Peer Review File

Manuscript Title: Asteroid Bennu's near-Earth lifetime is recorded by craters on its boulders

Redactions – Third Party Material

Parts of this Peer Review File have been redacted as indicated to remove third-party material.

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referees' comments:

Referee #1 (Remarks to the Author):

Review report of a manuscript entitled
"Asteroid Bennu's boulders record near-Earth history of impacts by millimeter- to centimeter-scale objects" by Ballouz et al., submitted to Nature

This study provides undoubtedly enduring results for centimeter-to-meter craters on an asteroid material. I would like to give my compliments to the authors. The original products are:

a. The morphology, the cross-sectional profile, and the ratio of depth to diameter of craters on boulders. All of these suggest that the boulders are porous.
b. The size-frequency distribution of craters in centimeter –to-meter scale. This gives a new information of the size-frequency distribution of impactor population.
c. The maximum crater diameter on a boulder of given diameter.

Based on these results,"impact strength" of boulders,impactor population in the near-Earth space, and the residence time scale of Bennu in the near-Earth space are discussed. These are all important and new contributions to our knowledge of the small body population and evolution in inner solar system. However, for these discussion, I have comments listed below. Although the current version of the manuscript is already very rich in contents and rather comprehensive in spite of the number of equations in methods, I hope that the paper is published after a revision with even richer and clearer descriptions, especially on (1) the ambiguity of the "impact strength" due to the ambiguity of the assumption of the weakest target radius (Rw), (2) the description of previous laboratory impact disruption study of carbonaceous meteorites, (3) the effect of oblique incidence in the main text (not at the end of a section of Methods), (4) a description of previous study of impactor population derived from seismic data on the Moon.

1. Robustness of the assumptions and parameters in this study

(1) On the assumption of Rw=80 m (the radius of Otohime boulder).

This is the crucial assumption in this study. If we chose larger Rw as in the previous studies, then the resultant "impact strength" becomes much larger.

Isn't there any C-complex fast-spinning object larger than Otohime boulder?

(2) On the derivation of μs by a curve fit to the maximum crater radii (Fig.2).

Due to the large uncertainty in the data of Rc/RT and small number of data points (five red triangles in Fig. 2b), we might only say that the value of μs allowed by the boulder data is about right (say, between 0.33 and 0.55) for porous target. Or can we say with more confidence? For example, if we only have 4 data points instead of the 5 in this study, can we still have the similar μs value?

(3) On the assumption of Rc=Rc, max at the impact of specific energy that is just below that required by the disruption threshold, I have only minor comments.

For a non-porous rock target (with spall fracture), the specific energy required for the largest crater and that for the disruption threshold is different in a few times.

Ref. Fig. 2 in Housen, Planet. Space Sci. 57, 142-153, 2009.

The assumption is maybe more suitable for a porous target or target without spall fracture (as boulders in this study).

Ref. Fig. 3 in Murakami et al., Planet Space Sci. 182, 104819, 2020.

2.Impact strength of boulders and cratering efficiency in the strength regime

The concept of impact strength is unfamiliar to most of readers, because it is not a widely recognized concept such as static tensile strength nor compressive strength. Below Eq. 2, it is described as the material strength that determines the diameter of crater. After Eq. M12, readers find that it probably denotes shear strength (or a kind of) in this study. More explanation of the "impact strength" in the main text may be of help for understanding. Moreover, using the relationship of Eq. M17, the cratering efficiency of the boulders on Bennu can be compared with various porous materials used in cratering experiments in the laboratory, for example, weak cemented basalt (WCB) has a μs value of 0.46 (similar to that of 0.47 in this study). Interestingly, the tensile strength of the material estimated using an empirical relationship was 0.45 MPa, roughly twice of that of the estimated tensile strength of the Ryugu boulder.

Ref. Table 4 in Housen and Holsapple, Icarus 211, 856-875, 2011.

3. Disruption threshold of boulders

The curves of basalt and pumice shown in Fig. 3 are not for normal impacts, but for oblique impacts. Therefore, a comparison between these curves and the result of this study needs discussion on the effect of impact angle on the disruption threshold. There are a few laboratory disruption experiments.

Ref. Yasui et al., Icarus 335, 113414, 2020.

Ref. Fig. 6 in Murakami et al., Planet Space Sci. 182, 104819, 2020 for extremely porous targets.

I'd like to suggest to refer to the result of impact disruption experiment of carbonaceous chondrites, Allende and Murchison in the literature. Because they are originally from C-complex objects and akin to the boulders on Bennu and Ryugu. Static strength was measured for these chondrites in the laboratory, therefore, showing the result (probably in Fig.3) will promote understanding.

Ref. Fig. 32 in Flynn et al., Chemie der Erde 78, 269-298, 2018, and references therein.

The catastrophic disruption threshold of Bennu boulders shown in Fig. 3a is weaker than basalt and pumice. Interestingly, however, it is very close to gypsum (with porosity of 65% which is more porous than the popular gypsum used for statue), which was shown to have a tensile strength of 0.3 MPa (similar to the estimated strength of a boulder on Ryugu). The μs value for the gypsum was 0.4 and close to the μs value (0.47) of the boulders of Bennu derived in this study, however, the ratio of depth to diameter of craters on gypsum targets are larger [36].

Ref. Eq.6 in Nakamura et al., Planet. Space Sci. 107, 45-52, 2015.

4. Size frequency distribution of the objects of millimeter-centimeter scale.
Apollo seismic data provided the information about the objects in this range.
Because lunar seismology (for the purpose of constraining the internal structure of the Moon) has long been an interest, the comparison of the present result with that of previous seismic data will be beneficial for the future planning of the seismology on the Moon.

Ref. Fig. 10.3 in “Impact Cratering: A Geologic Process” by Melosh, H. J., 1989.

An example of more recent analysis can be found here:

Ref. Fig. 2 in Kawamura et al., J. Geophys. Res. 38, L15201, 2011.

Specific comments

Line 29: "Ryugu revealed an unexpectedly low tensile strength [3]"
I'm afraid that this sentence is misleading. The thermophysical analysis directly revealed "high porosity" of a boulder on the surface of Ryugu. High porosity does not necessarily mean low tensile strength (imagine a metal honeycomb structure).

Lines 52-53: "Bennu and Ryugu ran contrary to expectations for surfaces that have experienced substantial collisional grinding and thermal degradation"
Where the ejecta from craters on Bennu? If fine particles have been lost from Bennu (Lauretta, 2019, Science), the degree of collisional grinding and thermal degradation cannot be accessed from the paucity of fine particles on the surface of Bennu.

Line 68-70: "however, the accumulation of centimeter- to meter scale craters on meter-scale boulders was thought to be infeasible as the energies required for their production should exceed the disruption strength of the targets"
I do not think this is the general view in the era after Hayabusa's observation of asteroid Itokawa. White spots on the surface of boulders on Itokawa were speculated to have been caused by impact of inter planetary dust particles (Ref. [20]). Because the boulders on Benne are porous, the craters are well developed and are easier to be identified. Laboratory craters on non-porous rocks are shallow when compared with those on porous targets.
In other words, the high porosity of the boulders helps this study.

Line 84-86: "there is a gap in our knowledge of the near-Earth object (NEO) size frequency distribution (SFD) at the millimeter to centimeter scale"
As mentioned in above, the Apollo seismic data provided the information of the size frequency distribution of objects at the millimeter to centimeter scale.

Line 99: "Fig. 1d-e"
Where are Fig. 1d-e? It seems the labels are missing, but there are no captions for Fig.1d-e, either.

Line 109: "We find that these craters have depth, d, to diameter ratios, d/D, that are approximately 0.25 (Fig. 1d)."
Where is Fig. 1d? It is easier to see the depth to diameter ratio from the graphs show in Extended Data Figure 1.

Line 111-112: "Craters with diameters greater than 50 cm and up to 5 m were mapped on boulders between ~ 1 and 50 m in the global Bennu basemap produced from the Detailed Survey campaign [35]"

I found that the slope of the cumulative number-boulder size relationship is about -2 for the largest 100 boulders (2.8 - 26 m in radius) in the Extended Data Table 1. This slope is shallower than those shown in Fig. 4 of previous study (DellaGiustina et al., Nature Astronomy 341-351). Why boulders with craters (or boulders with flat surfaces) have relatively larger number of large boulders? The boulders without craters (younger ones) are relatively smaller: they are fragments or they have mobility to change the faces of exposure?

