaged experiment WG18 developed a quasi-horizontal, system-spanning fracture during loading into the experimental device, and so it has a much larger initial fracture volume than the intact samples, and the other damaged sample. We have included these points in the caption of Figure S6 in the supporting information.

N.B. All numbered references refer to the corresponding numbers in the manuscript. Un-numbered references not given in full above are as follows:

Chang and Lee (2004), Estimation of cracking and damage mechanisms in rock under triaxial compression by moment tensor analysis of acoustic emission, *International Journal of Rock Mechanics and Mining Sciences*, 41(7), 1069-1086.

King et al. (2021), Source mechanisms of laboratory earthquakes during fault nucleation and formation, *Journal of Geophysical Research: Solid Earth* 126, e2020JB021059.

Lei et al. (1992), Fractal Structure of the Hypocenter Distributions and Focal Mechanism Solutions of Acoustic Emission in Two Granites of Different Grain Sizes, *J. Phys. Earth*, 40, 617-634.

Petružálek et al. (2018), Acoustic Emission in a Laboratory: Mechanism of Microearthquakes Using Alternative Source Models, *Journal of Geophysical Research: SolidEarth*, 123, 4965–4982.

Satoh et al. (1990), Fault Development in Oshima Granite under Triaxial Compression Inferred from Hypocenter Distribution and Focal Mechanism of Acoustic Emission, *Tohoku Geophys. Journ.* 33(3,4), 241-250.

Shah and Labutz (1995), Damage mechanisms in stressed rock from acoustic emission, *Journal of Geophysical Research* 100(B8), 15527-15539.

Stanchits and Dresen, (2010) Advanced acoustic emission analysis of brittle and porous rock fracturing, *EPJ Web of Conferences* 6, 22010.

Zang et al. (1998), Source analysis of acoustic emissions in Aue granite cores under symmetric and asymmetric compressive loads, *Geophysical Journal International*, Volume 135, Issue 3, December 1998, Pages 1113–1130.

Reviewer #2:

Review of manuscript “Deformation evolves from shear to extensile in rocks due to energy optimization” by McBeck et al.

This interesting manuscript reports results from triaxial compression experiments on a rock-mechanics classic – Westerly granite – observed with in-situ synchrotron micro-CT and applies numerical methods to examine how fracture propagation modes and fracture networks evolve during shortening and to test if the observed fracture orientation evolution can be explained with energy optimisation criteria. This work should be of interest to researchers concerned with brittle failure of solids, namely rock physicists, structural geologists, geophysicists, and mechanical as well as civil engineers. Overall, the results are novel and interesting, and the methodology seems fine. However, I believe that the paper requires some improvements before publication, mainly concerned with a better description of certain aspects of the methodology and raw data, adding information on errors and uncertainties, and better consideration of the broad body of existing experimental work on Westerly granite specifically, which has largely been ignored. In addition, I feel that the numerical experiments could have been tied much better to the synchrotron experiments

and would have benefited from careful calibration against existing rock-physics experiments on Westerly granite. Below, I elaborate on these recommendations and add some minor comments for the authors' and the editor's consideration.

Many thanks for the opportunity to study this interesting manuscript. Best wishes

Christoph Schrank

We have included more details about the methodology, and tests of the uncertainty of the experimental data. Please see the responses to comments #33-35 for details. We provide further justification about the design of the numerical models. Please see the response to comment #36 for details. We highlight our reliance on previous experimental work on Westerly granite when designing the numerical models and discussing the observations of fracture development in the experiments and numerical models. Please see the response to comment #40 for details.

32. The authors consider fracture orientation alone as criterion for shear versus extensional deformation. As Pollard (2000) points out: "the modes (of fracture propagation) may be distinguished on the basis of the displacement discontinuity measured between the two surfaces of a fracture in the near-tip region (Lawn and Wilshaw, 1975, pp. 51±56)." He further points out: "Furthermore, one cannot know from the geometry if the average deformation across the structure was an extension or a contraction. For example, internal fluid pressure in excess of the least compressive stress would cause a contraction in the adjacent host rock as the fracture opened." In other words, using fracture orientation as a proxy for the mode of propagation, or bulk strain around the fracture, can lead to erroneous fracture classification. For example, in a heterogeneous aggregate made of anisotropic minerals and/or minerals with different elastic (and inelastic) moduli, the local stress and displacement fields can be very complex and "overwrite" the stresses applied at the far-field boundaries (e.g., Raghmi et al., 2019). Grain boundaries generally crack first and play an important role for the local crack path evolution (e.g., Nasser et al., 2009a). It is possible that a polycrystalline aggregate produces purely extensile grain boundary cracks that are not parallel to the maximum principal stress. The authors used 3D DVC techniques in previous work, which – given suitable resolution – could be used for a proper kinematic classification of fractures in the present experiments. I understand that doing so would involve substantial work but it would be informative. At least, the limitations of using fracture orientation alone as kinematic proxy should be discussed. Moreover, it would be interesting to see post-mortem micrographs of the deformed specimens to study how crack paths interact with the microstructure of the material (proportion of inter- versus transgranular cracks, etc., as has been done for Westerly granite before (Nasser et al., 2009a; Nasser et al., 2009b; Nasser et al., 2007)).

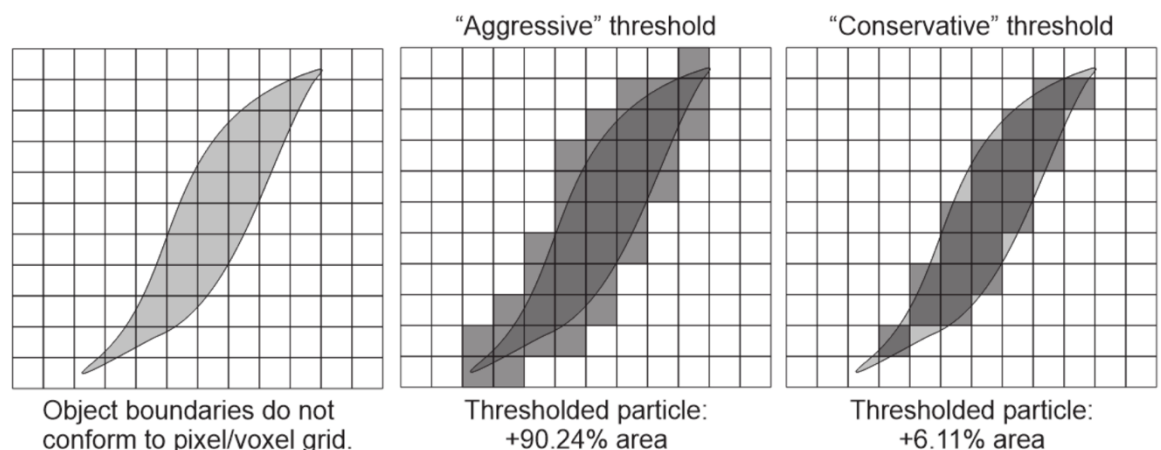
We agree that there are limitations in using the orientation of a fracture as a proxy of the dominant deformation mechanism, and we describe these points in the manuscript (paragraph highlighted with C#32). Following this comment, and #46 below, with digital volume correlation (DVC), it is possible to calculate the three-dimensional strain tensor at locations throughout a rock volume. From the strain tensor, we may calculate the volumetric strain using the first invariant of the tensor, I_1 , and quantify the shear strain using the second invariant of the deviator tensor, J_2 , for example. Although it is possible to compare the proportion of contractive and dilative volumetric strain with these data, it is not possible to compare the relative importance of dilative and shear strain in a given rock volume, or near a fracture. This is because comparing the dilative and shear strains is similar to comparing apples and oranges. We can find that the dilation is a certain value, and the strain is

