Manuscript number: esurf-2021-17
Authors: Sophie Bodek and Douglas Jerolmack
Title: Breaking down chipping and fragmentation in sediment transport: the control of material strength
Date: 6/20/2021

Responses to referee comments

*Below we deliver a point-by-point response to the two reviews of our paper. Reviewer #1 had a very positive review of the paper, and noted that addressing their comments would not require any additional experiments or major analysis. We have addressed each comment, and in almost every case, made changes and/or clarifications requested by the Reviewer. Reviewer #2 was decidedly more critical. Many of their critiques were substantive, fair and useful; we have done our best to respond to these points, either by improving our analysis and writing or by adding caveats and clarifications to assist the reader in interpreting our findings. Some of Reviewer #2's comments indicate a fundamental difference in philosophy or direction that cannot easily be reconciled with revision of a single paper. Nonetheless, we believe that our thorough reanalysis of the data, and the new alignment of our experiments with a very recent numerical simulation paper that reproduces our findings, should alleviate many of the Reviewer's concerns. The Reviewer's critiques about experimental protocol and control are reasonable, but we don't agree with the conclusion that these questions undermine the entire premise of the paper. The new re-analysis of the data follows most recommendations of this Reviewer and does not change the conclusions. In addition, we have added text to support that the contention that the mechanisms of fragmentation and chipping are rather generic and insensitive to details of material properties. The main objective of this paper was to experimentally demonstrate, for the first time, the proposed conceptual model [Ghadiri and Zhang, 2002]: that the transition from chipping to large-scale fragmentation is a continuous phase transition. This point is significantly bolstered in the revised version thanks to a new preprint (**https://arxiv.org/abs/2105.12198v1**) posted of high-fidelity numerical simulations, that find the same behavior that we observed in our experiments. Moreover, we demonstrate how changing material strength can drive such a transition, and make the implications for sediment transport clear. We do not argue that researchers can take our experimental findings and apply them directly to some field scenario; but we directly show how weak materials can disintegrate while strong materials can chip, which are fundamentally different modes of failure with important consequences for weathering and sediment production. Since none of the existing models for bedrock erosion/pebble attrition account for these different modes of failure, our findings suggest a re-examination of those models. A last general comment is that some of Reviewer #2's objections to the paper were opposite those of Reviewer #1. In such cases we acknowledge that we chose to follow Reviewer #1's suggestions.*

Referee #1: Alexander Beer

General comments:

The authors present a well-structured laboratory study on particle attrition do to impact energy during transport (applicable to gravitational and fluvial transport in surface processes). They specifically address the transience between chipping of small parts from and fragmentation of the whole particles for constant impact energies, but varying strength of artificial rock material. For a phase space of weak to hard material (applicable to natural rock), they delineate how hard rocks turn slowly round by chipping, weak rocks fast disintegrate and less round by fracturing, whereas the rocks of intermediate strength show major variability in both process' interactions. These results are well applicable to quantitatively interpret natural transport environments.

This work addresses a fundamental question in studying earth surface processes and well exams it – it is sound and right-suited for ESURF. I was intrigued to thoroughly reading the manuscript and in detail noted where I had problems to follow or where additional work is necessary to clarify issues. Generally, addressing these topics should be manageable by a thorough review of the manuscript by the authors, since no additional data nor major analysis is requested. I thus hope my comments are useful, instructive and are received as constructive criticism for the authors to improve readability.

Best,

Alexander Beer, Uni Tuebingen

We thank the reviewer for their thoughtful and insightful comments, which are addressed below:

Specific comments:

The term "particle" sounds misleading (to me), it always reminds me of natural grains etc. You could better use the term "sample", or "specimen", since you are dealing with artificial particles.

Particle is a rather generic term used in engineering, physics and geoscience. Indeed, in most other disciplines, the 'particles' are artificial. We believe it's best to keep the word 'particle' because of this precedence, but make clear in the introduction now that our particles are artificial.

In Fig.9 you show numbers of particle properties that you mention earlier – better have them in the table earlier or reference them to easier find that.

Figure 9 is referred to earlier in the manuscript. It is mentioned at the end of section 3.1, which discusses material properties for the different concrete compositions.

Part of the introduction is quite long (chapters 1.1 to 1.3); maybe be condensed for a more balanced general picture. The conclusions repeat a certain number of results and discussion thereof; they could more address the topics named in the introduction – e.g., how to determine past transport environments (fluvial, gravitational, planetary).

The introduction has been dramatically shortened and streamlined. The subsections have been eliminated and this section now builds in a straightforward and non-redundant manner toward *the point of our experiments. Also, the Discussion and Conclusions have been rewritten to revisit central concepts in the Introduction and make our findings more clear. Thanks for the suggestions.*

P1 L21: How does grainsize (specifically in a grainsize distribution) fits into the schematic understanding used here (and sketched in Fig.1)? Maybe it fits for the mean grainsize and specifically large grains are exempt?
The most direct effect of grain size is mass, through the collision energy. In fact, energy and material strength are the two most fundamental controls on whether particles chip, fragment or don't do anything at all. We have actually changed Figure 1 so that it reads 'Energy' (rather than velocity, which was the original variable in the Ghadiri and Zhang paper), and explained in the caption that grain size (through mass) and collision velocity contribute to this energy. Grain size may have an indirect control through yield stress, since yield of a sample does depend on its size (larger samples are actually weaker in a load test because of a higher probability of encountering flaws/initial cracks); but this is not relevant for our experiments since we measure the yield stress of samples that are the same size as those used in the experiments.

Fig.1: What about strong material and high impact energies? Is this for brittle material only?
For a very strong material, there will be no mass loss even from infinite collisions. This is represented by the "no breakage" area on the right. For very high impact energies, fragmentation occurs from one collision. This corresponds to "fragmentation" on the left of the plot. We did not make this clear because the y-axis was not labeled with any numbers; we have changed this so the plot now shows that the intercept occurs at N = 1 collision, not N = 0 as one might assume without any axis label.
This is not for completely brittle materials. The conceptual model was actually developed for semi-brittle materials with significant plastic deformation, and applied originally to impact attrition of more ductile salts and also more brittle ceramics.

P2 L10: might also find a reference for fluvial bedrock erosion
As far as we know, Hertzian fracture has not been proposed for fluvial bedrock erosion models. However, we have changed this portion as part of the sweeping revisions of the Intro.

P3 L10: Doesn't rock weathering also contribute to rounding particles? Rounded granitic grains might even become angular from transport when fragmenting. Don't know if this is of large contribution or only a specific case, but I would think discussing the influence of weathering in few sentences might round up the general picture.
*Fair point: **Most** fluvial pebbles are "born" from fragmentation of rock, typically because they are delivered to a stream by landslides or debris flows. So, even if they weathered to some kind of rounded shape on a hillside, these boulders are rarely (though occasionally) found in rivers. This is discussed in some of Jerolmack's previous papers, but pointing out the particular case of potentially rounded weathered boulders worthwhile. We have added a sentence and reference to this.*

Eq1: Why do you give this law in volumetric units, when you use it in mass units in fig.6?