Line 153-154:"The value of μs can be determined by considering an impact that delivers a specific impact energy that is just below that required by the disruption threshold"
Eq. (2) is for the craters on semi-infinite target.
Laboratory cratering experiments show that when the size of crater becomes large enough to be comparable with the size of target (boulder), the cratering efficiency increases. This is thought to be due to spalls, though, and since the craters in this study are pit craters, the curvature effect will not matter.

Ref. Suzuki, A. I., et al., Icarus 301, 1-8, 2018.

Line 181-190 and Fig. 2
The OLA diameter of craters of boulders #1-4 are given in the caption of Extended Data Figure 1, however, the crater diameter of boulder #5 is missing in the caption of Fig. 2c.

Line 233: "The lower range of these estimated values of Y are comparable to the tensile strength determined for boulders on the asteroid Ryugu"
Same as Line 29. The thermophysical analysis does not determine the strength but porosity.

Line 234-235: "Furthermore, our measured impact strength is lower than the strength of carbonaceous chondrite meteorites measured in the laboratory [48].
This is a comparison of apple with tomato, because the strength in [48] is static strength. However, the impact strength (or Q*) of boulders can be compared with those of carbonaceous chondrite meteorites measured in the laboratory as described in above (major comment 3).

Line 255-256
Please refer to the seismic data (Major comment 4).

Line 259-260: While the impact pits on Lunar samples are thought to have been created from 0.1-mm meteoroids originating from comets [23]
An older reference indicates impactors of densities between 2-4 g/cm3.

Ref. Horz et al., Planet. Space Sci. 23, 151-172, 1975.

Line 636-638
Please show the reference for n=2 and n=4 branches.

Referee #2 (Remarks to the Author):

General: This is an important paper for a range of asteroid, impact, and planetary defense science and entirely appropriate for Nature. However, the paper does need some work. I have detailed comments below. Some errors are serious and must be corrected. The organization of the paper was hampered by the strange reliance on dividing the imaging dataset into spacecraft operational phases. There are basically two datasets that are a function of resolution and the authors would be advised to state that in the beginning of the paper and drop the references to operational phases, which may mean a lot to them, but baffled this reader. The derivation of impact scaling parameters uses a lot of space. Most of the discussion and equations would be better put in the supporting material.

Line 50 "contrasting with the relatively fine-grained surfaces of S-type near-Earth asteroids (NEAs) previously visited by spacecraft [10,11]." Not really true and definitely misleading. Itokawa is very boulder rich. I seem to recall that particle size distributions suggest that Itokawa is actually courser than Ryugu. The difference, of course, are the "dust ponds" (actually gravel ponds) on Itokawa which are absent from Bennu and Ryugu. This may be a function of Itokawa's shape and gravity field. Eros has a much finer-grained surface, but this object is 5 orders of magnitude more massive than Bennu…. hence a much different gravity and regolith regime that does not directly compare with Bennu.

Line 107: "We therefore use OLA data to measure the depths of the largest craters in this set," What is the effective ground footprint of OLA in these cases?

Lines 101 and 111: "In the Orbital A and B imaging dataset" "in the global Bennu basemap produced from the Detailed Survey campaign" I don't understand why you are separating the craters mapped during Orbital A and B from the Detailed Survey. The point of the paper is the production of craters on Bennu boulders. The mapping phases of the mission seem a distraction at best and confuse your delivery of data since it seems that there are morphological differences between the small craters seen in A and B with the bigger craters mapped in Detailed Survey. The mapping is really a function of resolution which seems to be different in the different operation phases……just say that! Drop the operational phases which are a distraction. In the subsequent text you seem to assume that the larger craters were produced in the asteroid belt under impact speeds of 5 km/s while the smaller craters are from Bennu's time in NEA space with impact speeds of 20 km/s. If that is so, you need to explicitly state that from the beginning since you do apparently see two crater populations with different impact morphology. This is highlighted in the final section where you explicitly state the different properties of the two populations of craters. This needs to be done throughout the paper.

Line 137: "should represent the largest sub-disruption impact sizes allowable." I think the best you can say in this case is "may" instead of "should". Remember that you are dealing with scaling laws that cover many orders of magnitude.

Line 194-199: There is a lot of discussion of gravity regime scaling which is nice for completeness, but maybe not necessary in a Nature format. Put more of this into the methods section.

Line 234: "Furthermore, our measured impact strength is lower than the strength of carbonaceous chondrite meteorites measured in the laboratory [48]." Wrong! Cotto-Figueroa measured an anhydrous CV carbonaceous chondrite (Allende) which has NOTHING in common with the hydrated material on Bennu or Ryugu. The statement compares Bennu strength with materials we know do not exist on Bennu (apples to oranges). This has to be changed. There are a few measurements of CM materials…see Akira Tsuchiyama, Etsuko Mashio, Yuta Imai, Takaaki Noguchi, Yayoi N. Miura and Hajime Yano. Strength measurements pf carbonaceous chondrites and cosmic dust analogs using micro compression testing machine. In Japan Geoscience Union Meeting 2008 Proceedings, number P168-P002, 2008.

Line 240: "To assess the impact history of Bennu's boulders, we analyze the population of boulders with impact craters measured in the Orbital A and B data set.." Again, I question why the division of data sets into orbital periods. It seems to be that A and B are higher resolution, so you can see smaller craters? Why not just say "used highest resolution imaging to assess impact history…"? The reader does not really care or need to know about O-Rex operational phases.

Line 252: "As the measured CSFD exponent is similar to that of NEOs [25], the average impact speeds were ~20 km/s [24]." Just to be clear….you are not observing impact speeds of 20 km/s, you are assuming them? If that is so, please state that you are making that assumption.

Author Rebuttals to Initial Comments

Notes: Referee Comments are bolded text, and author responses are non-bolded text. Direct quotations from the updated manuscript are italicized.

We thank the reviewers for reading our manuscript carefully and providing valuable and meaningful feedback. We hope that our adjustments to the manuscript have made it more concise and comprehensive. Major changes in the main text and methods are highlighted with red font.

Referee #1:

Although the current version of the manuscript is already very rich in contents and rather comprehensive in spite of the number of equations in methods, I hope that the paper is published after a revision with even richer and clearer descriptions, especially on

(1) the ambiguity of the "impact strength" due to the ambiguity of the assumption of the weakest target radius (Rw),

(2) the description of previous laboratory impact disruption study of carbonaceous meteorites,

(3) the effect of oblique incidence in the main text (not at the end of a section of Methods),

(4) a description of previous study of impactor population derived from seismic data on the Moon.

1. Robustness of the assumptions and parameters in this study
(1) On the assumption of Rw=80 m (the radius of Otohime boulder).
This is the crucial assumption in this study. If we chose larger Rw as in the previous studies, then the resultant "impact strength" becomes much larger. Isn't there any C-complex fast-spinning object larger than Otohime boulder?

We agree with the referee that setting Rw=80 m, based on the size of Otohime boulder, is one of the critical assumptions of this paper. As we prepared our manuscript, we searched the asteroid spectral and photometric databases in search of fast-spinning C-complex objects that are larger than Otohime. We discovered that none of the asteroid with fast spins are C-complex objects; they are mostly higher-albedo S-complex or V-type objects. We also contacted Dr. Petr Pravec, an authority on light-curve observations of fast-spinning asteroids, and he also was unable to find any fast-spinners with diameters between 50 and 200 meters that were C-complex objects.

This is likely due to observational biases – as, for a given size, C-complex objects are more difficult to observe in visible wavelengths due to their low albedos. Assuming their mean geometric albedo p_V = 0.057, C-complex objects with diameters between 50 and 200 meters (where super-fast rotators predominate) have absolute magnitudes between 25.2 and 22.2. There are very few asteroids with magnitude > 22.2 that have their spectral type determined.

Therefore, our assessment is that Otohime is the largest monolithic C-complex object that has been observed either directly or inferred through lightcurve observations.

(2) On the derivation of ⍰s by a curve fit to the maximum crater radii (Fig.2).
Due to the large uncertainty in the data of Rc/RT and small number of data points (five red triangles in Fig. 2b), we might only say that the value of ⍰s allowed by the boulder data is about right (say, between 0.33 and 0.55) for porous target. Or can we say with more confidence? For example, if we only have 4 data points instead of the 5 in this study, can we still have the similar ⍰s value?

We thank the reviewer for their comment.
Our error estimate for μ'' of 0.02 is based on the simple least squares error found through fitting a straight line through the five data points. We considered three different approaches.