a certain value at a particular point, but this comparison does not indicate if a rock volume is dominated by dilation or shear strain.

To better highlight how the grain boundaries and microstructure influences fracture development, we now include new figures of the grayscale tomograms (**Figures 7 and 8**). Please see the responses to comments #1 and #2 for details. We cannot examine post-mortem images because the samples were completely crushed following the peak stress.

33. In general, experimental errors/uncertainties are not reported. The uncertainties for strain, stress, fracture volume, and fracture orientation should be stated. Volume errors associated with fracture segmentation in particular should be reported and discussed critically. As illustrated in Fig. 1, it is unlikely that the boundaries of segmented fractures coincide with the voxel grid of the tomograph. Therefore, segmented boundary voxels of individual fractures are subject to significant uncertainty due to this inherent discretisation error (which occurs regardless of the segmentation method if binary thresholding without refinement of the voxel grid is used, as is the case here). If the fracture width is on the order of a few voxels, the related volume uncertainty can become very large (see Fig. 1). Most fractures observed in the presented experiments probably consist of only tens of voxels (see manuscript Fig. 2) and are thus subject to substantial volume uncertainty. Indeed, the authors show such an effect in their own work implicitly (McBeck et al., 2021) where they compare tomograms of fracture networks with coarse (discretisation error is very large) and fine resolution (discretisation error is smaller in relative terms). In their coarse tomograms – which have the same resolution as the data presented here – fractures made of < 75 voxels overestimate the true volume by easily a factor of 2. Moreover, the statistical properties of these small fractures will also be highly sensitive to the segmentation (thresholding) approach.

Fig. 1: Left: A “natural” fracture (light grey) superposed onto a pixel grid. Centre: Binary thresholding with a method that segments all pixels overlapping with the fracture results in an object with grossly overestimated area (volume in 3D). Right: More conservative thresholding, segmenting only pixels with $> 50\%$ fracture content, produces a smaller error on the fracture area.



This is a good point that we originally did not fully address. The errors in the reported axial stress are smaller than the size of the markers on the plots (< 0.1 MPa).

The largest source of uncertainty in this work arises from the geometry of the fractures determined from segmentation. To remove noise from the segmented fracture networks, we exclude fractures with volumes less than 20 voxels. In general,

the fractures are much larger than this limit: the mean fracture volume in each experiment is 196, 210, and 320 voxels when the normalized differential stress is 0-50%, 50-75%, and >75%. And because we use the volume of the fractures to compare the dominance of shear and extensile deformation, the most volumetric fractures, with volumes >200 voxels, dominantly controls the resulting volume proportion of shear and extensile fractures. In response to this comment, we also tested the influence of removing fractures with volumes <100 voxels, and the key results do not change.

We suggest that the segmentation method may produce larger uncertainty than the discretization method. Consequently, we systematically varied the threshold at two representative scans (early in loading and immediately preceding failure) in each experiment, and show that the central conclusions of the analysis do not change. Please see the response to comment #34 below for details.

34. The choice of threshold will also have a large impact on the connectivity of “skinny” fractures. This is why it is important to examine how robust the data are in relationship to volume uncertainty (and threshold choice, see also (McBeck et al., 2021)). It is not straightforward to quantify the uncertainty of fracture volume due to the discretisation error. However, a very conservative upper-bound estimate can be obtained with the simple erosion-dilation method described by Liu and Regenauer-Lieb (2011). Considering that fracture-volume ratios are the key result of this study, I strongly recommend that volume errors be analysed and discussed. At early experimental stages, they certainly will play an important role.

We agree that this study relies on the fracture volumes calculated from X-ray tomograms, which are subject to uncertainties related to the segmentation technique and the discretization/spatial resolution. We also agree that at early stages in experiments, the potential uncertainties will be larger than later stages. However, even when the normalized differential stress is <50%, the average volume of the fractures, across all the experiments, is near 200 voxels. As the reviewer mentions, it is difficult to characterize the uncertainty that arises from discretization. In addition, it is not possible to estimate the accuracy of a segmentation technique using X-ray tomography data without additional “ground truth” laboratory measurements, such as porosity measurements. However, because we use the fracture volume to compare the dominance of shear and extensile deformation, the orientations of the largest fractures dominantly control the resulting proportion of shear to extensile deformation.

We suggest that the segmentation technique could produce larger uncertainty than the spatial discretization of the fractures. Consequently, we now show an analysis (new Figure S9) in which we systematically varied the threshold at two representative scans (early in loading and immediately preceding failure) in each experiment. We find that varying this threshold does not change the key results of the analysis. In particular, the relative volume proportion of the extensile to shear fractures remains unchanged in each experiment, and in particular, the extensile/shear fracture volume is larger in the tomogram immediately preceding failure than in the tomogram acquired earlier in loading.

Another source of uncertainty that we considered is the choice of the angle ranges used to classify the shear and extensile fractures. We show with Figure S4 that varying these ranges, using the limits derived from the expected range of internal friction, does not change the key conclusions of this work.

35. I noticed the link to the tomograms at the end of manuscript. The dataset is > 1.3 GB, and I did not have time to download it. My question: are these the raw tomograms before segmentation, filtering, etc.? It is important to show at least some representative examples to enable the reader to assess the quality of the unsegmented tomograms. The authors should also give documented examples of the various processing steps, including the relevant parameters, applied before segmentation. Otherwise, the study is not reproducible. The thresholds for segmentation should also be reported. Were they constant for a given experiment or did the authors recompute the threshold for each individual experiment? This information should be provided. In their method section, the authors state that “(the) thresholding technique [...] produces segmented scans with porosity similar to values measured with sample imbibition [60].” Misleadingly, they cite a paper examining a porous sandstone which contains mainly interconnected pores. For such a material, sample imbibition is expected to be representative of total porosity. However, this is certainly not the case in the examined Westerly granite where, at least in early stages of the experiments, porosity is low and not interconnected. Some empirical analyses on the effect of threshold choice would be useful.