We have removed this from the revised Intro as it was extraneous to our findings.

P4 L13: In my understanding this is the fracture toughness, right? This, though, all applies to brittle rock – I see you mention this in L17, but
In principle yes. However, measuring fracture toughness can be done a few different ways, each with their own drawbacks. Here we are trying to write conceptually for a reader that does not have a background in fracture mechanics; in order to avoid confusion between the concept of fracture growth (and fracture toughness), and measuring fracture toughness (and the associated controversy with that), we did not want to use the term here. Regardless, the revised Introduction no longer contains this line.

P4 L17: The actual fracture is "plastic" deformation; your sentence sounds confusing.
Indeed, yield is associated with a departure from the linear stress-strain relation, which by definition is plastic. The "purely elastic" should really be phrased as "brittle" as opposed to ductile. Regardless, the revised Introduction no longer contains this line.

P5 L18f: These are methods and results and should go their sections; maybe formulate a question on these things here.
We would prefer to leave this summary at the end of the Intro section here. It is common to have a high level summary at the end of the Intro, so the reader gets an overview before plunging down into granular details. In the past I have been asked (several times) by reviewers to include such a summary.

P6 L8: You could cite Sklar and Dietrich, 2001 here for creating these samples of different (tensile) strengths.
We have added sentence indicating that use of concrete as a synthetic, brittle rock is common in experiments of fluvial bedrock erosion and its mechanical behavior compares well to natural rocks [Sklar and Dietrich, 2001; Johnson and Whipple, 2007 (https://doi.org/10.1002/esp.1471)]

P6 L15: There was only one particle in the drum each time, right?
Yes, that is correct. Text has been clarified to indicate that only one particle was in the drum at a time.

P6 L19: How were the particles photographed – always orthogonal to one of their original plane sides for consistency?
Particles were photographed perpendicular to the initial plane side as long at the original side was distinguishable. This became more difficult at the particles became more spherical or broke down. We clarify this in the text.

P6 L24: Hard to understand, why synthetic CIRCLES should not have been recognized as circles. You mean you created pixeled or distorted ones? What was the resolution of the camera setup (or better the ~pixel resolution on the particles) you used for Fig.4A?

Synthetic circles were not distorted, but did have pixilation (even if at a very small scale). This pixilation results in up to a 10% error when measuring the circularity of ideal synthetic circles. We also mention the resolution of the camera setup – 12.2 MP.

P6 L28: I think you used the SAME 10 mixtures?
Yes, the same 10 mixtures were used in both the rotating drum experiments and Instron experiments. This has been clarified in the text.

P7 L6ff: Mega-Pascal (i.e. N/m2) is a standard unit for rock strength. Can you provide a range of typical rock's values to relate your laboratory material to the field?
We note that the cement with no sand is something like a limestone, and the sand with very little cement is like a weak sandstone. We have added this idea into the paper, to make some connection of our lab samples to geological materials. We also indicate in the text that our measured values for cement mixtures are comparable to those determined by Sklar and Dietrich (2001).

P8 L 19: Which "associated material properties"? Name them. You also mean best-fit lines from Figs5b and c?
"Associated material properties" refers to the ultimate strength and Young's Modulus for each different type of sample particle. However, this line has been removed from the revised text.

P8 L24: "Mass fraction" is not really intuitive – do you mean remaining mass fraction or so?
"Mass fraction" is defined and used consistently. The term "mass fraction" was defined in early papers by Domokos (back to 2012) so we use the same name for consistency.

Fig.6: You here set NR (which is not NR, but a normalized version) equal to x in eq.1 and also have a volumetric vs. a mass equation – settle and explain your reasoning and steps. Add R2 to the Sternberg law fits in panels (ace).
The emphasis on Sternberg's law has been reduced in the revised version. As an explanation – Sternberg's law is typically written in terms of volume, but we use it here in terms of mass because it is the more fundamental variable for attrition (in relation to energy). The equating of NR (or Ni in the revised version) to x is based on an assumption of one rotation equaling one particle hop. In the field, x is the measured variable but it is actually a stand in for the number of collisions, which we expect to increase linearly with x. So, the two should be proportional but not identical, which is why the main point in the figure is that the expected exponential scaling is observed; this has been clarified in the text. In the figure, the x-axis label is corrected to indicate rotations/impacts are normalized; R2 values are added.

P9 L1: In fig.7 are violin plots, not histograms. Though, here you should describe what data you calculated to then show in fig.7.
Correct, thanks. They are violin plots, which means a kernel density estimator has been used to smooth the data. We correct this, and clarify in the caption that what was measured was the fractional mass loss per impact for all impacts; this is a violin plot that shows the distributions of all measurements for each material.

P9 L6f: Sklar and Dietrich's relation refer to erosion of bedrock due to impacting particles and not to the attrition of particles by mutual collision (or collision on bedrock)!
Indeed! However, the rationale they develop for mass lost per impact, and how it should depend on material properties, applies equally to the target and the impactor (if they are both made of the same stuff).

How does a plot of attrition mass vs. energy/strength actually look like (you can plot that)?
This is plotted in figures 7 and 8. In Figure 8, the mass loss parameter k is the empirically determined number that relates mass loss to collision energy via equation 1. It relates to the parameter Ab via a simple scale factor. So, Figure 8 is a plot of the attrition mass loss vs. strength.
In this experiment we do not tune collision energy; the initial collision energy at the start of experiments is more or less fixed. But we indirectly account for changing energy because as the particles undergo attrition their mass (and hence energy) diminishes. Our attrition parameter, change in mass/collision energy, accounts for this as the particle loses mass. Hence, we are isolating the material control.

Would be very interesting to see how this would fit into bedrock erosion theory (specifically, since one may speculate on combined bedrock erosion and particle attrition during a flood or so)!
Indeed. But we think that this is mostly about choosing initial and boundary conditions and is beyond the scope of this paper. The Sklar and Dietrich theory (and variants of it) includes flux, saltation trajectories, etc. for the full treatment of erosion. But one element of it is that mass loss follows a single function of material strength and also of collision energy. Our data, and the new simulations by Pal et al. (2021), show that this is not the case.

I address this, since you show a largely increased particle attrition for increasing rock strength (Fig.8a), which would mean they decay fast and thence would faster lose their impact erosivity.
Well, the attrition number has tensile strength on the bottom; so, larger attrition number means weaker material.

Would your results (Fig.8a) mean, that with increasing rock strength particles decay fast, so would not contribute much (or better not for long) to bedrock erosion?
See point above. WEAK rocks break apart quickly and won't contribute much to bedrock erosion. Strong pebbles will round without breaking.
We have edited the text in many places to remind the reader that larger Ab means weaker material.

Also, I don't get your regressions misfit reasoning in Fig.8a – here you refer to k = AbC1, but Ab depends on Y and rho (eq.2) and the Sklar and Dietrich relation only depend on sigma_t. Hence it seems you compare different things – please better explain your steps and reasoning here.
We have removed the plot of tensile strength, and the Sklar and Dietrich prediction, from the plot. In other words, we removed the previous figure 8a. The second reviewer objected to this comparison since material properties were measured differently. In addition, Sklar and Dietrich

(and following papers) did not do experiments of single particle impacts to isolate the material control alone; so the inferred dependence from their data should not be directly compared to experiments like ours where single impacts are isolated.