1. We first considered the reviewer's recommendation of analyzing how the value of μ'' might change if we only had four data points, rather than five. For this, we removed each one of the five data points in turn and found μ'' = 0.41, 0.47, 0.50, 0.57, and 0.36. This is similar to the reviewer's suggested range of 0.33 to 0.55.

   We note that the mean and median of these values are 0.46 and 0.47, respectively, which is similar to the value we obtain with five data points. However, the standard deviation of these values is large, 0.08, compared to our initial error estimate.

2. Rather than ignoring a data point, we tried an approach where we would estimate the error in the value of μ'' through Monte Carlo trials.
   In this approach, we replaced one of the five data points with a data point that has a similar Rt value, but replaced the Rc/Rt with a new value drawn from a uniform distribution of values within the range given by the error bars of each measurement (see Extended Data Table 2). For a variable number of Monte Carlo trials, *N*, we find the following:

N	mean μ''	median μ''	Standard Deviation
10	0.434	0.440	0.079
100	0.472	0.472	0.078
1,000	0.474	0.473	0.076
10,000	0.471	0.469	0.074

100,000	0.471	0.470	0.074

3. Finally, we also considered an analytical estimate of the uncertainty in $\mu_{"}$, calculated through error propagation analysis.

   For the functional form: y = $x^{(}$, where y, x, and w are uncorrelated variables with associated uncertainties, then the standard uncertainty can be expressed as [e.g., see Table 2 of Farrance & Frankel 2012]:

$$(u(y)/y)^2 = w^2((u(x)/x)^2 + (\ln x)^2 (u(w)/w)^2)$$

where u(y), u(x), and u(w) are the uncertainties in y, x, and w, respectively. As we are trying to estimate the uncertainty in the exponent of this function form, we rearrange the equation:

$$u(w) = (1/\ln(x))((u(y)/y)^2 - w^2(u(x)/x)^2)^{1/2}$$

For our purposes, y = Rc, max/Rt, x = Rt, and w = $-5\mu_s/24$.

The uncertainty in $R_{c,max}/Rt$, $(Rc, max/Rt)$, is given by:

$$u(Rc, max/Rt) = Rc, max/Rt\ ((u(Rc)/Rc)^2 + (u(Rt)/(Rt))^2)$$

where $u(Rc)$ and $u(Rt)$ is the uncertainty in Rc and Rt , respectively. (Rc) is driven by uncertainty in the location of the crater rim. The values of (Rc) are given in Extended Data Table 2. We consider (Rt) ~ 15 cm, equivalent to ~ 3 pixels in the high-resolution image data.

The uncertainty in μ_s, $u(\mu_s)$, is given by:

$$(u(\mu_s))^2 = (u(w))^2(-24/5)^2$$

Finally, giving:

$$u(\mu_s) = (24/5)(1/\ln(Rt))((u(Rc, max/Rt)/(Rc, max/Rt))^2 - w^2(u(Rt)/Rt)^2)^{1/2}$$

Taking mean values for Rt, $(Rc, max/Rt)$, and $(Rc, max/Rt)$, we find $u(\mu_s)$ = 0.066 ~ 0.07, which is similar in value to the standard deviation found from our Monte Carlo trials.

To conclude, we find through these three methods that we did indeed likely underestimate the uncertainty in our value of $\mu_{"}$. We estimate that this uncertainty is actually ~ 0.07. We have updated mentions of this uncertainty in the text. We include the uncertainty analysis in Method3, as this is likely the most robust and reproduceable method, in the methods section, under **Uncertainty Estimate for μ_s.**

Ref:

Farrance, I. & Frankel, R. Uncertainty of Measurement: A Review of the Rules for Calculating Uncertainty Components through Functional Relationships. Clinical Biochemistry Review 33, 49- 75 (2012).

(3) On the assumption of Rc=Rc, max at the impact of specific energy that is just below that required by the disruption threshold, I have only minor comments. For a non-porous rock target (with spall fracture), the specific energy required for the largest crater and that for the disruption threshold is different in a few times.

Ref. Fig. 2 in Housen, Planet. Space Sci. 57, 142-153, 2009.

The assumption is maybe more suitable for a porous target or target without spall fracture (as boulders in this study).
Ref. Fig. 3 in Murakami et al., Planet Space Sci. 182, 104819, 2020.

Thank you for pointing out these important references. We have included this caveat in the main text, and expand on evidence that would corroborate our decision to use this framework:

“Laboratory impact experiments show that this is possible for impacts onto porous targets [11]; however, the largest possible craters on non-porous consolidated targets, produced through spall (fracturing and ejection of plate-like near-surface fragments), are created by impacts that are a factor of a few less than the disruption threshold [12]. Therefore, equating the formation of the largest possible crater to the catastrophic disruption threshold is a viable framework, as CM and CI meteorites—the meteoritic analogs to Bennu’s boulders—have high porosities (≳ 20% [13]) and Bennu’s boulders show little evidence for spalls”

2.Impact strength of boulders and cratering efficiency in the strength regime. The concept of impact strength is unfamiliar to most of readers, because it is not a widely recognized concept such as static tensile strength nor compressive strength. Below Eq. 2, it is described as the material strength that determines the diameter of crater. After Eq. M12, readers find that it probably denotes shear strength (or a kind of) in this study. More explanation of the "impact strength" in the main text may be of help for understanding.

We agree that the current discussion of impact strength is ambiguous, particularly to those researchers immediately outside of the field of impact cratering physics. We have modified the main text to clarify this point:

“We can then use our Q^*_{T} estimate to determine the size-dependent impact strength, , of monolithic C-complex objects (see Eq. (M28)). The size dependence of strength is a consequence of the increase in the number and size of internal cracks and flaws, which limit the strength of a material, with the size of an object. Strength is a measure of a material’s

ability to withstand a particular stress, such as compressive, tensile or shear [19]. An object's response to an impact, and therefore its crater formation mechanics, is dominated by one of these strengths. The formation of well-defined deep craters, as observed on Bennu's boulders, is typically dominated by shear [19] or compressive strength [11]."

Moreover, using the relationship of Eq. M17, the cratering efficiency of the boulders on Bennu can be compared with various porous materials used in cratering experiments in the laboratory, for example, weak cemented basalt (WCB) has a µs value of 0.46 (similar to that of 0.47 in this study). Interestingly, the tensile strength of the material estimated using an empirical relationship was 0.45 MPa, roughly twice of that of the estimated tensile strength of the Ryugu boulder.

Ref. Table 4 in Housen and Holsapple, Icarus 211, 856-875, 2011.

We thank the referee for pointing this out. We have revised the main text to provide comparison to the Ryugu boulder result by Grott et al. (2019), the strength of hydrated carbonaceous chondrite meteorites, and the terrestrial material gypsum and WCB.

"The lower range of these estimated values of Y are comparable to the tensile strength (inferred from high porosities) of boulders on Ryugu, which ranges from 0.2 to 0.28 MPa [20]. Our measured impact strength is lower than the 85 MPa compressive strength reported for the CM2 Sutter's Mill meteorite [21], but similar to the 0.25–0.7 MPa compressive strength estimated for the ungrouped C2 Tagish Lake fireball [26]. Weakly cemented basalt and gypsum have similar μ_W values of 0.46 [22] and 0.4 [23], respectively. Furthermore, gypsum has a similar tensile strength of 0.3 MPa; however, impacts on solid gypsum form deeper craters than those we measure on Bennu's boulders [23]."

3. Disruption threshold of boulders

The curves of basalt and pumice shown in Fig. 3 are not for normal impacts, but for oblique impacts. Therefore, a comparison between these curves and the result of this study needs discussion on the effect of impact angle on the disruption threshold. There are a few laboratory disruption experiments.

Ref. Yasui et al., Icarus 335, 113414, 2020.

Ref. Fig. 6 in Murakami et al., Planet Space Sci. 182, 104819, 2020 for extremely porous targets.

We agree with the referee that a discussion of the effect of an oblique impact should have been included in the main text, and not simply put in the methods. Comparing the change in Q*D in Benz & Asphaug 1999 (their Figs. 2-5) for impacts at 0-deg (projectile travels along a vector normal to the plane of impact) to 45-deg, we find that the value of Q*D varies by a factor of ~2, although this increases to a factor of ~ 3 when comparing a 0-deg impact to a 60-deg impact. Since the most likely impact angle is ~ 45-deg and ~ 75% of impacts have angles < 60 deg* [e.g., Shoemaker 1962], we can assume that the difference between our derived Q*D value compared to the basalt and pumice is indeed significant. We now include the following sentence in the main text to highlight the effect of oblique impacts as shown in simulations and experiments.