We have reworded this sentence to clarify that we tested the accuracy of the thresholding technique with imbibition on sandstone cores (C#35).

We have now described the dataset more specifically in the Data Availability section (C#35). In particular, these are the segmented data, following denoising and filtering.

We used different thresholds for each experiment, but the same threshold for all the scans in a given experiment. We have now provided a new dataset that includes the processing scripts that we used to denoise, filter, and segment the data. These scripts include the precise value of the thresholds, and the other parameters.

36. Almost all curves for the evolution of fracture volumina (Fig. 3, Figs. S2-S4) show very significant oscillations. Why does the volume ratio of extension over shear fractures oscillate by a factor of 2 several times over small stress increments? Total fracture volume also oscillates, most strongly in the experiment with $\sigma_2 = 20$ MPa and $P_f = 10$ MPa (Fig. S2). The physical origin of these oscillations needs to be discussed. Are these processing artifacts? Do fractures really open, close, and open again?

Regarding the variations in the ratio (Fig. 3), please see the response to comment #54 below.

Regarding the variations in the total fracture volume (Figs. S2-S4), some fractures do open and close with increasing differential stress loading. Previous work has documented the closure of preexisting pores and fractures with increasing confining stress, in particular (e.g., Stanchits et al., 2006). We describe in the caption of Figure 2: “The fracture volume consistently increases in some experiments (WG19_18, WG18_14, WG20_19). In contrast, the other experiments show temporary episodes in which the total fracture volume decreases (WG10, WG06_02, WG14). These decreases occur because increasing differential stress can close fractures.”

Another source of these variations could be the limit of the fracture volume that we impose. To remove artefacts, we exclude fractures with volumes less than 20 voxels, and this could produce some unrealistic oscillations as new fractures nucleate. The fact that the oscillations occur over several (tens) of scans, and not between one or two scans suggests that artefacts produced by X-ray acquisition did not produce these oscillations. We expect that the transitory artefacts related to differences

between the acquisition of subsequent tomograms would produce differences between one or two scans, and not the longer term variations. We now include these points in the caption of Figure S2.

Stanchits, S., Vinciguerra, S., & Dresen, G. (2006). Ultrasonic velocities, acoustic emission characteristics and crack damage of basalt and granite. *Pure and Applied Geophysics*, 163, 975-994.

37. Why do the authors choose highly simplified geometries in their numerical models, such as 6 perfectly aligned fractures of equal length in an orientation consistent with the Mohr-Coulomb angles, when their code can handle 100's of fractures (Fattaruso et al., 2022). Why not make use of the actual tomograms? One could have chosen slices through an experimental cylinder at a suitable (shear-dominated) time before and after peak stress (extension-dominated) and discretised those for the numerical experiments to then compare the modelled and observed orientation and frequency statistics of the fracture network. In the authors' stage 1, about half of the fractures have orientations expected to represent mode-II/III but there is a large proportion of fractures representing mode-I. This geometry differs substantially from 6 aligned shear cracks of equal length. In lines 303 – 304, it is stated that “The fracture geometry that produces the largest ΔWW [...] shares key aspects with the experimental fracture networks. At this early stage of fracture development, the most efficient fracture network is not composed of fractures with sub-vertical orientations, but instead includes fractures oriented at $c=35^\circ$ from the σ_1 direction, consistent with the experiments.” Which early stage of the experiment in terms of normalised differential stress? How can 6 perfectly aligned model “shear fractures” be deemed consistent geometrically with the natural sample that has both vertical and inclined fractures?

Regarding the design of the numerical models, and the similarity between the numerical model results and the key components of the experiments, reviewer #1 also shared these concerns. Please see the responses to comments #6 and #2 above. Your suggestion of a model that simulates the observed network is a good one for a future project that has different goals than this paper.

38. Information on how the geometric characteristics of fractures are measured is missing (except for volume, which is defined as the sum of number of voxels multiplied by voxel volume). For example, is the orientation measured relative as that of the plane spanned by largest and second- largest eigenvector of the 3D best-fit ellipsoid to a fracture? What are fracture length and width?

We have now included the key information about how we calculate the fracture orientation. Please see the response to comment #20 above for details. We do not report the fracture length and width in this work, but in general, these properties are defined by the eigenvalues of the best-fit ellipsoids.

39. The Discussion opens with the statement: “Our experimental and numerical results indicate that shear is the dominant form of deformation preceding about 90-95% of the differential stress at macroscopic failure (Figure 3, Figure S4).” Voxel size is 6.5 μm . Therefore, fracture with apertures smaller than the voxel size cannot be resolved. Fracture aperture analysis of natural rocks shows clearly that micro-fractures with widths less than instrument resolution are common (Hooker et al., 2014). Because of the resolution limit, the experimental data shown here do not reveal insights into fracture processes at the onset of cracking. Experiments on Westerly granite demonstrate that grain-boundary cracks form first, with aperture well below the resolution of the tomograph used here (Schrack et al., 2012; Tapponnier

and Brace, 1976). This limitation should be stated and discussed. Fracture statistics of unheated and heat-treated Westerly granite are readily available in the literature (e.g., Nasser et al., 2007, and references therein).

We agree that due to the limitations imposed by the spatial resolution, we cannot observe the nucleation of fractures when their aperture is below the spatial resolution. However, we are able to observe fracture nucleation at the edges of grain boundaries, as observed in previous work. The new Figures 8 and 9 show these fractures. We now describe these points in the manuscript (C#39).

40. Mechanical properties (elastic moduli, UCS, fracture toughness, etc.) for Westerly granite, both heat-treated and undamaged, are also widely available in the literature (see, for example, the papers by Nasser et al. and co-workers in the reference list below). Why have these data not been used in the numerical models? I also note that none of this body of work is discussed. The authors use very low values for tensile and shear strength and argue: "The rock near the tips of growing fractures is heavily damaged, and thus the strength of this rock is much lower than the bulk strength measured on intact granite samples. The Methods section and the supplementary information (Table S3) further justify the choice of these three strength parameters." In contrast to the assertion that further justification (other than "rock near the tips of growing fractures is heavily damaged") can be found elsewhere, this is not the case. What results do the models give if higher strength values are used? These must be somehow unfavourable, otherwise there would be no reason to use very low strength values. It would be interesting to know. A very useful approach for justification would be to calibrate the numerical models such that they can reproduce the existing experiments on (unheated and heated) Westerly granite by Nasser et al. and many others, and then apply those to the experiments presented here. These authors all studied fracture toughness, which lends itself very nicely for integration with GROW.