P9 L11: Your formula assumes (mi-1) ~ m, right? Can you elaborate this?
It is unclear to us what is meant by this comment. In general, we measure particle mass frequently enough that the change in mass between collisions is small compared to the mass of the particle. But this is not always the case; weak particles fragment significantly, which by definition means that the change in mass can be a significant fraction of the mass of the particle. Either way, there is no assumption (or need for one) that the change in mass is small compared to the overall mass.

P9 L14ff: Split and better declare the content of this sentence. What is the reason for C1 to be 3 orders of mag smaller for natural particles (weaknesses)?
This portion of the results section has been edited and no longer contains that specific line. Since this is the Results and not Discussion section, however, we do not want to speculate here what reasons there might be for the difference. We do address this in the Discussion (see below). The significantly revised Discussion now centers on the bi-partite relation between mass loss and material properties; this means that it's not a linear relation over the entire range, and therefore the coefficient C1 is not really relevant. We addressed this in more depth in responses to Reviewer 2 below.

P9 L15: The independence might be seen in Fig.7 (and Fig.6), but how in Fig.8?
We have significantly clarified and bolstered the case for two regimes of mass loss, where for strong materials the weak control of material properties is now better seen. We have also discussed at length a new numerical paper (cited above) that found the same thing.

Fig.7: I think the caption should be delta_m* vs. Ab. The distribution is not an average. What is the y-axis – the same as in Fig.6? You could indicate your explanations from the caption also in the figure: strong vs. weak for the x-axis. A second transition for small Ab is not obvious. Further, how the mass of the removed fragments looks like (can show this as a second panel)? This would also give a quantitative feeling relating to your description of chipping vs. fragmentation.
Caption has been modified. We have removed mention of a "second transition" at small Ab. We did not directly measure the mass of the fragments themselves. We only measured the mass lost from the main particle. For strong particles, the mass distribution of the pieces broken off likely is close to the distribution of mass loss from the main particle. For weak particles, however, the 'fragments' broken off would often beak apart themselves into loose sand, since the cement content was so low; measuring the mass distribution for this material would not have been meaningful. We added text describing this in Results and in Discussion.

Fig.8: To which data is the regression for Sklar and Dietrich in (a) fit to – the black data? Is the parameter b in the cyan equation equal to 0.838? Where is the data in (b) from – which material in Miller and Jerolmack, 2020? I don't easily see how you got the value for C1, only

that it is three orders of magnitude larger than for this data – you got it from (a) with solving the k-equation for C1? Then you should describe this in the text, not in the figure caption.
We removed the panel with the Sklar & Dietrich relation (Panel a), and corrected the error in C1 value. Text was changed to highlight bi-partite relation of mass loss and material strength. Since it is not a single linear relation, C1 is no longer reported and discussed in the text.

P13 L3: Having a low amount of sand in the particles (i.e. high strength) means a very fine texture (mostly Portland cement) here. This is different to a hard rock that though consists of large minerals, like a Granite, i.e. there might be a difference in fracture toughness. Is this issue addressed/settled in the natural particle attrition studies?
This was a most interesting finding from the Miller and Jerolmack (ESurf Discussions, 2020) paper that examined natural rocks undergoing binary collisions. The mass of fragments produced by collisions was essentially the same for the different rock types, regardless of the size of grains making them up. They tested materials including granite, sedimentary rocks, bricks and limestone. It seems that the chipping process can and does break grains/crystals, and that mass loss in that chipping regime can be related to bulk material properties without considering grain/cystal size – at least, to first order. One may think of our samples here as spanning two end members: the cement with no sand is something like a limestone, and the sand with very little cement is like a weak sandstone. We have added this idea into the paper, to make some connection of our lab samples to geological materials.

P14 L1ff May the weaker particles have had (more) initial cracks from curing?
Likely all of the samples did. This, however, is also true of natural rocks. Either way, it is likely that each stress drop is related to crack formation or growth, but that those cracks did not percolate through the sample and therefore catastrophic failure did not occur. We have added to the Results and Discussion that heterogeneity in the samples likely led to a distribution of inherent cracks in particles at the start of experiments, and that stress drops correspond to formation or activation/growth of cracks.

P14 L22: Maybe at low Ab, there is only chipping on sharp edges and where fracture toughness is more important than strength?
In Miller and Jerolmack (2020) we found that, at the very beginning stages of attrition of cubes with sharp edges, that mass loss was independent of particle properties and was constant for all of different natural rock samples (and bricks). This was indeed related to a geometric effect: when edges were sharp enough (radius of curvature approaching zero), any collision could locally exceed yield because local pressures were very large. However, this effect went away once very sharp edges became rounded.
In this paper, our numbers come from measuring particles over the entire stage of rounding, and even the hardest particles undergo significant rounding. So, we do not think it's the geometric effect that we saw in the Miller and Jerolmack paper.

P15 L2: Are >50% VCM at Ab<0.014? You could indicate both regimes (pure chipping vs. pure fragmentation) in Fig.12 (maybe showing a gradient)

The new analysis changed the numerical values of Ab, but not the general patterns. We now refer to particles by their value for Ab rather than VCM, for clarity.

P16 L18f: So you could even show a 3D regime with x=material strength, y=impact energy, z=impact number, extending the 2D-version of Fig.1?
One could, yes. However, it is our hope that we showed here that increasing material strength has the same effect as decreasing collision energy. Therefore, we would prefer a 2D representation where one axis is a product or ratio of impact energy and material strength, and the other is impact number.
We have added in the Discussion the idea of the Basquin law of fatigue failure, and that the equation of the attrition mass loss model (old equation 2, new equation 1) equates energy and material, and therefore that the presentation in old figure 13 essentially demonstrates the Basquin law.

P17 L2: That would mean that the attrition regime is predicted by a mass (delta_m; eq.2)? I don't get this.
Yes! Actually, the attrition regime is DEFINED by the mass lost per collision. Or, alternatively, the number of collisions required to make a larger 'fragment' rather than a small chip. If the mass lost per collision is very small (or the number of collisions to make a 'fragment' is very large), we are in the limit where chipping occurs.
The Reviewer's comment caused us to notice something that we hadn't before: in the strong material limit (low Ab), the number of collisions required to make a large 'fragment' does not change with material strength. This is apparent in the flat line on the left of the plot in Figure 11 (previously Figure 13). This is also the regime where the mass lost per collision does not appear to change much with Ab.
We note that our dynamic range of Ab in THESE experiments is actually larger than the observed range in Miller and Jerolmack (2020) for natural rock samples. In that study, they noted that the strongest rock type (a very fine grained 'volcanoclastic' rock) was approaching a condition where the mass lost per collision was nearly zero for a finite Ab (Ab = 0.2). Perhaps coincidentally, or perhaps not, our data here indicate a small and roughly constant mass loss fraction for values of Ab < 0.2. This may be coincidental because material properties were measured using different tests in the present study, so values of Ab may not directly correspond to reported values in Miller and Jerolmack (2020).
The substantially revised Discussion addresses this directly. We suggest that our data show a bi-partite relation, where mass loss changes only gradually with Ab for strong particles (this is chipping), and then increases rapidly beyond a threshold as material weakens (this is fatigue failure). The experiments in Miller and Jerolmack are purely in the chipping regime, and therefore they only observed the first gradual dependence.