"The values of Q_D^* measured for basalt and pumice were for oblique impacts, which lead to higher Q_D^*, as demonstrated in laboratory impact experiments [18]."

And modified the Figure 3 caption to now say:
"We compare this threshold to those of basalt (black dotted line) and pumice (black dashed line) derived from numerical simulations of impacts at angles of 45° [16,17]."

*This was found by taking the differential probability that a meteoroid from an isotropic flux impacts at angle x between x + dx, is given by: $d(x) = 2\sin(x)\cos(x)\,dx$, and integrating from 30 deg to 90 deg.

(since the definition of impact angle is flipped from that of Benz & Asphaug 1999, where x= 90 deg corresponds to a 0-deg head-on impact in Benz & Asphag 1999)

Ref.:
Shoemaker, E.M. Interpretation of lunar craters. Kopal, Z. (Ed.), Physics and
Astronomy of the Moon Academic Press. New York, New York, 283–359
(1962).

I'd like to suggest to refer to the result of impact disruption experiment of carbonaceous chondrites, Allende and Murchison in the literature. Because they are originally from C-complex objects and akin to the boulders on Bennu and Ryugu. Static strength was measured for these chondrites in the laboratory, therefore, showing the result (probably in Fig.3) will promote understanding.

Ref. Fig. 32 in Flynn et al., Chemie der Erde 78, 269-298, 2018, and references therein.

We now include a reference to Flynn et al. (2018) in the summary. Though, we do note that Allende is a CV, unlike boulders on Bennu and Ryugu which are expected to be hydrated carbonaceous chondrites such as CM and CI. Flynn et al. (2018) do discuss impact experiments to measure Q*D for meteorites (citing previous work by Flynn and Durda 2004), as well as new experiments into “wet carbonaceous meteorites”. However, Flynn et al. (2018) do not provide a detailed description of the samples tests, their sizes, or the experimental setup. Therefore, it is difficult to directly compare their results with ours, as we suggest a size dependence in the values of both Q*D and Y. Nevertheless, in the interest of referencing this other important data point in the literature, we now state:

“Furthermore, we find that 1-cm radius targets have a Q_D^*~1.5×10^7 *erg/g, which is comparable to the* $Q_D^* > 1.419 \times 10^7$ *erg/g found for hydrated carbonaceous chondrites [3].”*

The catastrophic disruption threshold of Bennu boulders shown in Fig. 3a is weaker than basalt and pumice. Interestingly, however, it is very close to gypsum (with porosity of 65% which is more porous than the popular gypsum used for statue), which was shown to have a tensile strength of 0.3 MPa (similar to the estimated strength of a boulder on Ryugu). The µs value for the gypsum was 0.4 and close to the µs value (0.47) of the boulders of Bennu derived in this study, however, the ratio of depth to diameter of craters on gypsum targets are larger [36].

We thank the reviewer for pointing this out, and have added a statement summarizing the similarity of impact parameters and tensile strengths between Bennu boulders and gypsum.

In response to these the two preceding comments, we have expanded our discussion on comparisons of our estimated strength properties of Bennu’s boulders to those found for terrestrial materials and meteorites. We now state:

“We can then use our Q_D^* *estimate to determine the size-dependent impact strength, , of monolithic C-complex objects (see Eq. (M28)). The size dependence of strength is a consequence of the increase in the number and size of internal cracks and flaws, which limit the strength of a material, with the size of an object. Strength is a measure of a material’s ability to withstand a particular stress, such as compressive, tensile or shear [19]. An object’s response to an impact, and therefore its crater formation mechanics, is dominated by one of these strengths. The formation of well-defined deep craters, as observed on Bennu’s boulders, is typically dominated by shear [19] or compressive strength [11]. In contrast, impacts onto brittle material lead to shallow spall craters formed by tensile failure. For impacts with* $U = 5$ *km/s and* $\mu_G = 0.33 - 0.36$*, a 1-m-diameter boulder on Bennu has* $Y = 0.44 - 1.70$ *MPa, which may approximate its shear strength and/or compressive strength. The lower range of these estimated values of* Y *are comparable to the tensile strength (inferred from high porosities) of boulders on Ryugu, which ranges from 0.2 to 0.28 MPa [20]. Our measured impact strength is*

lower than the 85 MPa compressive strength reported for the CM2 Sutter's Mill meteorite [21], but similar to the 0.25–

0.7 MPa compressive strength estimated for the ungrouped C2 Tagish Lake fireball [26]. Weakly cemented basalt and gypsum have similar μ_W *values of 0.46 [22] and 0.4 [23], respectively. Furthermore, gypsum has a similar tensile strength of 0.3 MPa; however, impacts on solid gypsum form deeper craters than those we measure on Bennu's boulders [23]. "*

4. Size frequency distribution of the objects of millimeter-centimeter scale. Apollo seismic data provided the information about the objects in this range. Because lunar seismology (for the purpose of constraining the internal structure of the Moon) has long been an interest, the comparison of the present result with that of previous seismic data will be beneficial for the future planning of the seismology on the Moon.

Ref. Fig. 10.3 in "Impact Cratering: A Geologic Process" by Melosh, H. J., 1989.

An example of more recent analysis can be found here:

Ref. Fig. 2 in Kawamura et al., J. Geophys. Res. 38, L15201, 2011.

We thank the reviewer for suggesting this point, and include a reference to the work done on using seismic data to estimate the SFD of impactors.

We point out that the Apollo seismic data does not capture the mm impactor population, rather they are able to observe the centimeter-decimeter population – similar to the impact flash data.
As show in Melosh 1989, Fig. 10.3, the seismic data described in Latham et al. 1973 is caused by impactors with estimated masses of 100 g – 1000 kg. Assuming a spherical projectile with densities between 1 to 3 g/cm^3, this corresponds to an impact diameter between ~ 4 to 6 cm.

Following Kawamura et al. 2011, who present the SFD of impact energies more clearly, we see that population statistics are complete for energies > 10^9 Joules, at a minimum (see their Fig. 3 and 4). For impact speeds of 10-20 km/s, an impact energy of 10^9 Joules corresponds to an impactor with a mass of 5 to 20 kg. Assuming a spherical projectile with density of 3 g/cm^3, this corresponds to an impact diameter between ~ 14 to 23 cm. This would put the seismic data squarely in the decimeter size range, rather than the mm-cm range.

We have updated the text to now state:

The centimeter- to decimeter-scale population is also measured by observations of impact flashes

[27] and seismic events [Lognonne et al. 2009, Kawamura et al. 2011] on the Moon [27].”

We also include the range of NEO sizes that have been observed in the lunar seismic data in Fig. 4, and state in the caption:

“The centimeter- to decimeter-scale near-Earth object (NEO) population is also measured by observations of impact flashes [25] and seismic events [26] on the Moon.”

Specific comments
Line 29: "Ryugu revealed an unexpectedly low tensile strength [3]"

I’m afraid that this sentence is misleading. The thermophysical analysis directly revealed "high porosity" of a boulder on the surface of Ryugu. High porosity does not necessarily mean low tensile strength (imagine a metal honeycomb structure).

Thank you for pointing this out. It is true that rather than revealing or measuring a low tensile strength for Ryugu boulders, the tensile strength was inferred. We have removed this sentence from the summary, and updated our comparison in the main text to now state:

“The lower range of these estimated values of Y are comparable to the tensile strength (inferred from high porosities) of boulders on Ryugu, which ranges from 0.2 to 0.28 MPa [20].”

Lines 52-53: "Bennu and Ryugu ran contrary to expectations for surfaces that have experienced substantial collisional grinding and thermal degradation” Where the ejecta from craters on Bennu? If fine particles have been lost from Bennu (Lauretta, 2019, Science), the degree of collisional grinding and thermal degradation cannot be accessed from the paucity of fine particles on the surface of Bennu.

We have removed this sentence from the manuscript, and do not include any discussion on the outcome of fine regolith production on the surface of an asteroid.
Certainly, the balance between fine particle production (e.g. collisions, thermal degradation) and loss (e.g., electrostatic lofting, brazil nut effect) on Bennu is still unknown as the timescales for the different mechanisms that contribute to each are currently poorly constrained.

Given the added uncertainty of the main belt residence time of Bennu itself [Walsh et al. 2019] in light of the results from the SCI experiments suggesting a gravity-dominated crater scaling relationship for the surface of small regolith-covered asteroids [Arakawa et al. 2020], addressing the reason for the paucity of fine particles on Bennu is out of the scope of this paper, and may distract from the main results.