We did use the bulk material properties (e.g., Young's modulus) of granite reported in previous work. Please see the captions of Table S2 and Table S3 for details.

We also justified the choice of the low tensile and shear strength in the main manuscript (paragraph at C#40) and supporting information, and in particular in the caption of Table S3. Then, we tested the influence of larger and smaller strength parameters, and show these results in the old Figure S10 and now new Figure S11, and describe them in the main manuscript (C#40). Rather than disagreeing with the key observations, as suggested by the reviewer, these models all show an increasing proportion of extensile to shear deformation, consistent with the experiment. In general, increasing the values of the tensile and shear strength, and keeping the same proportionality between the two values, will only increase in the axial load required to propagate the fractures, and will not change the resulting fracture geometry.

Because these models are two-dimensional, they will not be able to simulate all of the details of fracture development observed in triaxial compression experiments. Moreover, the aim of the models was not to reproduce the experiments exactly, but rather to test the idea that energy optimization can help explain key aspects of fault development observed in the experiments. In order to test this idea, we needed to build simple fracture geometries that enable parameterization, such as modifying the fracture orientation and basic spatial arrangement. Your suggestion of a simulation study that tunes to observations is good for a future study that has different goals than this paper.

41. In Fig. 5, the authors report numerical results for the ratio of extension over shear fractures. First, why do these plots not show the normalised differential stress on the x-axis, as in the synchrotron experiments? This would permit a much better comparison. Second, the authors do not explain how fracture orientation and length are measured in the 2D numerical experiments. In the few model snapshots shown in Fig. 5, there one usually observes one interconnected fracture with several vertical and inclined branches (rather than multiple single fractures). So how can the authors determine the plotted ratio for such models? Ideally, the measurement strategy should be the same as that for the segmented tomograms (albeit in 2D). Whatever it is, it needs to be documented.

As described in the Methods (C#41): “The low strain simplification of the linear elastic models prevents simulating the full loading of a triaxial compression experiment, and thus requires simulating snapshots of the experiment at particular stress steps.” Consequently, the models only produce one or two unique macroscopic differential stresses, and so the best representation of “time” in these numerical models is the total fracture length, rather than the differential stress.

In the numerical models, we calculate the total fracture length by summing the length of all the linear elements of all the fractures at a given model time step. We calculate the fracture orientation using individual elements, and so one connected fracture may have unique orientations for each element. Ideally, we would have calculated the fracture orientations in the experiments with this method also, but this is not feasible. However, the shape anisotropy of the fractures in the experiments is rather high, and so there are not many highly tortuous fractures that would produce varying orientations if we measured the orientation segment by segment, every 10 microns, for example. We now mention these points in the manuscript (C#41).

Detailed comments/questions

42. Lines 49 – 50: A pre-existing common kinematic (“strain”) term in the literature is “extension fracture”(Marrett and Peacock, 1999) rather than “extensile fracture”, as proposed here. That said, I personally tend to agree with the measured critical response of Pollard (2000) to the paper of Marrett and Peacock (1999), who prefers the nomenclature related to mode of propagation.

We follow the approach of Paterson and Wong (2005) for the terminology, as described in the manuscript (C#42).

43. Lines 127 – 131: I am not sure if alternating step-overs would form in nature. The initial Griffith cracks might be randomly distributed, certainly grow at different times, and would have varying initial lengths. Once two cracks link up, one obtains a larger stepped crack. Such a crack would propagate more easily than its smaller unlinked siblings because of the inverse relationship between crack length and critical stress for propagation. In addition, due its stepped geometry, the crack-tip stress fields would become asymmetric. In short, there could be some dynamic effects that lead to a preferred step-over direction.

We agree with this idea, and now mention it in the manuscript (C#43).

44. Line 167: The term “unparalleled” is rather sensational.

Although it is sensational, it is true: with X-ray tomography we can observe fracture development directly in situ, unlike previous methods.

45. Line 169: The expression “partitioning of the local strain tensors” appears a bit awkward semantically. Better perhaps: local deformation is partitioned. An individual strain tensor quantifies locally homogeneous deformation and is not in itself “partitioned”.

We have modified the sentence for clarity.

46. Line 173: This line states “previous work has not been able to directly compare the contribution of shear and extensile deformation to the system”. However, the authors employed DVC in their own previous paper and report on contraction/dilation/shear strain. Why is that not “a direct comparison of shear and extensile deformation”?

Please see the response to comment #32 above.

47. Line 188: The adverb “explicitly” is not needed.

Corrected as suggested.

48. Line 197: I recommend replacing “unparalleled”.

Corrected as suggested.

49. Lines 231 – 232: What is the width of the median filter?

The size of the median filter is 5 data points. We now include this information in the manuscript (C#49).

50. Lines 251 – 254: This has all been said at the beginning of the paragraph and should thus be removed.

Here, we wish to emphasize that despite the minor difference between experiment WG10 and the other experiments described here (C#50), for all of the experiments, the volume of extensile fractures exceeds the volume of shear fractures for the majority of the accumulated time.

51. Line 264: Replace “during loading” with “before peak stress”. After peak stress, loading continues for a while, after all.

We have reworded this sentence accordingly.

52. Line 288: The parameter “g” for grid geometry is not well explained, here or in the supplement. Why are numbers used rather than clear statements such as “left-dipping”, “right-dipping”, “alternating dip”, etc. What would $g = 3$ or 5 denote?

We have added an additional sentence to clarify the meaning of g.

53. Line 334 ff: The authors name material parameters without sufficient explanation. Are these the mechanical properties of the intact matrix rock or the pre-existing fractures themselves? It needs to be stated in the main text.

We have modified this sentence to clarify that these parameters are those of the intact rock.

54. Line 356 ff: The authors state:” The high temporal resolution provided by the models shows that the fracture networks experience some temporary phases in which the

ratio of the extensile/shear fracture length decreases.” This is also the case in the synchrotron data. Why?

In both the models and the experiments, the variations in the ratio arise from the influence of fracture propagation and coalescence. We have modified this sentence to describe this point more clearly (C#54).

55. Line 426: The authors do not present any observations of randomly oriented cracks at early deformation stages. Hence, the statement “consistent with our observations” seems incorrect. As argued above, their tomograms are too coarse to resolve pre-existing and early cracks.