Technical issues:
P1 L19: ... is "of" lower energy, ...
This suggestion has been implemented in the text.

Fig.1: I think bedload (or bed-load for consistency) should be bedload transport

While both “bed load” and “bedload” are used in the literature, most researchers use “bed load”. We changed to “bed load” in the figure. If we adopt “bed load”, then we must also write “bed-load transport”.

Fig.1 caption: “a” debris flow and “bedload transport”; what transport mechanism images do you refer to?
We are referring to the row of photographs across the top of the figure. The caption has been revised for clarity.

P2 L8: ..., and “with” the related phenomenon...
The Introduction has been revised and no longer contains this line.

Eq.2 Why do you use the notation “C_1” when there is no C_2? Better use a more descriptive variable name.
No other variables use “C” in the manuscript. C_1 is used for consistency with Miller & Jerolmack (2020), which is mentioned in the text.

P5 L20: Dropped from the same height in a drum? Can you elaborate here (only one grain per experiment hitting the steel frame or hitting other grains)?
SB: The text has been clarified to indicate that there was only one particle in the drum at a time.

P5 L11: ...“sample” particles... This would be clearer
The introduction has been revised and no longer includes this line.

Fig.5: caption: In (b) and (c) better have the (n = 5) at the end of the sentences, respectively. The equations are missing units.
(n = 5) has been added to the end of the applicable sentences. Units are defined in the figure caption.

Fig.6: Panels (ace): Is this “Mass fraction” on the y-axis? As I said above, better use a more intuitive (and consistent) terminology here. Note the VCM-% in the figures; the lines are not visible in the legend. You could show the individual lines in gray (doesn’t matter which is which) and so could improve the trend picture and reduce the legend. The x-axes are not NR, but normalized. Remove the “We can see”.
Y-axis label has been corrected to read “mass fraction”, X-axis label indicates that the number of rotations/impacts are normalized (N_i / N_{total}), and the key has been corrected to show the colors for each particle; “we can see” is removed.

Fig.8: caption: Instead of 8.38e-1 write 0.838 – or is it e to the power of -1?
The equation describing the line of best fit in Figure 8 has been modified for clarity.

P11 L4: What were about extending drum run time - would this have destroyed the weak particles or could you have reached higher circularity?

For very weak materials, the particles disintegrated or became extremely small. So longer run times would likely have destroyed the weaker particles. We clarify in the Methods section that weak particles were rotated until very little remained of the initial particle.

Fig.9: You may write/name in the figure that you show the shape evolution with number of rotations (NR=0 is initial shape).
This has been clarified in the figure caption.

Fig.10: Use a color gradient to highlight increase of particle strength. In the caption, combine the last two sentences. You use "transport" here, but write cumulative mass loss on the x-axis (it actually is relative); before you spoke of rotations. Be consistent and explain, if you move from rotations to something else. In panel (b) you could show the relative change of the aspect ratio relative to the initial conditions (i.e. 1). Combine Figs 10 an 11 into panels to not have a confusion with referencing and figure numbering.
These figures have been removed as they were superfluous and distracted from making a clear and concise point.

Fig.12: Have the increasing VCM as a gradient in color for a clearer picture.
This suggestion has been implemented. The increasing VCM is now illustrated through a color gradient.

Fig.13: Since you refer to Fig.1 wit this, you may also have it in the same x-axis orientation as this (chipping regime on the right). You could also add some of its arrows or so. Call this phase space (as in the discussion)?
We refer to this figure as a phase space, arrows are added, and the x-axis is reversed.

P15 L16: "… shape evolution of PARTICLES in order …". After that you speak of "transport" – have this defined before and be consistent throughout with this (impact, mass loss, transport).
The sentence has been modified to indicate that the shape evolution observations were for concrete particles.

P16 L6: "... that a TRHESHOLD exists..."
The Conclusion has been edited and the revised version no longer contains this line.

Referee #2: Anonymous

This paper explores experimentally the erosion processes of particles during sediment transport, in particular the transition between abrasion and fragmentation. To this end, the authors have carried out a series of experiments in a drum equipped with a paddle that causes a series of drops (about 40-50cm high) applied to artificial particles made of a mixture of variable proportions of sand and concrete.

The experimental approach is not new in itself (the use of drums has been quite classical for more than 50 years to study the abrasion of particles) but the idea of focusing on the transition between abrasion and fragmentation is original.

On the basis of their experimental results, in particular the interpretation of the variability of the mass loss undergone at each impact and that of the evolution of the roundness index of the particles, the authors propose a threshold value of the mechanical strength of the particle material which could correspond to the transition between abrasion and fragmentation.

Some parts of the introduction could be improved. Explanations on calculations or variables present only in the figure legends should be developed in the text and some ambiguities could be clarified. But on general, the text follows correctly, the general idea and the figures are understandable.

In spite of its relative clarity and its potential interest, I see this study more as a preliminary study or as a trial run allowing to set up a more elaborate study in the future, and not as a study sufficiently completed to lead to a scientific article that will advance the knowledge on the tackled problem. I give the reasons for this in what follows by exposing the different conceptual and interpretative problems I encounter in this paper.

We thank the reviewer for the comment, and we understand also that this is a first experimental study of this transition. We have made major revisions to the paper along the following lines:

1. We have redone the analysis to get rid of the regressions used to determine the susceptibility parameter Ab, and now that parameter is computed directly for each material from the measured values for tensile strength, modified Young's modulus and density. We have also added substantial text that indicates our patterns of material properties, despite their variation, are in general agreement with observations of concrete. The conclusions drawn from the material control do not change with this new analysis.

2. Our results are substantially bolstered by a new paper just deposited on arXiv (Pan et al., 2021) in which direct numerical simulations of impact attrition show a continuous transition from chipping to fragmentation that shows all the major features observed in our experiments. We had revised the paper substantially based on comparisons to these new simulations. We have also added many statements to make clear which parts of the dynamics we think are generic (chipping and fragmentation do not depend on the details of the materials), and which *parts are sensitive to details and/or error. The main conclusion is that in the strong material limit, there **should** be a dependence of attrition rate on material strength that we are unable to see in the current experiments due to noise. But, comparing to experiments of Miller and Jerolmack (2020) and the new simulations of Pan et al. (2021), this dependence is small and it is likely we can't resolve it in the present experiments due to error/noise in material properties. But we **do** resolve a bi-partite dependence where, beyond some threshold of material strength, we see mass loss grow rapidly with decreasing material strength. This was also seen in the simulations mentioned.*

3. We have made major efforts in the Discussion and Conclusions to make clear what one can and cannot conclude from our experiments.