Line 68-70: "however, the accumulation of centimeter- to meter scale craters on meter- scale boulders was thought to be infeasible as the energies required for their production should exceed the disruption strength of the targets"

I do not think this is the general view in the era after Hayabusa's observation of asteroid Itokawa. White spots on the surface of boulders on Itokawa were speculated to have been caused by impact of inter planetary dust particles (Ref. [20]). Because the boulders on Benne are porous, the craters are well developed and are easier to be identified. Laboratory craters on non-porous rocks are shallow when compared with those on porous targets. In other words, the high porosity of the boulders helps this study.

We include a reference to the white spots observed on Itokawa's boulders in the summary. We have also removed the quoted sentence as we removed our introduction and condensed the main points into the summary.

However, white spots on boulders from inter-planetary dust particles would fall under the first mode of boulder degradation, "sand-blasting", mentioned in the introduction of the first version of the manuscript.

We agree with the reviewer, that this may not have been the view in the post-Hayabusa era. However, this certainly was not a *universal* view by the community. The particular comment: "*centimeter- to meter scale craters on meter-scale boulders was thought to be infeasible as the energies required for their production should exceed the disruption strength of the targets*" comes from the well-cited Hörz and Cintala 1997 retrospective: Impact experiments related to the evolution of planetary regoliths.

Many modern works still use the related Hörz et al. 1975 prescription for impact-induced boulder degradation (For example, to compare against other degradation mechanisms like thermal fatigue).

55 citations for Hörz and Cintala 1997, and 21 citations for Hörz 1975 since 2008.

Line 84-86: "there is a gap in our knowledge of the near-Earth object (NEO) size frequency distribution (SFD) at the millimeter to centimeter scale" As mentioned in above, the Apollo seismic data provided the information of the size frequency distribution of objects at the millimeter to centimeter scale.

As mentioned in our reply above, the seismic data does not adequately capture the millimeter size range, rather the centimeter-decimeter scale. We have included a reference to the seismic data in the following:

"The centimeter- to decimeter-scale near-Earth object (NEO) population is also measured by observations of impact flashes [25] and seismic events [26] on the Moon."

Line 99: "Fig. 1d-e"
Where are Fig. 1d-e? It seems the labels are missing, but there are no captions for Fig.1d-e, either.

Line 109: "We find that these craters have depth, d, to diameter ratios, d/D, that are approximately 0.25 (Fig. 1d)." Where is Fig. 1d? It is easier to see the depth to diameter ratio from the graphs show in Extended Data Figure 1.

We apologize for not including complete panel labels for Fig. 1, we have now included them as well as associated captions. Figs. 1d-e represent the colored OLA point cloud data. The d/D values can be evaluated by comparing the colors between the crater centers and crater rim. We now also refer the reader to the crater profile in Fig. 2 and the crater profiles in Extended Data Figure 2.

Line 111-112: "Craters with diameters greater than 50 cm and up to 5 m were mapped on boulders between ~ 1 and 50 m in the global Bennu basemap produced from the Detailed Survey campaign [35]"

I found that the slope of the cumulative number-boulder size relationship is about -2 for the largest 100 boulders (2.8 - 26 m in radius) in the Extended Data Table 1. This slope is shallower than those shown in Fig. 4 of previous study (DellaGiustina et al., Nature Astronomy 341-351). Why boulders with craters (or boulders with flat surfaces) have relatively larger number of large boulders? The boulders without craters (younger ones) are relatively smaller: they are fragments or they have mobility to change the faces of exposure?

We thank the reviewer for these interesting questions. At the moment we can offer some speculation for the reason why the boulders that express craters have a shallower CSFD than the global boulder population.

First, we note that the CSFD for boulders in the DellaGiustina and Emery et al. (2019) is complete down to boulder diameters of 5 m. The team has yet to complete a final assessment of the global boulder CSFD down to smaller sizes. However, analysis of the boulder CSFD at specific regions of interest has shown a shallowing of slope the CSFD to values lower than the -2.5 found globally, for sub-meter completeness limits. This suggests that the collisional processing of the Bennu surface may lead to a shallower CSFD than the

Bennu-formation event, represented by the CSFD of the largest boulder on the surface (see Supplementary Figure 5 of DellaGiustina and Emery et al. (2019) which shows that the largest boulders on the surface, > 20 m, must originate from the catastrophic disruption of the Bennu parent body).
As we illustrate in Fig. 3b, the mean collisional lifetime of a 50-m boulder is about an order of magnitude larger than a 5-m boulder. Depending on Bennu's residence time in the main asteroid belt, its surface may have experienced substantial comminution of the smaller boulder population, compared to the large boulder population. Therefore, the boulder SFD we observe may have a break from collisional equilibrium at the small sizes we report here.

Finally, the shallow SFD may simply represent an incompleteness of our survey. It is easier to observe craters on larger boulders than smaller ones. Therefore, our mapping of craters on boulders will certainly be biased towards having a shallower CSFD compared to the global boulder CSFD.

The reviewer's questions do provide us with interesting additional lines of inquiry; however, we think that the speculations we have just outlined may distract from the main results of the paper, and are also partly based on unpublished results. Therefore, we do not include them in the main text.

Line 153-154:"The value of μs can be determined by considering an impact that delivers a specific impact energy that is just below that required by the disruption threshold"
Eq. (2) is for the craters on semi-infinite target.

Laboratory cratering experiments show that when the size of crater becomes large enough to be comparable with the size of target (boulder), the cratering efficiency increases. This is thought to be due to spalls, though, and since the craters in this study are pit craters, the curvature effect will not matter.

Ref. Suzuki, A. I., et al., Icarus 301, 1-8, 2018.

We thank the referee for this insight. We have included a reference to Suzuki et al. 2018 in the methods section.

"Laboratory impact experiments have shown that cratering efficiency may increase when the size of the crater becomes comparable to the target boulder size [45]; however, this is thought to be due to spallation, which is not a dominant mechanism for the craters we observe on Bennu's surface."

Line 181-190 and Fig. 2
The OLA diameter of craters of boulders #1-4 are given in the caption of Extended Data Figure 1, however, the crater diameter of boulder #5 is missing in the caption of Fig. 2c.

We thank the referee for spotting this oversight. We have now included the crater diameter of boulder #5 in the caption of Fig. 2c.

Line 233: "The lower range of these estimated values of Y are comparable to the tensile strength determined for boulders on the asteroid Ryugu"
Same as Line 29. The thermophysical analysis does not determine the strength but porosity.

As we describe above, we have changed our description of the Grott et al. 2019 findings to state:

"The lower range of these estimated values of Y are comparable to the tensile strength (inferred from high porosities) of boulders on Ryugu, which ranges from 0.2 to 0.28 MPa [20]."

Line 234-235: "Furthermore, our measured impact strength is lower than the strength of carbonaceous chondrite meteorites measured in the laboratory [48]. This is a comparison of apple with tomato, because the strength in [48] is static strength. However, the impact strength (or Q*) of boulders can be compared with those of carbonaceous chondrite meteorites measured in the laboratory as described in above (major comment 3).

We thank the reviewer for this comment. We have removed the reference to [48] since it is a study of the static strengths of a CV rather than a hydrated carbonaceous chondrite such as a CM and a CI. There is no laboratory measurement for hydrated carbonaceous chondrite meteorites. Flynn et al. 2018 report the results of impact experiments into Allende, a CV - dry carbonaceous chondrite, to be 1419 J/Kg, and performed three experiments on to a "wet carbonaceous chondrite" but provide no experimental details, or a value of Q*D.

They state:
"the three wet carbonaceous chondrite (CM2)data points all plot to the right of the ordinary chondrite data, suggesting these wet carbonaceous chondrites require even more specific impact energy for disruption than the ordinary chondrites."

So we may say that Q*D > 1419 J/Kg, but for what size object? This would be a similar situation of comparing apples to tomatoes.

In conjunction with reviewer 2's suggestion to compare against static strength measurements of hydrated carbonaceous chondrite meteorites, we write the following (restricting our comparisons to static strength measurements found in peer-reviewed articles):

"Furthermore, we find that 1-cm radius targets have a Q_D^* ~1.5 × 10^7*erg/g, which is comparable to the* Q_D^* > 1.419 × 10^7*erg/g found for hydrated carbonaceous chondrites [3]."*

And

"Our measured impact strength is lower than the 85 MPa compressive strength reported for the CM2 Sutter's Mill meteorite [21], but similar to the 0.25–0.7 MPa compressive strength estimated for the ungrouped C2 Tagish Lake fireball [26]."