We have reworded this sentence for clarity (C#55). We also now include images of the fracture networks at earlier stages of loading when the shear fractures dominate. Please see the responses to comments #1 and #2 for details.

Based on the average volume of the fractures identified in the early stages of loading in each experiment, we are confident that these data and methods reveal a significant volume fracture of the total fracture network, and that the key conclusions are not biased by the minor influence of the smallest fractures. In addition, please see the responses to comments #33 and #34 above.

56. Line 427 ff: The authors state:” The evolution observed in the present analysis agrees with microstructural observations that suggest that after about 75% of the maximum differential stress, new transgranular fractures develop at high angles to interfaces of different minerals and generally subparallel to the σ_1 direction [22].” This is not how I interpret the data in Fig. 3. In the interval 75 to 85%, only 1 experiment shows, more or less, a monotonic increase in the proportion of extensile fractures. The authors also do not report any micrographs and hence cannot ascertain if the new fractures are transgranular cracks or not.

This sentence refers to the development of transgranular fractures identified in previous work, which occur after above 75% of the maximum differential stress (ref. 22), and not in the current work. We now include new figures that show how the fractures interact with the microstructure of the rock. Please see the responses to comments #1 and #2 for details.

In both the previous work and the present work, a higher volume or number of extensile fractures develops when the differential stress >75%. We agree that the ratio of the extensile/shear fracture volume becomes greater than one generally after 90% of the differential stress at failure. So only the chosen threshold of the differential stress changes between the present and previous work.

57. Lines 469 ff: “Constraining the geometry of first few fractures that fail during triaxial compression using energy optimization indicates that these fractures will not be aligned sub-parallel to the σ_1 direction, but instead at 30-35° from the σ_1 direction (Figure 4), consistent with the higher proportion of shear fractures observed early in loading in the experiments (Figure 3).” I disagree with this statement because the numerical model does not model fracture nucleation (= the “first few fractures that fail”) and thus does not show that energy optimisation would produce this geometry. In real rocks, the available evidence shows that the “first few” fractures are entirely controlled by the complex heterogeneous microstructure of natural rock.

We agree and have reworded this sentence in order to emphasize that this step of the modelling analysis aims to capture the development of the first few mm-scale fractures, rather than the nucleation of fractures (C#57).

58. Lines 571: “Consequently, we follow the approach of decades of previous work that have used the orientation of fractures relative to the global directions of the principal stresses to compare the relative proportions of extensile and shear deformation”. Just because something has been done for a long time, does not mean that it is correct/applicable. As stated above, the mode of propagation is strictly defined via the displacement continuity at the fracture tip – a purely kinematic quantity that can, in principle, be derived from the velocity fields obtained with DIC. The definition is not associated with stress-tensor orientation.

Indeed, the displacement discontinuity on the fracture define whether it may be classified as mode I or mode II/III. With digital volume correlation (DVC), it is possible to calculate the three dimensional strain tensor at locations throughout a rock volume. From the strain tensor, we may calculate the volumetric strain using the first invariant of the tensor, I_1 , and quantify the shear strain using the second invariant of the deviator tensor, for example, J_2 . Although it is possible to compare the proportion of contractive and dilative volumetric strain with these data, it is not possible to compare the relative importance of dilative and shear strain in a given rock volume, or near a fracture. While comparing the contractive and dilative strains is similar to comparing apples of different sizes, comparing the dilative and shear strains is more similar to comparing apples and oranges. We describe this point in Section 1.2 (C#58). Moreover, the spatial resolution of the DVC calculation must be generally 10 to 20 times smaller than the resolution of the scan in order to perform a reasonable correlation between tomograms. And so in practice it is difficult, if not impossible, to define the displacement discontinuity on a fracture from the strain tensor measured with DVC.

59. Line 596: You probably mean σ_2 here, rather than σ_1 .

With the sentence “perpendicular to the σ_1 direction (horizontal) and 4 mm parallel to the σ_1 direction (vertical)” we indicate that the horizontal direction is perpendicular to the σ_1 direction, and the vertical direction is parallel to the σ_1 direction. So we intended to use σ_1 here.

60. Line 650ff: “The two-dimensional system prevents the development of polymodal faulting [69].” That reminds me: it would be useful to plot the 3D fracture orientation distribution. Do you obtain radially symmetrical fracture orientations? This adds additional complexity to interpreting the data in terms of Mohr-Coulomb plasticity.

We do not generally observe the polymodal faulting, or conjugate fault structures, as described in previous work in these experiments. The fracture geometries shown in Figure 2 are more representative of the fracture networks that develop in these experiments, rather than conjugate, system-spanning shear fractures.

61. Line 651ff: “The low strain simplification of the linear elastic models prevents simulating the full loading of a triaxial compression experiment, and thus requires simulating snapshots of the experiment at particular stress steps.” From the model snapshots shown in the paper, it appears that these models contain fractures that propagated through the whole model. At the fracture tips of shear fractures, the low-strain assumptions most certainly breaks down. The stresses there will be greatly overestimated (see Schrank et al. (2017)) and thus lead to unrealistic propagation timeframes.

Indeed, we agree that these models do not accurately capture aspects of the timing of fracture propagation. The conclusions derived from the models about the fracture geometries do not depend on how quickly the fractures propagate in the models.

62. Fig. S9: There are dots with error-bar-like features. It is not explained what they represent (supposedly some characteristic moments of some frequency distribution).

We have now clarified in the caption of this figure that “The plots show the mean $\pm$ one standard deviation of the shape anisotropy of the extensile and shear fractures at each loading step in each experiment.”

