Authors' Comments

We would like to thank the editor, reviewers, and community members for their commentary on our manuscript, "Chemical abrasion: The mechanics of zircon dissolution" that will ultimately help to strengthen this contribution. We respond to each of their comments below. If given the opportunity to submit a revised version of this manuscript, some of the major changes we would make in response to feedback include:

1) Focus on the heart of the manuscript more strongly – *textural evidence* for the mechanics of zircon dissolution. Many comments reference the lack of geochemical and geochronological data. We would like to emphasize that a complimentary manuscript that focuses on the geochemical and geochronological evolution of chemically abraded samples is currently in preparation. We would shorten and refocus Sections 4.2 (Implications for ID-TIMS U-Pb Geochronology) and Section 5 (Conclusion) to emphasize that the effectiveness of any chemical abrasion protocol for ID-TIMS U-Pb geochronology will ultimately be sample-dependent and reflect a sample's radiation damage and inclusion content and distribution. We would refrain from prescribing any specific chemical abrasion protocol, since no geochronological and geochemical data are presented in the current work.
2) We would also remove Section 4.3 (Implications for radiation damage annealing models) since it is tangential to the discussion and in need of additional supporting data.
3) Streamline the writing to eliminate wordy text and shorten the manuscript length. We would add two small tables that more succinctly summarize Raman data and basic sample descriptions.

We address each reviewer's specific comments below. Reviewer comments are in black text, and our responses are in blue text.

Reviewer #2 – Nasdala Lutz

The main shortcoming of the present research is perhaps that there is no chemical information on the samples studied. The present paper presents results of rather sophisticated studies including micro-CT and precise dimension estimates, but the most basic information remains hidden. There are quite a few journals that would refuse to publish a study on samples whose chemical compositions are unknown; for good reasons, as I think.

In particular, there are no U and Th concentrations available that, along with U–Pb ages, could have been used to calculate realistic time-integrated alpha doses. Instead, authors fiddle with equation #2 of Váczi and Nasdala (2017) in attempting to estimate self-irradiation doses from Raman FWHMs. Which is based on a serious fallacy of thinking: The equation of Váczi and Nasdala (2017) – that by the way has never been proposed as a "calibration", as present authors claim in line 223 – refers solely to Sri Lankan gem zircon with its particular damage-annealing history, whereas other zircon populations may have appreciably different FWHM-dose relationships.

In contrast, comparing alpha doses that were calculated from U, Th and age, with the observed FWHM values, would have opened up valuable opportunities for characterising the initial samples and their annealing histories. This chance was missed. Also, because of the unavailability of chemical compositions, it cannot be evaluated whether HF etching has removed completely some interiors of the grains but has left the rest fully unchanged, or did un-dissolved remnants experience some chemical leaching?

We appreciate the reviewer's concern regarding the absence of geochemical data for these samples. Our plan is to pair this microstructural investigation with a forthcoming study (the manuscript is currently in preparation) that presents geochemical and geochronological data for three of the zircon samples studied. This current contribution is very long (19 figures plus supplementary tables and videos); efforts to include both microtextural and geochemical data in a single contribution would produce a paper of unreasonable length. As such, the dataset has been split. In the second forthcoming contribution, we will compare alpha dose estimates based on zircon age and actinide concentrations with the Raman-derived alpha dose estimates presented here as suggested by the reviewer; the opportunity is delayed rather than missed. Further, because alpha dose estimates based on U and Th concentrations and U-Pb crystallization ages do not consider radiation damage annealing, they are not necessarily accurate representations of present-day damage in the zircon samples analyzed. Directly measuring the present-day radiation damage densities with Raman is more relevant to the present solubility experiments.