Problems with the experimental protocol (probably the major issue):

At the first reading of the paper and its figures, I was struck by the strong dispersion on the strength data, and especially on the density data (between 1.6 and 2.7!). There is very little information on the preparation protocol of the particles and the specimens. Even the composition of the concrete material that will be mixed with the sand is not given (what is its content in cement?). The literature on concrete is more than a century old, and it is well known that many factors can modify the strength of a concrete. In this study only the ratio between concrete and sand is considered. However, the proportion between cement and water (beyond a w/c ratio of about 0.5, the quality of concrete degrades in relation to residual porosity after drying), the proportion of occluded air, the duration of curing before use of the concrete (a minimum of one month is recommended to achieve a certain constancy of the strength value), the granulometry of aggregate, the surface of the aggregate particles. Parametric studies (varying the % of sand, water, etc.) published in the literature generally show a much smaller dispersion of values than that presented by the authors of this study for their data. Also the measured strength values (between 3 and 10 MPa) for mixtures with a majority of concrete (VCM>60%) are clearly lower than the values commonly given in the literature (from 20 to 50 MPa after one month of curing).

We acknowledge that we can provide significantly more detail on the preparation protocol for the concrete blocks we created in the lab. In particular, we have added detail about water content, curing time, and size of the sand added. For sure particles that have cured for a month could be stronger than our particles, but our experimental time was sufficiently short that, at least, it is highly unlikely strength changed during the experiments from curing. Our measured values for strength are consistent with other cement materials created for fluvial bedrock erosion and also overlap with some natural rock materials (Sklar and Dietrich, 2001). They may be smaller than reported values the Reviewer states due to curing time, but also because "normal" cement includes "aggregate" (gravel) in addition to sand. Sand/water/cement mixtures are more like "mortar" than "concrete", in that they are weaker, but they are still brittle solids.

We also agree that the dispersion in the strength data is large, and acknowledge that a more careful and consistent preparation protocol could have reduced that scatter. We do not believe that this is disqualifying, however, for two reasons. First, natural rock samples are heterogeneous and have differences in pre-existing flaws, such that they also exhibit a wide dispersion in strength measurements. This dispersion is the reason that we analyze multiple samples and average them. Second, the shape evolution and mass loss data still show trends consistent with each other and with the notion that there is a 'pure chipping' regime, and then a continuous transition toward fragmentation – and, that this can be controlled or tuned with material strength. The trends would likely be even cleaner if efforts were made to reduce the spread in material strength; but we believe they are already clear enough to confirm the main hypothesis of the paper.

At this stage I can only speculate on the origin of such a dispersion in between measurements on the same mixture, or between mixtures with nearby characteristics. Considering the strong variations of density and considering that the average density here is lower than that of a well-made concrete (density from 2.2 to 2.4), I imagine that occluded air is a dominant issue here, and that from one mixture to another the quantities of trapped air have strongly varied. It is possible that strong variations of the water/cement ratio are also at the origin of small bubbles and strong variations of density. In the first case in particular, the presence of large bubbles that vary from one sample to another could explain the very variable resistances. But it is also possible that the curing times were not respected, which poses a problem for example if a specimen and a particle were prepared jointly but were passed to the press and in the drum several days or weeks apart.

We have added more detail about our preparation protocols. It seems likely that air bubbles were contained in the samples since we did not shake/compact them as is sometimes done. Surely, variation in air content could explain part of the variation among samples. We don't disagree that this may make greater scatter in strength, but there is still a large dynamic range in measured strengths that exceeds the dispersion. Additionally, a sentence was added to indicate that particles used in the drum and in the Instron were prepared at different times following the same methods (with some variations in water content and curing time). We have also added very explicit caveats in the beginning of the Discussion, with recommendations for what should and should not be interpreted based on our observations.

Another source of error seems to be related to the preparation of the specimens for the uniaxial press. It is fundamental to obtain quality measurements to have specimens with smooth and parallel top and bottom surfaces. If the specimens have been passed as they were when they came out of the mould, I can see that this can create an additional source of dispersion and also explain why the loading curves in the press are difficult to use and not very suitable for calculating a Young's modulus.

Sentence was added to indicate that attempts were made to have smooth and parallel surfaces in contact with the Instron (although this was not always achieved).

Perhaps the authors have taken all these "difficulties" into consideration, but if so, they should specify this and be able to explain why the density or strength data are so scattered, and why the press measurements are of poor quality. If not, I strongly urge the authors to repeat their experiments, making sure to produce a bubble-free concrete, to respect the optimal or constant proportions of water vs. cement, to let their concretes undergo a minimum of one month of curing before using them, and to rectify their specimens before doing strength test or runs within the drum. Having a reproducible preparation protocol seems to me to be necessary to be able to answer the question asked without bias and without approximation.

We respect the reviewer's clarity and desire for rigor. To be perfectly honest, this project began as an exploratory study but then we were impressed with the resulting trends in terms of a 'pure chipping' regime at high material strength that behaves precisely as observed in natural rocks, and a continuous transition to fragmentation as hypothesized (but never before observed in experiment) by Ghadiri and Zhang (2002). Because none of the existing models for bedrock erosion (or pebble attrition) account for these different modes of breakage we believe the result is significant – not because of the exact numerical values of mass loss or strength, but because we observe a transition in the failure mode. Again, being honest, the student has moved on and it is unlikely that anyone in the foreseeable future will repeat these experiments. Our opinion is that it is better to put out this new finding, and have other researchers build on and refine it, then to not put it out at all.

Problems of estimation and taking into account the errors of σ_U and A_b

In an attempt to overcome the large dispersions in strength, the authors replace the measured values with those from a linear regression passing through the middle of the points. This approach is not appropriate for three reasons:

1) If the specimens and particles are prepared with the same mixture (with the same water/cement ratio, or the same quantities of occluded air) then the abrasion resistance of the particles should be related to the mechanical resistance measured in the press, and not to a value interpolated from other mixtures with distinct preparation biases.

2) If the dispersion is due to a variable curing time, then depending on the time of preparation, the order of passage of the particles in the press and in the drum, the biases may not be corrected by the choice of an average value.

3) by considering a linear regression, the authors implicitly consider that a linear relationship exists between the VCM and the ultimate strength. However, the mortar experiments I could find in the literature (Singh et al., 2015; Bu, Tian, Zheng et al., 2017) show that strength (both compressive and tensile) is not a monotonic function of the sand proportion. Bu et al. indicate that for reduced content in sand (<66%) the strength increase with the sand content, whereas Singh et al. describe an opposite trend for sand content >75%. Similarly but for concrete mixed with aggregate, Stock et al (1979) also describe non-linear trends. In other words, wanting to pass a straight line has no experimental or necessarily physical reality. Wanting to pass a linear

relationship will not reduce the noise on the data, it could instead add error and systematic bias.

The comments related to dispersion in the measurements are similar to above. The comment regarding using linear fits to tensile strength and Young's Modulus to compute Ab is new, and this is a valid critique. It is likely obvious, but the particles measured for strength in the Instron are NOT the same particles that were used in the experiments since both are destructive procedures. So the strength of particles in the drum could have been within the range measured for the 5 tested in the instron for each VCM… or could have been different. Nonetheless, we have removed the fits from the analysis and simply compute Ab for each VCM directly from the averaged measured values of sigma_m, Y, and rho. While there is scatter, the main result is unchanged.