References

- Fattaruso, L., Cooke, M., McBeck, J., 2022. The Influence of Fracture Growth and Coalescence on the Energy Budget Leading to Failure. *Frontiers in Earth Science* 10.
- Hooker, J.N., Laubach, S.E., Marrett, R., 2014. A universal power-law scaling exponent for fracture apertures in sandstones. *GSA Bulletin* 126, 1340-1362.
- Liu, J., Regenauer-Lieb, K., 2011. Application of percolation theory to microtomography of structured media: Percolation threshold, critical exponents, and upscaling. *Physical Review E* 83, 016106.
- Marrett, R., Peacock, D.C.P., 1999. Strain and stress. *Journal of Structural Geology* 21, 1057-1063.
- McBeck, J.A., Cordonnier, B., Renard, F., 2021. The influence of spatial resolution and noise on fracture network properties calculated from X-ray microtomography data. *International Journal of Rock Mechanics and Mining Sciences* 147, 104922.
- Nasseri, M., Tatone, B., Grasselli, G., Young, R., 2009a. Fracture Toughness and Fracture Roughness Interrelationship in Thermally treated Westerly Granite. *Pure and Applied Geophysics* 166, 801-822.
- Nasseri, M.H.B., Schubnel, A., Benson, P.M., Young, R.P., 2009b. Common evolution of mechanical and transport properties in thermally cracked westerly granite at elevated hydrostatic pressure. *Pure and Applied Geophysics* 166, 927-948.
- Nasseri, M.H.B., Schubnel, A., Young, R.P., 2007. Coupled evolutions of fracture toughness and elastic wave velocities at high crack density in thermally treated Westerly granite. *International Journal of Rock Mechanics and Mining Sciences* 44, 601-616.
- Pollard, D.D., 2000. Strain and stress: Discussion. *Journal of Structural Geology* 22, 1359-1367.
- Raghami, E., Schrank, C., Kruhl, J.H., 2019. 3D modelling of the effect of thermal-elastic stress on grain-boundary opening in quartz grain aggregates. *Tectonophysics*, 228242.
- Schrank, C.E., Fousseis, F., Karrech, A., Regenauer-Lieb, K., 2012. Thermal-elastic stresses and the criticality of the continental crust. *Geochem. Geophys. Geosyst.* 13, Q09005.
- Schrank, C.E., Karrech, A., Boutelier, D.A., Regenauer-Lieb, K., 2017. A comparative study of Maxwell viscoelasticity at large strains and rotations. *Geophysical Journal International*.
- Tapponnier, P., Brace, W.F., 1976. Development of stress-induced microcracks in Westerly Granite. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts* 13, 103- 112.

17th Aug 23

Dear Dr McBeck,

Your manuscript titled "Deformation evolves from shear to extensile in rocks due to energy optimization" has now been seen by our reviewers, whose comments appear below. In light of their advice we are delighted to say that we are happy, in principle, to publish a suitably revised version in Communications Earth & Environment under the open access CC BY license (Creative Commons Attribution v4.0 International License).

We therefore invite you to revise your paper one last time to address the remaining concerns of our reviewers. At the same time we ask that you edit your manuscript to comply with our format requirements and to maximise the accessibility and therefore the impact of your work.

EDITORIAL REQUESTS:

Please review our specific editorial comments and requests regarding your manuscript in the attached "Editorial Requests Table".

*****Please take care to match our formatting and policy requirements. We will check revised manuscript and return manuscripts that do not comply. Such requests will lead to delays. *****

Please outline your response to each request in the right hand column. Please upload the completed table with your manuscript files as a Related Manuscript file.

If you have any questions or concerns about any of our requests, please do not hesitate to contact me.

SUBMISSION INFORMATION:

In order to accept your paper, we require the files listed at the end of the Editorial Requests Table; the list of required files is also available at <https://www.nature.com/documents/commsj-file-checklist.pdf> .

OPEN ACCESS:

Communications Earth & Environment is a fully open access journal. Articles are made freely accessible on publication under a [CC BY license](http://creativecommons.org/licenses/by/4.0) (Creative Commons Attribution 4.0 International License). This license allows maximum dissemination and re-use of open access materials and is preferred by many research funding bodies.

For further information about article processing charges, open access funding, and advice and support from Nature Research, please visit <https://www.nature.com/commsenv/article-processing-charges>

At acceptance, you will be provided with instructions for completing this CC BY license on behalf of all authors. This grants us the necessary permissions to publish your paper. Additionally, you will be asked to declare that all required third party permissions have been obtained, and to provide billing information in order to pay the article-processing charge (APC).

Please use the following link to submit the above items:

[Link redacted]

** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

We hope to hear from you within two weeks; please let us know if you need more time.

Best regards,

Joe Aslin

Senior Editor,
Communications Earth & Environment
<https://www.nature.com/commsenv/>
Twitter: @CommsEarth

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

See PDF.

Reviewer #2 (Remarks to the Author):

Dear colleagues,

please, find all comments in the attached PDF.

Many thanks again for the opportunity to study this work.

Best wishes
Chris

**Second review of 'Deformation evolves from shear to tensile in rocks due to energy optimization'
by McBeck et al.**

Dear Dr. McBeck,

Thank you for the careful consideration of my comments and clarification of my queries, and the associated adjustments to the manuscript. I have two further comments:

1. C#04 – 'This numerical result is consistent with the experiments; all of the six or five of six (for experiment WG10) most volumetric fractures identified in the experiments when $\sigma_H = 45\text{MPa}$ have orientations within the range expected for shear.' I suggest including a table in the supplementary information showing the measured orientations of the six most volumetric fractures in each experiment.
2. I noticed when going through the revised manuscript that the statement 'using different combinations of orientation ranges does not change the general trend of the curves for each experiment' in the caption for Figure S4 is not entirely accurate. Yes, they all show an increasing volume of extensile fractures toward failure, but they do not all show a transition from shear to extensile deformation. The range 23-32 degrees (based on friction coefficient values 0.5-1.0) shows rather a competition between shear and extensile deformation in four out of the six cases, and predominantly extensile deformation in the other two. The other three ranges do show a transition from shear to extensile, but two of the ranges (<12, 17-32 and <12 23-32) bias the results towards a shear to extensile evolution by ignoring potential extensile fractures in the range 12-17 and 12-23 degrees. It is not clear why 12 degrees was used for the tensile crack threshold here, nor why the intermediate orientations were ignored – some clarification on this would benefit the paper. Given that the sentence in lines 221-223 states 'previous work suggests that the maximum internal friction coefficient may be 1.0 rather than 1.5', that the models do not use friction coefficients larger than 1.0, and that experimentally-obtained friction coefficient values for granites are in the range 0.5-0.7 while the representative value for a range of rock types is 0.85 from Byerlee's Law, my question is then, why present experimental results using the range 0.5-1.5 (shear crack orientations of 17-32 degrees) in the main manuscript instead of the more realistic range of 0.5-1.0 (shear crack orientations of 23-32 degrees)? The differences in how the volume proportion evolves between these two ranges, as shown in Figure S4 affects the reliability of the main claim of the paper that deformation evolves from shear to extensile. A sentence or two addressing why the wider range was used would help to strengthen the paper.

Best,

Alexis

Comments on revised manuscript and rebuttal

Dear editor, dear authors,

I now had a chance to study the revised manuscript and rebuttal letter and am happy to note that the authors worked hard to address both reviewers' comments.