Line 255-256
Please refer to the seismic data (Major comment 4).

As mentioned in our reply above, the seismic data does not adequately capture the millimeter size range, rather the centimeter-decimeter scale. We have included a reference to the seismic data in the following:

"The centimeter- to decimeter-scale near-Earth object (NEO) population is also measured by observations of impact flashes [25] and seismic events [26] on the Moon."

Line 259-260: While the impact pits on Lunar samples are thought to have been created from 0.1-mm meteoroids originating from comets [23]
An older reference indicates impactors of densities between 2-4 g/cm3. Ref. Horz et al., Planet. Space Sci. 23, 151-172, 1975.

We thank the reviewer for catching this error in citation. Indeed Horz et al. 1975 do argue for impactor densities between 2-4 g/cm3 based on the depth to diameter ratio found in laboratory impact experiments compared to that of the impact pits on lunar samples (their Fig. 4). We note that their low density projectile is polystyrene, which may not be the best analog for impact by icy particles. Our original citation [23], does argue for a possible cometary origin for sub-mm meteoroids – though they do not come to a firm conclusion either way.

The actual argument for a cometary origin for the sub-mm meteoroid population comes from Nesvorny et al. 2010. We have included an additional reference [Nesvorny et al. 2010], who perform a dynamical analysis of the dust population and conclude that the majority (~85%) of the near-Earth sub-mm dust population originate from Jupiter family comets. In the main text, we now state:

"Whereas the impact pits on lunar samples are thought to have been created from sub-mm meteoroids [27] likely originating from comets [31], we find that the NEO population of objects

> 1 mm have an asteroidal origin, as Bennu's boulders show that their CSFD has a larger exponent."

Line 636-638
Please show the reference for n=2 and n=4 branches.

We have included references to these branches in the methods section:

"n = 4 for normal craters dominated by shear strength, and n = 2 for spall craters. Y_0 is the impact strength of a target with a radius of 1 cm. As we see little evidence for spall-dominated craters on Bennu's boulders, we adopt a value of n = 4 [43,40]."

Referee #2 (Remarks to the Author):

General:

The organization of the paper was hampered by the strange reliance on dividing the imaging dataset into spacecraft operational phases. There are basically two datasets that are a function of resolution and the authors would be advised to state that in the beginning of the paper and drop the references to operational phases, which may mean a lot to them, but baffled this reader.

We thank the reviewer for this comment. Our original intention was to highlight the mission phases in order to describe the uniqueness of the observational geometry, which allowed us to measure impact pits on flat boulder faces – a feature that previous missions to spacecraft have not been able to do. Just as it is easier to observe craters on the moon along the terminator, the orbital phases of the mission, where the spacecraft was in a terminator orbit, made it easier to identify small craters on the surface of flat boulders. We realize now that this attempt ended up detracting from the clarity of the manuscript.

We have removed mention of orbital phases in the main text.

The derivation of impact scaling parameters uses a lot of space. Most of the discussion and equations would be better put in the supporting material.

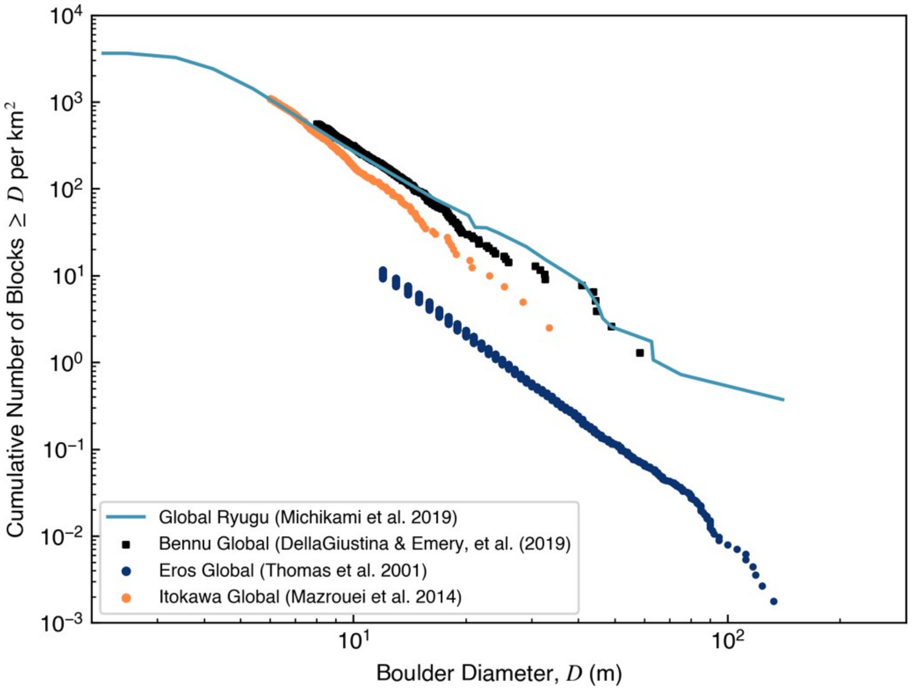
This section was substantially reduced.

Material in the main text was moved to the methods section.

Line 50 "contrasting with the relatively fine-grained surfaces of S-type near-Earth asteroids (NEAs) previously visited by spacecraft [10,11]." Not really true and definitely misleading. Itokawa is very boulder rich. I seem to recall that particle size distributions suggest that Itokawa is actually courser than Ryugu. The difference, of course, are the "dust ponds" (actually gravel ponds) on Itokawa which are absent from Bennu and Ryugu. This may be a function of Itokawa's shape and gravity field. Eros has a much finer-grained surface, but this object is 5 orders of magnitude more massive than Bennu…. hence a much different gravity and regolith regime that does not directly compare with Bennu.

We remove this sentence from the manuscript as we think it distracted from the main points and did not add anything substantive to the discussion. However, we do not think it was incorrect.

As stated in DellaGiustina & Emory et al. (2019) and Michimaki et al. (2020). The absolute number density of large boulders (≥20 m) on Bennu and Ryugu is greater than that on Itokawa. We reproduce Supplementary Fig. 4 from DellaGiustina & Emory et al. (2019), adding in the Ryugu data from Michikami et al. (2020).



Nevertheless, as we state in our response to reviewer 1:
The balance between fine particle production (e.g. collisions, thermal degradation) and loss (e.g., electrostatic lofting, brazil nut effect) on Bennu is still unknown as the timescales for the

different mechanisms that contribute to each are currently poorly constrained.

Furthermore, the difference between the C-complex asteroids and Itokawa may indeed be a strong function of the shape of the two asteroids.

Given the added uncertainty of the main belt residence time of Bennu itself [Walsh et al. 2019] in light of the results from the SCI experiments suggesting a gravity-dominated crater scaling relationship for the surface of small regolith-covered asteroids [Arakawa et al. 2020], addressing the reason for the paucity of fine particles on Bennu is out of the scope of this paper, and may distract from the main results.

Line 107: "We therefore use OLA data to measure the depths of the largest craters in this set," What is the effective ground footprint of OLA in these cases?

As stated in the methods section, the effective ground footprint of OLA is about 5 cm, globally. This is also true for these cases.

Lines 101 and 111: "In the Orbital A and B imaging dataset" "in the global Bennu basemap produced from the Detailed Survey campaign" I don't understand why you are separating the craters mapped during Orbital A and B from the Detailed Survey. The point of the paper is the production of craters on Bennu boulders. The mapping phases of the mission seem a distraction at best and confuse your delivery of data since it seems that there are morphological differences between the small craters seen in A and B with the bigger craters mapped in Detailed Survey. The mapping is really a function of resolution which seems to be different in the different operation phases……just say that! Drop the operational phases which are a distraction.

We remove all mention of mission phases, and instead make a distinction between the two data sets by describing the image resolution each were mapped in, as the reviewer suggested. In the main text, we now state:

"We studied images of asteroid Bennu taken by the PolyCam instrument, part of the OSIRIS-REx Camera Suite (OCAMS) [8] onboard the Origins, Spectral Interpretation, Resource Identification, and Security–Regolith Explorer (OSIRIS-REx) spacecraft. These images resolved circular cavities that we interpret as craters with diameters, DC, between 0.03 and 5 m, on boulders with diameters between 0.5 and 50 m (Fig. 1a-c). These images are divided into two datasets that differ in resolution and thus in resolvable crater size [9]."