Within the scatter of our measurements, it is possible that there is some non-monotonicity of strength and sand content – but likely not that strong. The literature is mixed in terms of the effect of changing sand content on the strength (tensile, compressive, etc.) of mixtures of cement+sand+water. As the author notes, these mixtures (that do NOT include gravel "aggregate") are typically considered "mortar". Some papers indicate a non-monotonic relation of sand content and strength, but others have not. One recent paper (https://link.springer.com/content/pdf/10.1007/s11595-017-1607-9.pdf) found an increase in mortar strength with increasing sand content (opposite our general trend); however, this change in strength was small (about 20%) compared to the range of strengths we make in our samples. A big challenge is that few engineering studies report strength measurements over a the full range from 0% to nearly 100% sand, since mortar/cement usually uses a much narrower range due not only to strength considerations, but also to 'workability'.

Our summary for all of this is that we now simply compute Ab from MEASURED values of sigma, Y and rho rather than any fits. This means we are no longer using any forced relation (linear, monotonic or otherwise) but simply using our combined strength parameter from measured values. Whether or not the computed Ab from measured parameters has a monotonic relation with cement content, or not, therefore is no longer relevant. Nonetheless, we indeed observe some non-monotonicity of strength parameters with cement content that is in qualitative agreement with studies from others. We followed the Reviewer's advice and now simply compare the mass loss parameter k to Ab directly. We also added discussion in the Results of the trends for the material properties, how they compare to other studies with cement, and the sources of error.

This issue is important: the authors insist that the mass loss parameter becomes independent on the mechanical strength for the most resistant particle. However, the basic observation is that the mass loss parameter becomes independent on the %VCM, and authors' conclusion is fully dependent on the assumed increasing relation between %VCM and strength. If this increasing relation is not verified for VCM > 50% (and graphically an increasing trend is not really observed) then this whole authors' conclusion becomes pointless.

We see the reviewer's point. Again, we have now computed Ab for each VCM directly from measured parameters (rather than any fit). This does not change the basic conclusion, however. The result is not as dire as implied by the Reviewer. Looking at the more "raw" plot of tensile strength against %VCM (making no assumption about a 'fit'), we can see that all samples with %VCM >= 50% or at least twice as strong as all samples with %VCM < 50%. And when we look at the rounding data (Fig. 12 in previous version, Fig. 10 in current version), we see that ALL samples with %VCM >= 50% follow the same evolution in terms of shape vs. mass loss – which also follows the theory of the Szabo et al. (2018) paper. In other words, all samples with %VCM >= 50% seem to be in a pure chipping regime, and all samples less than that show varying degrees of fragmentation. (We can also see this in the mass loss distributions). So, there is still evidence of a 'limit' of strength and chipping; and that all of the samples that seem to be only chipping are all definitely stronger than the particles that are partially fragmenting.

As discussed generally above, the new numerical study from Pal et al. (2021) substantially confirms and bolsters our observations, since they observe the same behavior. This comparison to their model results is a major change in the revised manuscript, and is so thoroughly woven into the paper that we can't point out all the revisions here. The Discussion section is completely re-written.

In any case, it is fundamental for the figures including σ_U and A_b to add the error bars and to take them properly into account to calculate a regression coefficient.

See above, we have changed things so that we do NOT use regressions for computing Ab.

Fragmentation is poorly defined

The very notion of fragmentation is defined in this paper in an indirect way. As presented, all we know is that particles with a high % of sand show a more irregular abrasion pattern. But is this necessarily the result of fragmentation? One can imagine that these weakly cohesive particles give essentially sand and that the variations are the fact that the particle falls on an angle, a face or a vertex. The video put online on Youtube (https://www.youtube.com/watch?v=UsW8TMxfiqI) seems to me to show during the impacts essentially the production of sand, which seems to be confirmed by the authors page 14 (lines 10 to 14).

Yes, the weak materials had chunks that broke off and often disintegrated into sand. Indeed, we mentioned this in the text. This does not mean, however, that fragmentation and fracture did not happen. First of all, wet and even dry granular materials fracture (https://pubs.rsc.org/en/content/articlelanding/2017/sm/c6sm02600a#!divAbstract), in a manner often modeled using the Mohr-Coulomb criterion (https://link.springer.com/article/10.1007/s10035-013-0477-x). Our weak particles are behaving as a weakly cemented soil or sandstone (https://link.springer.com/chapter/10.1007/978-94-011-4659-3_23), and the mortar mixtures of sand and water are similar to mixtures of sand and polymer (https://hal.archives-

ouvertes.fr/hal-01720762/document) where the strength can be tuned by the polymer/cement content. Whether the chunks that break off disintegrate into sand grains or remain intact, the formation of these chunks is still related to the formation of a fracture/failure plane that connects one point on the particle surface to another. Also, the 'chunks' that broke off often not disintegrate into sand IMMEDIATELY. Rather, large fragments broke off of the main particle, and then got battered by the large particle from repeated drops in the drum; and this battering reduced them to sand. So, at least qualitatively, we can say with some certainty that fragments were often produced – they were not 'explosions' of sand pieces being ejected from the parent particle, but rather fragments that then got pulverized by repeated collisions. Because these fragments got pulverized, however, we were not able to measure the mass distribution of the fragments themselves – only the distribution of mass loss per collision.

More broadly, we examine the intermediate range of strengths in which numerous collisions are required to produce a large 'fragment' (acknowledging, as we do in Figure 11 [previously Figure 13], that the definition of fragmentation is arbitrary and various thresholds may be chosen). There are particles that must collide 10s to 100s of times before a large piece breaks off. Clearly, these pieces are experiencing fatigue failure and crack growth.

We note that the new simulation results of Pal et al. (2021) actually found that in the high energy limit, fragments disintegrate into powder. This also happens in experiments with gypsum balls, which are a model system for dynamic fragmentation. We have made major revisions to the Discussion section to make this comparison to our experiments clear, and to justify that the main results do not depend on those details.

In the introduction, the authors mention the Hertz contact for chipping. The notion of Hertz contact zone is defined for an elastic medium. But here we can imagine that for low resistance particles, during the contact, the whole contact zone will be plastically deformed, fragmented and the important residual kinetic energy will induce a widening of this contact zone and an extension of the deformation/abrasion, until the kinetic energy is completely absorbed. In other words, a very strong abrasion can be localized around the contact zone, leading to a wide surface that is rougher and less round than before (and thus explaining that the shape cannot converge to a sphere) and without fracturing or fragmenting along a fracture that crosses the sample as is usually conceived for fragmentation in natural pebbles.