We intend only to use the v_3SiO_4-alpha dose relationship presented by Váczi and Nasdala (2017) for Sri Lankan zircon to broadly characterize damage in these samples as low-to-intermediate (KR18-04 and BOM2A) and intermediate-to-high (SAM-47 and AS3) and to demonstrate approximate magnitudes of inter- and intra-crystalline variations in damage. The v_3SiO_4-alpha dose relationship presented by Palenik et al. (2003) (their Fig. 5) for Sri Lankan zircon (which assumes an equivalent damage accumulation interval for the Sri Lankan zircon of 375 Ma to correct for geological

annealing) nicely fits the dataset of unannealed zircon presented by Nasdala et al. (2001), suggesting that the relationship is broadly appropriate for zircon from a wide range of geological environments. We can add caveats in the revisions that better clarifies that the derived alpha dose estimates are equivalent alpha doses based specifically on the $v_3 SiO_4$-alpha dose relationship established for Sri Lankan zircon. However, since we do not establish a firm alpha dose–dissolution mathematical model or the like, we do not think that using Raman-derived alpha dose damage equivalents is inappropriate.

Present authors rely on the results of Palenik et al. (2003) and Váczi and Nasdala (2017) who both have claimed that the FWHM of the $n_3(SiO_4)$ Raman band of Sri Lankan zircon has a maximum ("saturation") value of about 35 cm^{-1}. On the other hand, sample AS3 yielded FWHMs of about 48 cm^{-1} (Fig. 5a, Table S1). This apparent contradiction should be discussed, and actually such extremely broadened Raman spectra should be shown (at least in a supplementary figure). It has been suspected that Sri Lankan zircon of elevated degrees of radiation damage have experienced preferred annealing, which might explain the "saturation" as a particular feature of Sri Lankan zircon. In conclusion, it is not surprising that other zircon, not affected by the particular "Sri Lankan" annealing history, may indeed show further band broadening. This has been rarely observed thus far but is valuable; so it should been shown and discussed.

The reviewer is correct that previous studies have shown that the maximum $v_3(SiO_4)$ width is ~35 cm^{-1}. The maximum FWHM value recorded for the $v_3(SiO_4)$ band for AS3 is ~35 cm^{-1}, consistent with previous studies. The 48 cm^{-1} width that the reviewer cites is for the E_g band and not v_3. As demonstrated in Fig. 5, Fig. 6, and in previous studies, the E_g band is always broader than the v_3 band (Härtel et al., 2020; Anderson et al., 2017).

Line 183, it is not really of relevance for the reader to learn about the equipment of the Princeton Raman system; instead, it should merely be stated which particular laser was actually used in the present study. Presumably red, as the green laser tends to induce Er^{3+}-related photoluminescence obscuring the Raman spectrum? The laser power at the sample surface (so not laser output but power measured behind the objective) needs to be reported.

We can rephrase this line in the revised text to better clarify. Both red and green laser wavelengths are reported because both lasers were used for Raman measurements. Most measurements were made using the red laser, since the green laser often produced fluorescent artifacts as the reviewer suggests. The laser power to the samples surface was 8.5 to 17 mW and ~175 to to 60 mW for the red and green lasers, respectively.

- Line 185: I wished Horiba would stop implementing the main silicon band in their auto-calibration procedure. However, if authors believe the friendly Horiba service person who did so was right, they should provide a published reference for the 520.7 cm^{-1} value they used for calibration. Other authors have used 520.5 or 521.0 cm^{-1}, and none of these values has ever been appropriately supported. Note that the oriented Si wafer is provided for alignment purposes but pretty unsuitable for wavenumber calibration: Visible light penetrates only a few µm into Si, which is why increasing the laser power, or changing the laser colour, changes absorption-induced heating and causes a shift of the Raman band. Why not using a truly known signal instead? The Rayleigh line has a Raman shift of 0 cm^{-1}; a perfect calibration means.