In the subsequent text you seem to assume that the larger craters were produced in the

asteroid belt under impact speeds of 5 km/s while the smaller craters are from Bennu's time in NEA space with impact speeds of 20 km/s. If that is so, you need to explicitly state that from the beginning since you do apparently see two crater populations with different impact morphology. This is highlighted in the final section where you explicitly state the different properties of the two populations of craters. This needs to be done throughout the paper.

We thank the reviewer for highlighting this point. In its original form, the choice of using 5 km/s or 20 km/s for different parts of the analysis was poorly justified. In reality, our analysis required us to test whether the formation of each subset of craters on boulders was viable in the main asteroid belt or near-Earth space. We are then able to develop constraints (Bennu's NEA lifetime and the collisional lifetime of boulders) that validated these assumptions (smallest observed craters formed in NEA-residence, and large craters formed in Main Belt residence).

The first instance of an assumed speed occurs when we are estimating Q*D for Bennu's boulders: "*Using the derived prescriptions for the catastrophic disruption parameters, Fig. 3a* T *compares* Q^* *for monolithic C-complex objects at typical main belt impact speeds, U = 5 km/s [2] with simulation results [4,5] for disruption against basalt and pumice targets."*

This assumption is valid to make here as we are comparing against numerical studies that were performed at impacts speeds of 5 km/s.

Then, we assume 20 km/s when discussing the mm-cm NEO population, where we state:
"We find that the best fit has $\alpha = -2.69 \pm 0.07$ *for a completeness limit of 13 cm (Fig. 4a). This value of* α *is consistent with that reported for larger near-Earth objects with diameters between 1 and 10 m, based on observations of fireballs and bolides [24]. Therefore, we postulate that these craters were created during Bennu's residence time in near-Earth space.*

…

As the measured CSFD exponent is similar to that of NEOs [24], we assume that the average impact speeds are ~20 km/s [30]."

Here, our assumption is validated as the SFD of the small impactors is similar to that of larger NEOs.

We then revisit the U=5 km/s for craters close to the maximum limit when discussing the collisional lifetime of Bennu's boulders:

"During Bennu's residence time in the main belt, its surface boulders would have collisionally evolved more quickly. Once a C-complex asteroid dynamically and collisionally decouples from the main belt, impact cratering, rather than disruption, becomes the primary mechanism for impact-induced breakdown."

Finally we conclude the main text with a discussion that validates these assumptions:
"We conclude that the large craters on Bennu's boulders ($R_{C}/R_{T} > 0.3$*, for *$R_{T} > 1$* m) were created by impacts with energies that are close to the boulder's disruption limit and formed during Bennu's residence time in the main asteroid belt. Conversely, the small craters (DC < 50 cm) observable on flat boulders were formed more recently, during Bennu's residence time in near- Earth space over the past 1.75 ± 0.75 Myr."*

Line 137: "should represent the largest sub-disruption impact sizes allowable." I think the best you can say in this case is "may" instead of "should". Remember that you are dealing with scaling laws that cover many orders of magnitude.

We agree with the author that this statement may need to be softened. We have revised that sentence, and described additional evidence that support our assumption:

"The largest craters relative to the host boulders may represent the largest sub-disruption impact sizes allowable. Laboratory impact experiments show that this is possible for impacts onto porous targets [11]; however, the largest possible craters on non-porous consolidated targets, produced through spall (fracturing and ejection of plate-like near-surface fragments), are created by impacts that are a factor of a few less than the disruption threshold [12]. Therefore, equating the formation of the largest possible crater to the catastrophic disruption threshold is a viable framework, as CM and CI meteorites—the meteoritic analogs to Bennu's boulders—have high porosities (≳ 20% [13]) and Bennu's boulders show little evidence for spalls."

Line 194-199: There is a lot of discussion of gravity regime scaling which is nice for completeness, but maybe not necessary in a Nature format. Put more of this into the methods section.

This has been significantly shortened, and a more detailed description is found in the methods section. We now only introduce some key assumptions which are required to formulate a complete description of the catastrophic disruption threshold of monolithic C-type objects up to the transition diameter between strength and gravity dominated objects. This section in the main text now simply is written as:
"The gravity regime parameters can be estimated from the results of numerical simulations and spacecraft observations of large asteroids (see Methods Eq. (M22-M23))."

Line 234: "Furthermore, our measured impact strength is lower than the strength of carbonaceous chondrite meteorites measured in the laboratory [48]." Wrong! Cotto-Figueroa measured an anhydrous CV carbonaceous chondrite (Allende) which has NOTHING in common with the hydrated material on Bennu or Ryugu. The statement compares Bennu strength with materials we know do not exist on Bennu (apples to oranges). This has to be changed. There are a few measurements of CM materials…see Akira Tsuchiyama, Etsuko Mashio, Yuta Imai, Takaaki Noguchi, Yayoi N. Miura and Hajime Yano. Strength measurements pf carbonaceous chondrites and cosmic dust analogs using micro compression testing machine. In Japan Geoscience Union Meeting 2008 Proceedings, number P168-P002, 2008.

We agree with the reviewer that comparing the static strength of a CV carbonaceous chondrite to the estimated impact strength of hydrated carbonaceous chondrite is imperfect.

However, we found it to be one of the few peer-reviewed work that adequately described the testing of the static strength of any carbonaceous chondrite. We would ideally compare against the measurements of Tsuchiyama et al. (2008), but we hesitated as that work appears in a non-referreed conference abstract. Though, their results are cited in peer-reviewed articles that have aggregated physical property measurements of meteorites such as Ostrowski & Bryson (2019), and more recently Pohl & Britt (2020).

For example, Ostrowski & Bryson (2019) tabulate the following information for carbonaceous chondrites:

[tables 3 and 4 from Ostrowski & Bryson (2019) redacted]

Of the 6 unique references in these two tables:
Miura et al. (2008) and Tsuchiyama et al. (2008) are non-refereed conference abstracts.

Brown et al. (2002) and Svetsov (1995) estimate the strength of the pre-impact fireball using ablation models.

Jenniskens et al. (2012) provide no information on the methodology for measuring the compressive strength of Sutter's Mill.

Cotto-Figueroa et al. (2016) is the only peer-reviewed publication that provides a clear description of the methodology used to measure the static strength of Allende.

We would have liked to compare against the results of Tsuchiyama et al. (2008, 2009) – but it is difficult to trust the provenance of these numbers as the work is not peer reviewed.

In order to make an “apples-to-apples” comparison, we decided to compare our estimated strength to those of measured by Jenniskens et al. (2012) for the CM2 Sutter’s Mill, and Brown et al. (2002) for Tagish Lake.

The main text now reads (the text is broken up into segments here for clarity):

“We can then use our Q_D^* *estimate to determine the size-dependent impact strength, , of monolithic C-complex objects (see Eq. (M28)).*

The size dependence of strength is a consequence of the increase in the number and size of internal cracks and flaws, which limit the strength of a material, with the size of an object. Strength is a measure of a material’s ability to withstand a particular stress, such as compressive, tensile or shear [19]. An object’s response to an impact, and therefore its crater formation mechanics, is dominated by one of these strengths.

The formation of well-defined deep craters, as observed on Bennu’s boulders, is typically dominated by shear [19] or compressive strength [11]. In contrast, impacts onto brittle material lead to shallow spall craters formed by tensile failure.

For impacts with U *= 5 km/s and* $\mu_G = 0.33 - 0.36$*, a 1-m-diameter boulder on Bennu has*

$Y = 0.44 - 1.70$ *MPa, which may approximate its shear strength and/or compressive strength.*

The lower range of these estimated values of Y *are comparable to the tensile strength (inferred from high porosities) of boulders on Ryugu, which ranges from 0.2 to 0.28 MPa [20].*

Our measured impact strength is lower than the 85 MPa compressive strength reported for the CM2 Sutter’s Mill meteorite [21], but similar to the 0.25–0.7 MPa compressive strength estimated for the ungrouped C2 Tagish Lake fireball [26].

Weakly cemented basalt and gypsum have similar μ_W *values of 0.46 [22] and 0.4 [23], respectively. Furthermore, gypsum has a similar tensile strength of 0.3 MPa; however, impacts on solid gypsum form deeper craters than those we measure on Bennu’s boulders [23].”*

Line 240: “To assess the impact history of Bennu’s boulders, we analyze the population of boulders with impact craters measured in the Orbital A and B data set..” Again, I question

why the division of data sets into orbital periods. It seems to be that A and B are higher resolution, so you can see smaller craters? Why not just say "used highest resolution imaging to assess impact history…"? The reader does not really care or need to know about O-Rex operational phases.