This may be possible, but it does not seem likely for our case. First of all, even weakly bonded/cemented granular materials are fairly brittle. While our Instron tests sometimes showed several stress drops rather than one, they still did not show any evidence of large-scale plastic deformation that typically stretches the envelope beyond failure. In other words, even weak samples exhibited catastrophic failure. ***While anecdotal, we have examples from images of particles breaking like pebbles do, with fragments that are not pulverized.*** *In other words, the reviewer raises a reasonable doubt; but, if correct, it would require essentially that all large mass removal events are associated with pulverized sand grains detached by essentially creating a yield cone that goes into the material.*

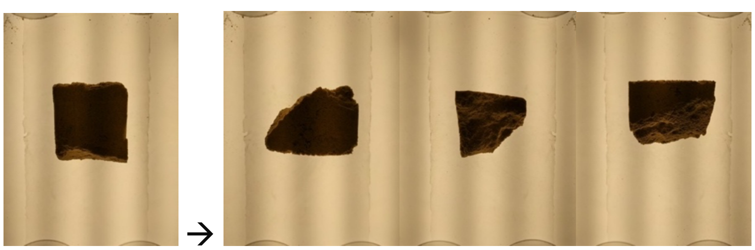
Figure 7 has been modified to illustrate the size of fragments removed from strong, intermediate, and weak particles to indicate the general size and shape of daughter particle from each general material strength. We have also made many modifications to the Discussion section, again related to the new simulations of Pal et al. (2021) and classic gypsum ball experiments.

It seems to me therefore necessary for the authors to present some pictures of the products of what they call fragmentation in their experiments, to present the fragment size distributions, to demonstrate if it is the case that fragmentation in its classical definition (splitting in two or more large fragments, and not a shower of sandy particles) is indeed occurring, and finally to discuss the relevance or not of the behavior of their material to account for the fragmentation process in natural pebbles.

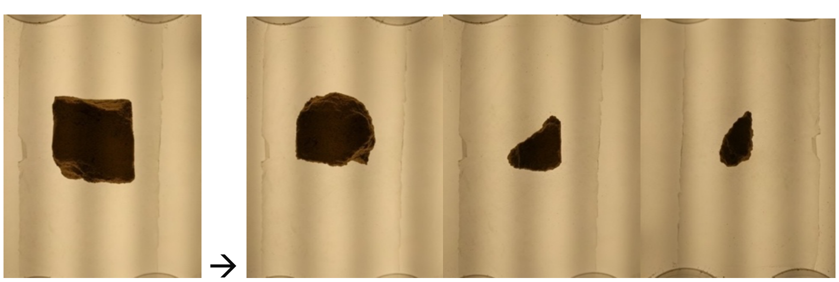
Because many of the fragments did eventually get pulverized by repeated collisions of the parent particle with them after they detached, we did not have a chance to measure the size distributions of the attrition products. However, we do have some images that we can provide.

Below are images showing fragmentation in particles:

1) Example of fragmentation in a 14.3% VCM particle (1st photo shows the parent, 3 subsequent photos show daughters)

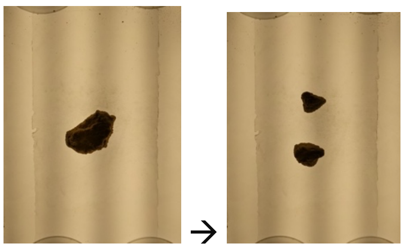


2) Example of fragmentation in a 16.7% VCM particle (1st photo shows the parent, 3 subsequent photos show daughters)



3) Example of "cracks" visible in a 12.5% VCM particle

4) Example of fragmentation in a 12.5% VCM particle (1st photo shows parent, 2nd photo shows daughters)



Attrition product size distribution is illustrated in Figure 7, where we added outlines of parent to daughter particle transition for strong, intermediate, and weak materials.

The discussion is not mature enough

The discussion needs to be reworked and deepened. Experiments and their results are proposed, a relation (fig. 13) is deduced graphically, but nothing is really said about the transposition of these results to natural cases.

- Can the results of experiments done in the open air be transposed to collisions occurring in water?

In short, yes they can. We have provided a lot of justification in a related paper, Miller and Jerolmack (2020), but we summarize here. First, strong particles in a drum show a relation between rounding and mass loss that is the same as what is observed in fluvial pebbles and even aeolian sands (Szabo et al., Nature Comm. 2015; Science Adv. 2018). Second, impact attrition becomes significant and independent of the fluid when collision Stokes numbers are large enough. For pebbles underwater, this condition is satisfied for particles larger than a few centimeters (previous references, plus Jerolmack and Brzinski, Geology 2010).

The Discussion section has been completely re-written. Again, the new simulation paper by Pal et al. (2021), which shows the very same behavior to what we see in experiments, confirms that the behaviors we observe are truly general. Also, bed-load transport is one kind of collision, but landslides and rock falls are another and those do not involve collisions underwater. So, it is not obvious that examining collisions underwater would be an improvement if we are also considering other kinds of transport that may not be underwater.

- Are the impact velocities realistic? It is not specified but from my calculations, I assume that the impact velocity is 3m/s, which is higher than most river environments;

Impact energy is the most relevant parameter, not velocity per se. (But we do note that impact velocity is at least of similar order (in the field it is roughly terminal fall velocity in water, ~sqrt(RgD) where R = relative submerged density = 1.65, g = gravity and D equals grain size. For 10cm pebbles one expects a velocity of order 1 m/s; our velocity is larger but the same order.) Particle mass also matters. Moreover, we hope to have shown that it's not just energy, but material properties as well. To go a bit deeper, collision energy at the beginning of experiments is relatively large because particles are large. But as they erode they lose mass. So, each experiment actually transits a range of collision energies as the particle loses mass. This is why we compute mass lost per collision/collision energy rather than just mass lost; because collision energy keeps changing as particles get smaller.

Given the experimental configuration of a single (initially cubic) cement particle being dropped in a drum, and the ambiguities in the experimental coefficient that comes from it C1, we would not want readers to take our exact numerical values and port them to the field. But we believe that it is the simplicity of this setup that allows us to directly observe this transition from chipping to fragmentation, and to show that changing material strength alone can do that.

In the substantially revised Discussion we explicitly advise the reader on what should and should not be used from our experiments in terms of application to field and/or models. And, in the revised paper, we focus now on the bi-partite nature of mass loss vs. material strength. It is bi-partite because of the regime transition from chipping to fatigue failure; so fitting a single linear relation to these two regimes is actually not appropriate. This is explicitly spelled out in the revised Discussion.

- are the abrasion phenomena transposable to river environments? In the drum, the particles fall almost at right angles on a smooth surface (steel), while in nature most impacts will be made with a significant obliquity and the roughness of the impacted surface will lead to scratching which is not reproduced in the experiments.