We used the Si 520.7 cm^{-1} band for instrument calibration because as the reviewer points out it is what the instrument manufacturer recommends, and it is the calibration standard supplied by Horiba for the LabSpec6 software's autocalibration function. Itoh and Shirono (2020, Journal of Raman Spectroscopy) demonstrated that a Si wafer has a Raman peak position of 520.45 ± 0.28 cm^{-1} which is within uncertainty of the value recommended by Horiba. We also monitored the accuracy of measured peak positions daily by measuring a reference spectrum for quartz (Krishnam, 1945 Nature). We appreciate the reviewer's suggestion about using the Rayleigh line, and we will bear this in mind for future Raman investigations.

- Line 189–190, quoting ~2 µm spatial resolution upon using a "400 to 100 µm confocal pin hole" is too general and imprecise. With the confocal hole set to 400 µm, the spatial resolution will be much poorer than 2 µm lateral, and with a 100 µm confocal hole, the LabRAM Evolution should – provided the beam alignment is well done – reach its maximum performance close to the Rayleigh or Sparrow criterions, well below 1 µm (see detailed study by Kim et al., 2020, Current Applied Physics 20:71–77). The above resolution estimate does of course only apply if the Olympus 100× / n.a. 0.90 objective was used (this essential information is missing as well).

The ~2 µm spatial resolution sited is meant to be generalized because measurements were made using a range of conditions including confocal pin hole size and objective lenses. Most measurements were made using an Olympus 100x/0.9na lens or Mitutoyo 50x or 20x long working distance objective lens. We can update our spatial resolution to <1 µm to ~5 µm to better reflect the range of spatial resolutions expected. We do not attempt to target or correlate our measurements across specific, fine-scale zones within a crystal. Rather, we use Raman measurements to broadly characterize the range of radiation damage densities in these samples, so precise spatial resolution is not critical.

- Page 8, caption of Fig. 3a: The wording "SEM images…" is too imprecise. State explicitly which images are CL, BSE, SE.

Each image in the figure has a symbol indicating whether the image is a CL or a BSE image. The figure includes a symbol legend. This was done to help shorten the text in the figure caption.

- Lines 218–219, the expression "Si–O asymmetric stretching band" is a contradiction in itself. To vibrate asymmetrically, it takes at least two bonds, whereas the stretching vibration of a single Si–O bond does not have any symmetry. Should be reworded to "asymmetric SiO_4 stretching".

The reviewer makes a valid point. We will make the suggested correction in revisions.

- Fig. 18, in subfigures a–d the main x-axis label is missing.

The x-axis labels should simply be "Leaching Temperature." We will add these labels in revisions.

- Lines 768–769, "…suggest that annealing or radiation damage is not simply the inverse of damage accumulation"; and again lines 891–892, "Raman analyses of annealed grains suggest that dry annealing is not the inverse of radiation damage accumulation". These statements are presented as if they were a result of the present study, which is decidedly not the case. In contrast, it has been emphasised may times already during the last 20 years that annealing does not inverse the damage accumulation process (mismatch of Raman FWHMs and Raman shifts, mismatch of unit-cell parameters a and c, etc.). Something like "our observations confirm again that…" would be more appropriate and more honest…. - Same in lines 250–252, "We note that none of the samples have achieved complete structural recovery after annealing at 900 °C for 48 h, since all measured peak widths are broader than that of synthetic zircon". This is not a new finding but merely another well-known fact, here merely reconfirmed again. People who wish to achieve near-complete annealing of zircon do this at 1400°C for four days or so.

Our intention was not to imply by that we were the first to draw these conclusions, but simply to describe the basic implications of the Raman data presented. We can rephrase these lines and add additional references as the reviewer suggests to more appropriately assign credit (e.g., Zhang et al., 2000; Geisler et al., 2002; Ginster et al. 2019; Huyskens et al., 2016).

By the way, Raman analyses do not suggest anything; only the results of Raman analyses suggest something (wording in line 891 should be improved).

We can rephrase this.

-Line 800, a typo? Instead of "verses" there should be "versus"…. - Line 983, the second author is Dr Chutimun (first name) Chanmuang N. (family name), so her name should be abbreviated "Chanmuang N., C."; same in line 1012. Similarly, in line 1050 the second author should be quoted as "Van den haute, P."