I feel that the methodology is much better explained now, enabling reproducibility, and studying uncertainty related to the thresholding of the tomograms by varying thresholds systematically is very useful. Since I have been given the tempting chance, I cannot refrain from adding a couple of comments on the rebuttal and changes for the authors' and editor's kind consideration:

1) Fig. S9 shows nicely by how much the volume ratios depend on the segmentation choice. The authors state that the overall outcome does not seem to be affected by this uncertainty. However, one can note that the associated uncertainty is large enough to change the volume ratio from <1 to >1 for at least two experiments (WG10, WG14) at failure. Isn't that important to note? Moreover, it would be nice to see the empirical upper-bound uncertainties from Fig. S9 as error bars in Fig. 3. Are stress errors now also mentioned in the paper?

2) The authors replied to my suggestion to use their 3D DVC capabilities for a better mode characterisation of the fractures: " [...] it is not possible to compare the relative importance of dilative and shear strain in a given rock volume, or near a fracture.

This is because comparing the dilative and shear strains is similar to comparing apples and oranges. We can find that the dilation is a certain value, and the strain is a certain value at a particular point, but this comparison does not indicate if a rock volume is dominated by dilation or shear strain."

I do not agree with this statement. Fig. 1 illustrates graphically that one can simply use the displacement-gradient vectors adjacent to fractures (rather than full strain tensors) to characterize the presence/absence and nature of a displacement discontinuity across any fractures (as per the proper definition). For each fracture one could thus record volume and any displacement gradient across it (DVC would work really nicely near fractures because of the large X-ray attenuation gradient with the matrix) as function of load and thus apply the true definition for fracture mode of propagation. The newly added tomogram slices highlight the well-known important control of microstructure on fracture nucleation, which could have – as argued in the original reviews of both reviewers – an important effect on the results, including potential misclassification of fractures.

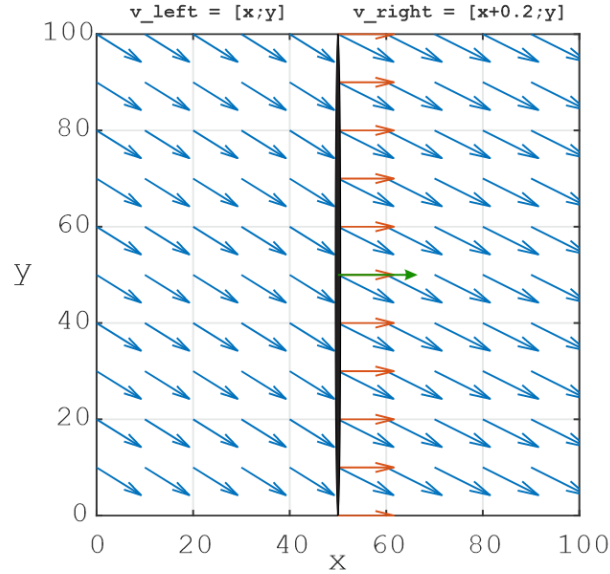


Fig. 1: Displacement field (blue arrows, scaling arbitrary but consistent) around a hypothetical fracture (black ellipse centred on $x = 50$). On the right side of the fracture, the displacement vectors have a slightly larger x -component, leading to a non-zero displacement gradient across the fracture (brown arrows). If one compares the orientation of the displacement gradient vector to the unit-normal vector (green arrows) near the fracture tip, one can easily apply Pollard's astute definition of modes I and II/III for fracture characterisation. In the example here, the displacement gradient is perfectly parallel to the crack's unit-normal vector, denoting a mode-I crack. Note that this approach lends itself to a very efficient computational implementation because one can focus on the regions immediately around fractures only. One could then even study in which mode fractures nucleated and when and which fractures eventually changed mode of propagation! What wealth of new insights that would potentially give!

Of course, none of these issues are proverbial "deal breakers", and I gladly leave the decision on potential additional changes of the paper in the editor's and authors' expert hands. I believe this work will make a valuable contribution to the literature.

Best wishes

Christoph Schrank

Dear Dr. Schrank, and Dr. Cartwright-Taylor,

Thanks for these helpful suggestions. We have clarified these remaining points. We respond below in bold font, and indicate where we have changed the manuscript using red font and the comment function of Word, where the change following comment #1 is indicated with C#01.

Best,
Jess McBeck

Second review of 'Deformation evolves from shear to tensile in rocks due to energy optimization' by McBeck et al.

Dear Dr. McBeck,

Thank you for the careful consideration of my comments and clarification of my queries, and the associated adjustments to the manuscript. I have two further comments:

1. C#04 – 'This numerical result is consistent with the experiments; all of the six or five of six (for experiment WG10) most volumetric fractures identified in the experiments when $\sigma = 45\text{MPa}$ have orientations within the range expected for shear.' I suggest including a table in the supplementary information showing the measured orientations of the six most volumetric fractures in each experiment.

We have included this table in the supporting information, and refer to it in the main manuscript (C#01).

2. I noticed when going through the revised manuscript that the statement 'using different combinations of orientation ranges does not change the general trend of the curves for each experiment' in the caption for Figure S4 is not entirely accurate. Yes, they all show an increasing volume of extensile fractures toward failure, but they do not all show a transition from shear to extensile deformation. The range 23-32 degrees (based on friction coefficient values 0.5-1.0) shows rather a competition between shear and extensile deformation in four out of the six cases, and predominantly extensile deformation in the other two. The other three ranges do show a transition from shear to extensile, but two of the ranges (<12, 17-32 and <12 23-32) bias the results towards a shear to extensile evolution by ignoring potential extensile fractures in the range 12-17 and 12-23 degrees. It is not clear why 12 degrees was used for the tensile crack threshold here, nor why the intermediate orientations were ignored – some clarification on this would benefit the paper. Given that the sentence in lines 221-223 states 'previous work suggests that the maximum internal friction coefficient may be 1.0 rather than 1.5', that the models do not use friction coefficients larger than 1.0, and that experimentally-obtained friction coefficient values for granites are in the range 0.5-0.7 while the representative value for a range of rock types is 0.85 from Byerlee's Law, my question is then, why present experimental results using the range 0.5-1.5 (shear crack orientations of 17-32 degrees) in the main manuscript instead of the more realistic range of 0.5-1.0 (shear crack orientations of 23-32 degrees)? The differences in how the volume proportion evolves between these two ranges, as shown in Figure S4 affects the reliability of the main claim of the paper that deformation evolves from shear to extensile. A sentence or two addressing why the wider range was used would help to strengthen the paper.

Best, Alexis

We chose the threshold for the extensile orientation following the ranges of internal friction coefficient and resulting shear orientations (17 or 23°), and using a more vertical orientation that reflects a more conservative estimate of the volume of extensile fractures (12°). We now state this point explicitly in the caption of Figure S4.