As previously stated, we remove all mention of mission phases. This part of the main text now reads:

"In the higher-resolution image dataset, which resolves multiple craters on a given boulder, we measured 367 craters with diameters between 3 and 50 cm on 36 boulders with diameters 0.5 to

2.5 m. These boulders have a total surface area of 160 m^2*."*

Line 252: "As the measured CSFD exponent is similar to that of NEOs [25], the average impact speeds were ~20 km/s [24]." Just to be clear….you are not observing impact speeds of 20 km/s, you are assuming them? If that is so, please state that you are making that assumption.

We thank the reviewer for catching this error. We have updated the text to now read:

"As the measured CSFD exponent is similar to that of NEOs [24], we assume that the average impact speeds are ~20 km/s [30]."

Reviewer Reports on the First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

The authors have responded reasonably to my questions and comments and the revised version of manuscript looks fine to me except for the following three points.

1.Cratering age of boulders.
Line 41 and Line 178.
The age of 1.75 Myr is based on the assumption that all the craters on boulders were created during Bennu's residence time in near-Earth space (Line 160-161). The assumption is based on the similarity of slope of the cumulative size frequency distribution of the crater diameter with the slope of those of larger near-Earth objects. This assumption is possible and acceptable, however, can we deny the possibility of crater formation, at least partially, during Bennu's

residence time in the MBA? In this sense the derived age of 1.75 Myr is the maximum.

2.Q* of CM2.
Line 115-116:"we find that 1-cm radius targets have a Q*_D ~ 1.5 x 10^7 erg/g, which is comparable to the Q*_D > 1,419 x10^7 erg/g found for hydrated carbonaceous chondrites [3].

First, the reference [3] does not mention about the hydrated carbonaceous chondrites.The data of CM2 are shown in the other paper published in 2018 (Flynn et al., 2018, Chemise der Erde 78, 269–298). As authors pointed, the detail of the experiments is not described in the paper, though. At least one of the data (of Murchison target) is in a published paper (Flynn et al.,2009, Planet. Space Sci. 57, 119–126),in which the experiment condition is given.
Second, I wonder if there is better way of citation.
The two detailed and similar numbers of 1.5 x 10^7 and 1,419 x10^7 (especially the latter) are misleading given the range of the value of Q*_D of the boulders of Bennu (1-3 MPa according to Fig.3a) and the fact that 1.4 x10^7 is just a lower limit.

3. Strengh of gypsum.
Line 133-135: "Weakly cemented basalt and gypsum have similar values of 0.46 [22] and 0.4 [23], respectively. Furthermore, gypsum has a similar tensile strength of 0.3 MPa; however,impacts on solid gypsum form deeper craters than those we measure on Bennu's boulders [23]”

The gypsum target with 0.3 MPa tensile strength is not the usual one (porosity:50%, tensile strength: a few MPa.), but of very high porosity (65%) and is not mentioned in the reference [23] but in another paper (Nakamura et al., 2015, Planet. Space Sci. 107, 45–52). Referring the example of weakly cemented basalt may be enough.

Minor comment.
L405: reference [23] is wroing.
L622: reference [58] is missing.
L623: typo “!”

Referee #2 (Remarks to the Author):

The authors have done an excellent job of addressing my comments. Publish as soon as possible

Author Rebuttals to First Revision:

We thank reviewer 1 for their final comments on our manuscript. We have responded to these point-by-point.

The authors have responded reasonably to my questions and comments and the revised version of manuscript looks fine to me except for the following three points.

1.Cratering age of boulders.
Line 41 and Line 178.
The age of 1.75 Myr is based on the assumption that all the craters on boulders were created during Bennu's residence time in near-Earth space (Line 160-161).

The assumption is based on the similarity of slope of the cumulative size frequency distribution of the crater diameter with the slope of those of larger near-Earth objects.

This assumption is possible and acceptable, however, can we deny the possibility of crater formation, at least partially, during Bennu's residence time in the MBA?

In this sense the derived age of 1.75 Myr is the maximum.

We thank the reviewer for bringing this point up. It is true that our justification for a record of the NEA lifetime is based on the similarity of the CSFD slope of craters to that of decimeter-to- meter scale bolides entering Earth's atmosphere. The CSFD slope of objects > 100 m in the main belt has a somewhat shallower slope (see Bottke et al. 2005); however, there is no observational evidence as to whether this shallow slope extends to the smaller cm-scale objects that formed the craters (< 50 cm) on Bennu's boulders. In our calculation of the collisional lifetime of boulders on Bennu's surface, we use the modeled main belt SFD at these small sizes.

It may be that the larger meter-scale craters on boulders formed in the main-belt. As we state in the last two paragraphs of our manuscript, where we discuss the relative difference in the collisional lifetime of boulders:

During Bennu's residence time in the main belt, its surface boulders would have collisionally evolved more quickly. Once a C-complex asteroid dynamically and collisionally decouples from the main belt, impact cratering, rather than disruption, becomes the primary mechanism for impact-induced breakdown.

We conclude that the large craters on Bennu's boulders (R_c/R_T > 0.3, for R_T > 1 m) were created by impacts with energies that are close to the boulder's disruption limit and formed during Bennu's residence time in the main asteroid belt.

However, we do not use these larger craters when deriving the NEO lifetime.

2.Q* of CM2.
Line 115-116:"we find that 1-cm radius targets have a Q*_D ~ 1.5 x 10^7 erg/g, which is comparable to the Q*_D > 1,419 x10^7 erg/g found for hydrated carbonaceous chondrites [3].

First, the reference [3] does not mention about the hydrated carbonaceous chondrites.

The data of CM2 are shown in the other paper published in 2018 (Flynn et al., 2018, Chemise der Erde 78, 269–298).

As authors pointed, the detail of the experiments is not described in the paper, though. At least one of the data (of Murchison target) is in a published paper (Flynn et al.,2009, Planet. Space Sci. 57, 119–126),in which the experiment condition is given.

Second, I wonder if there is better way of citation.

The two detailed and similar numbers of 1.5 x 10^7 and 1,419 x10^7 (especially the latter) are misleading given the range of the value of Q*_D of the boulders of Bennu (1-3 MPa according to Fig.3a) and the fact that 1.4 x10^7 is just a lower limit.

We thank the reviewer for clarifying this pointing out that we had mistakenly included an incorrect reference.

We have updated ref [3] to the Flynn et al. (2018) to Chemis der Erde 78, 269-298, as that was the intended reference.

In order to clarify the comparison between our estimated values of Q*D and those of ref [3] laboratory experiments, we give a range for Q*D for Bennu's boulders, and indicate that the values for the hydrated carbonaceous chondrites are a lower limit:

Furthermore, we find that 1-cm radius targets have a Q_D^* = 1.1×10^7 – 3.0×10^7 *erg/g, which is comparable to the experimentally determined values of* Q_D^* *for hydrated carbonaceous chondrites, which have a lower limit of* $Q_D^* > 1.4 \times 10^7$ *erg/g.*

3. Strengh of gypsum. Line 133-135: "Weakly cemented basalt and gypsum have similar values of 0.46 [22] and 0.4 [23], respectively. Furthermore, gypsum has a similar tensile strength of 0.3 MPa; however, impacts on solid gypsum form deeper craters than those we measure on Bennu's boulders [23]" The gypsum target with 0.3 MPa tensile strength is not the usual one (porosity:50%, tensile strength: a few MPa.), but of very high porosity (65%) and is not mentioned in the reference [23] but in another paper (Nakamura et al., 2015, Planet. Space Sci. 107, 45–52). Referring the example of weakly cemented basalt may be enough.

We thank the reviewer for pointing out this inconsistency in our citations. We have included the reference to Nakamura et al. (2015), when making the comparison to the value of \mu_s (previously Lines 101-103); however, at the suggestion of the reviewer, we have removed the specific discussion on the crater morphology and tensile strength of very porous gypsum.

"This value of μ_4 *is slightly larger than that determined from laboratory impact experiments into weakly cemented basalt and highly porous gypsum, which have* $\mu_4 = 0.46$ *[15] and 0.4 [16], respectively."*

Minor comment.
L405: reference [23] is wroing.

We have fixed this reference to [6].

L622: reference [58] is missing.

We have fixed this reference to [44].

L623: typo "!"

We have removed the equation in parenthesis.