We will correct all the referenced typos.

- Table S1, the "empirical correction" of Raman shift values for band downshifts that are claimed to be caused by laser-induced heating remains fully vague (unclear). Something like this has never been observed and reported before; so at least some comprehensible documentation is needed. How about simply decreasing the laser-power density, to reduce laser-induced heating to a negligible extent??

We noticed that $v_3(SiO_4)$ and E_g peak positions were slightly down-shifted in carbon-coated samples compared to uncoated samples, so we tested how different laser-power densities affected Raman FWHM and Raman shift values in carbon-coated samples. We found a correlation between Raman shifts and laser power density. As such an empirically-derived correction was applied to affected Raman shift values measured at too high of laser-power densities. There was no correlation between FWHM values and laser power. We recommend that all Raman analyses be made prior to carbon coating and SEM imaging to avoid this problem. Unfortunately, the carbon coat could not be removed from our fragile samples. Because of the additional uncertainty involved with these Raman shift corrections, we do not base any of the conclusions presented in this study based on Raman shift values, nor do we present Raman shifts in any figures in the main text. We can add additional text to the footnote in Table S1 to clarify this.

- Table S1, in view of the fact that many authors present and discuss uncorrected FWHMs, applying FWHM correction in the present article is meritorious. However, a correction assuming a FWHM of the IPF (instrumental profile function) of 1.5 cm^{-1} (I assume this is the meaning of the unitless "1.5" in footnote c?) most likely has led to an overcorrection. The FWHM of the IPF of the LabRAM Evolution with its 800 mm focal length, in case the 1800 grooves per millimetre diffraction grating is used, is about 0.8 cm^{-1} in the red and about 1 cm^{-1} in the green range of the electromagnetic spectrum. Authors should test this – and hence the experimental band broadening of their system

– by measuring the width of the Rayleigh line, Kr lamp or similar emission lines, or even better narrow Raman bands of known widths (unstressed diamond, Ba nitride).

We derived the 1.5 cm^{-1} value used for band broadening corrections as cited in our footnote by measuring the width of the Rayleigh line as the author suggests. That is simply the value we measured, so we consider it to be the most appropriate estimate for experimental band broadening for our Raman system. We acknowledge that it is broader than , but we cannot offer a satisfactory explanation as to why.

Unfortunately I was unable to find Raman shifts and corrected FWHMs of synthetic zircon as measured by present authors. However, data points in Figs. 5 and 6 imply that synthetic zircon would have a FWHM of the $n_3(SiO_4)$ band of about 1 cm^{-1}, which is well below the real value of about 1.7–1.8 cm^{-1}. This seems to support my above suspicion of FWHM overcorrection. Let me assume authors have fitted an FWHM of 2.1 cm^{-1}. This value, corrected for a (too large) 1.5 cm^{-1} IPF, yields an overcorrected FWHM of 1.0 cm^{-1}. In contrast, a fitted FWHM of 2.1 cm^{-1}, corrected for a 0.8 cm^{-1} IPF, yields a quite realistic FWHM of 1.8 cm^{-1}. Anyway, data for synthetic zircon should be added to Table S1.

Measured (uncorrected) $v_3(SiO_4)$ and E_g FWHM values for synthetic zircon are 2.1 cm^{-1} and 4.0 cm^{-1}, respectively. Corrected for a 1.5 cm^{-1} bandpass, these values are 1.0 and 3.4 cm^{-1}, respectively. We can add these values to Table S1. It is possible that these values are overcorrected as the reviewer argues, however, 1.5 cm^{-1} is the measured bandpass for the instrument that we used.

- Table S1: Reporting FWHM values with two decimal digits fakes an unrealistic accuracy. As the uncertainty is 10% for narrow FWHMs, reporting 4.4 instead of 4.38 cm^{-1} will do the job.

We can change this in revisions.