We did not focus on the intermediate orientations for the combinations of orientation ranges with a gap between the extensile and shear orientations in order to ease comparison between the fractures dominated by extensile deformation and dominated by shear deformation, as this is the focus of the manuscript.

There is a disagreement in the literature about the maximum limit of the internal friction coefficient (1 vs 1.5). Consequently, we tested ranges the two selected ranges of shear orientations. We describe this choice in the manuscript (C#02).

We agree that using the different ranges produces minor changes in the precise value of the ratio of the extensile and shear fractures for some of the experiments. However, we maintain that the central conclusions remain unchanged, and we describe the minor differences in the caption of Figure S4: “Using different combinations of orientation ranges does not change the general trend of the curves for each experiment, except for the experiment WG20_19 and the angle ranges of #2. In the other experiments, all of the combinations produce a similar trend with an increasing volume of extensile fractures toward failure (WG19_18, WG06_02, WG18_14, WG14), or extensile-dominated deformation throughout loading (WG10). For the remaining experiment (WG20_19), using the orientation ranges of $\theta_e=0-23^\circ$ and $\theta_s=23-32^\circ$ produces a higher volume of extensile fractures than shear fractures throughout loading than the other orientation ranges. However, for all of the combinations, the volume of extensile fractures increases toward failure after about 97% of the σ_D at macroscopic failure. The consistent trends of r with increasing σ_D for the four different combinations lend confidence to the results described in the main manuscript using $\theta_e=0-17^\circ$ and $\theta_s=17-32^\circ$.”

Reviewer #2: Comments on revised manuscript and rebuttal

Dear editor, dear authors,

I now had a chance to study the revised manuscript and rebuttal letter and am happy to note that the authors worked hard to address both reviewers' comments.

I feel that the methodology is much better explained now, enabling reproducibility, and studying uncertainty related to the thresholding of the tomograms by varying thresholds systematically is very useful. Since I have been given the tempting chance, I cannot refrain from adding a couple of comments on the rebuttal and changes for the authors' and editor's kind consideration:

3) Fig. S9 shows nicely by how much the volume ratios depend on the segmentation choice. The authors state that the overall outcome does not seem to be affected by this uncertainty. However, one can note that the associated uncertainty is large enough to change the volume ratio from <1 to >1 for at least two experiments (WG10, WG14) at failure. Isn't that important to note? Moreover, it would be nice to see the empirical upper-bound uncertainties from Fig. S9 as error bars in Fig. 3. Are stress errors now also mentioned in the paper?

We now state that the stress errors are smaller than the size of the markers in Figure S1 in the caption.

We agree that varying the segmentation threshold does change the volume ratio from below to above one for one experiment of the six (WG14). However, for experiment WG10, and all the other experiments, the difference between the volume ratio at the beginning of experiment and immediately preceding failure does not change. As described in the caption of Figure S9, for all of the experiments: “the relative volume proportion of the extensile and shear fractures remains unchanged as the threshold changes. In other words, the light blue line remains above the dark blue line (or vice versa) for each experiment. Consequently, we have a high degree of confidence in the experimental results.”

4) The authors replied to my suggestion to use their 3D DVC capabilities for a better mode characterisation of the fractures:” [...] it is not possible to compare the relative importance of dilative and shear strain in a given rock volume, or near a fracture. This is because comparing the dilative and shear strains is similar to comparing apples and oranges. We can find that the dilation is a certain value, and the strain is a certain value at a particular point, but this comparison does not indicate if a rock volume is dominated by dilation or shear strain.”

I do not agree with this statement. Fig. 1 illustrates graphically that one can simply use the displacement-gradient vectors adjacent to fractures (rather than full strain tensors) to characterize the presence/absence and nature of a displacement discontinuity across any fractures (as per the proper definition). For each fracture one could thus record volume and any displacement gradient across it (DVC would work really nicely near fractures because of the large X-ray attenuation gradient with the matrix) as function of load and thus apply the true definition for fracture mode of propagation. The newly added tomogram slices highlight the well-known important control of microstructure on fracture nucleation, which could have – as argued in the original reviews of both reviewers – an important effect on the results, including potential misclassification of fractures.

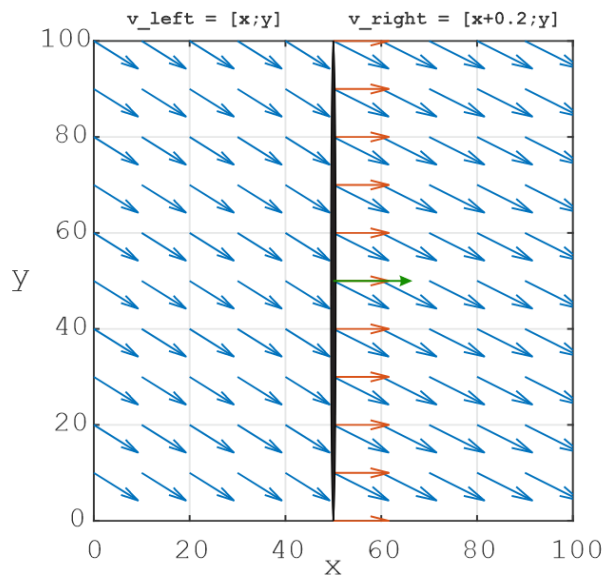


Fig. 1: Displacement field (blue arrows, scaling arbitrary but consistent) around a hypothetical fracture (black ellipse centred on $x = 50$). On the right side of the fracture, the displacement vectors have a slightly larger x-component, leading to a non-zero displacement gradient across the fracture (brown arrows). If one compares the orientation of the displacement gradient vector to the unit- normal vector (green arrows) near the fracture tip, one can easily apply Pollard’s astute definition of modes I and II/III for fracture characterisation. In the example here, the displacement gradient is perfectly parallel to the crack’s unit-normal vector, denoting a mode-I crack. Note that this approach lends itself to a very efficient computational implementation because one can focus on the regions

immediately around fractures only. One could then even study in which mode fractures nucleated and when and which fractures eventually changed mode of propagation! What wealth of new insights that would potentially give!

Of course, none of these issues are proverbial “deal breakers”, and I gladly leave the decision on potential additional changes of the paper in the editor’s and authors’ expert hands. I believe this work will make a valuable contribution to the literature.

Best wishes Christoph Schrank

We agree that this approach could be beneficial to future work. In practice, the implementation could be difficult because of potential differences between the spatial resolution of the tomogram and the DVC calculation. Performing this calculation may provide useful information about the evolving dominance of extensile and shear failure, however, it is outside the scope of the present work.