We have recently addressed this at some length in the Miller and Jerolmack (2020) paper, *where for a setup that looked a binary collisions of rocks on a double pendulum we wrote: "A note of caution is in order regarding the geometry and kinematics of our binary collisions, compared to the situation of bed-load transport. Fluvial pebbles impact the bed at shallow angles, typically on the order of theta ~ 10°. Such shallow angles reduce the bed-normal collision velocity by a factor sin(theta) (Sklar and Dietrich, 2004; Beladjine, et al., 2007; Larimer, et al., 2021), and proportionately reduce the mass lost per impact (Larimer, et al., 2021; Francioli, et al., 2014). Bed-load particles may also rotate (Francis, 1973) adding a further tangential velocity component to collisions. The effect of this rotation on mass attrition, however, has not been studied. Moreover, it has been suggested that rotation is small compared to the magnitudes of horizontal and vertical velocities associated with saltation (Nino and Garcia, 1998).* ***The rounding of fluvial pebbles in nature indicates that bed-normal***

chipping, rather than tangential (sliding) abrasion, is the dominant attrition mechanism under saltation (Noval-Szabo, et al., 2018). *The usual assumption in bed-load attrition studies is that collision energy is determined by the bed-normal component of saltation velocity, which is roughly the terminal fall velocity of the particle (Sklar and Dietrich, 2004). Despite the simplified collision scenario of our experiments, collision velocities are comparable in magnitude to computed terminal fall velocities for similar-sized particles in water. We expect then that experiments can be used to examine material and energy controls on mass loss, but that observed trends will include an empirical prefactor that is related to the specific details of our configuration."*

While we cannot write the same text in a different (this) paper, we have added a few sentences of justification in the Discussion and referenced this other paper for more details. More importantly, we have now explicitly described that readers should NOT try to do this extrapolation, and why. The new Discussion explains how the main purpose of this paper is the experimental validation of the chipping to fragmentation regime, and what this means for both inferring pebble transport distance from shape (in rivers) and applying "saltation abrasion" models (which do not include fatigue failure). As we mention now in the revised Discussion, and in other places in this response file, rockfalls and landslides are also relevant situations for our experiments. So, choosing to align our results with bed load conditions is not necessarily the thing to do.

- Is the fragmentation process invoked in this paper representative of what occurs in most rocks? It seems to me that here the defects that lead to fragmentation are related to punctual defects, mostly related to the presence of air bubbles or heterogeneities of the binder between grains during drying, while in natural rocks the fragmentation will result from the distribution of essentially planar defects (fracture, schistosity, layering..)

We don't know that it matters particularly whether the defects on the "rock" are due to heterogeneities in our prepared particles or natural heterogeneities. However, fragmentation is very generic so long as ductile behavior is not dominant (Domokos et al., Scientific Reports 2015; PNAS 2020; Pal et al., arXiv 2021). In particular, even in heterogeneous natural rocks, fragmentation shapes and mass distributions have more to do with geometry and the applied stress field than they do with internal structure of the rock (Domokos et al., PNAS 2020). Surely there is some sufficiently strong structure like layering that could break the generic behavior and produce sample-specific patterns, but this does not appear to be the rule.

Most importantly, the new numerical simulation paper by Pal et al. (2021) observes the very same behavior in terms of the transition from chipping to fragmentation that we do. Their simulations use a high-fidelity and well accepted discrete element method (DEM) technique that has been benchmarked against experiments. By extension, the agreement of our experiments with their simulations indicates that our cement particles' behavior is representative (enough) of natural rocks.

- Are the attrition rates representative of most lithologies? It is difficult to translate jumps into essentially horizontal travel distance. However, if we assume that these vertical jumps of ~40cm correspond to a succession of horizontal hops, a number of experimental jumps ranging from 25 to 4400 correspond to transport distances of 10m to 1.8km. Roughly speaking, the particles can barely travel a 1km slope before being totally reduced. What are we looking at in the end? Erosion on the slopes, in the colluvial parts? or do the authors think they are reporting fluvial processes? Another way of highlighting this mismatch is to look at the equivalent abrasion rates of these experiments which would be of the order of 75%/km to 25000%/km (considering the number of jumps/distance correspondence mentioned above). The rates of grain size reduction are therefore 3 to 6 orders of magnitude higher than the experimental rates measured on most natural rocks.

Numerous other experiments, notably those of Sklar & Dietrich and Attal & Lave, were designed around the premise of measuring erosion or attrition rates that could be directly applied to the field. Our intent is not to reproduce those experiments. It is instead to focus on the dynamics of the impact attrition process that is the unit process driving attrition.

We note that our particles erode faster than most natural rocks BY DESIGN. Mass loss per collision in our experiments here is orders of magnitude larger than even the simple binary collision experiments of Miller and Jerolmack. Those experiments examined stronger, natural rocks and lower energies. We did this so that we could expeditiously examine many rocks of different strengths without waiting for months. Despite these very large differences in mass loss, the chipping regime dynamics are equivalent for our cement particles, the binary collisions of Miller and Jerolmack, the multi-body collisions of limestone in a drum (Szabo et al., 2015), natural pebbles in rivers and on beaches (Novak-Szabo et al., 2018) and even sand grains in the desert (Novak-Szabo et al., 2018). In other words, demonstrating that the signatures of chipping – Sternberg's law for mass loss, and the universal rounding pattern – do not vary with orders of magnitude change in the RATE of mass loss is overwhelming evidence of dynamical similarity. In essence, the only difference between these different materials and energies comes down to a kinematic rate constant.

For these reasons, the reader is not advised to try to translate these findings numerically into river transport distance; we have added significant text to this effect in the Discussion*. We reiterate that our main purpose was to demonstrate a transition from chipping (which we already know is ubiquitous for pebbles undergoing bed-load transport in rivers [Szabo et al., Science Advances 2018]) to fragmentation (which we already know happens in landslides and debris flows), and that it can be forced by changing material strength and have a similar effect to other studies that changed impact energy. We also validated the conceptual model of Ghadiri and Zhang (2002), and showed a continuous phase transition from the chipping limit to fragmentation, spanning all the way to instantaneous fragmentation. We do not know of another study that has attempted this, for any system. Very importantly, the new results from the Pal et al. (2021) numerical simulations confirm our experiments which significantly elevates our findings.*

Summarizing all the points mentioned above, it seems that the experiments explore conditions quite distant from those of natural pebbles and river environments. It is therefore legitimate and necessary to ask the question of the transposition, or even the usefulness, of the results of this study to natural systems.

We have addressed this above. First, since we are transiting chipping to fragmentation, we are also transiting across corresponding natural environments. So, our results are not only to understand bed load, but also landslides/debris flows and rock falls. Second, the bed load regime has been examined exhaustively already by previous studies, and the signatures of chipping by bed load are well known – from the work of Kuenen in the 1950s to more modern work (Szabo et al., 2013; Miller et al., 2014; Domokos et al., 2014; Szabo et al., 2015; Novak-Szabo et al., 2018). In these experiments it is important to confirm that there is a chipping limit, where our particles behave just like all those other studies. But the novelty is then to show the transition from chipping to fragmentation systematically, which previous studies have not done.

More concretely, even if the proposed formalism (Fig. 13 and link with Ab) is potentially interesting and could represent a first step, it has first the disadvantage of being dependent on the experimental set-up (the characteristics of the impacts are directly linked to the experimental set-up). In addition, it is still far from taking into account all of the results found in previous work. For example Kodama (1994) show that andesite and flint do not behave in the same way depending on the size considered. And also that the chert which presents a resistance in compression higher than the andesite will fragment whereas it is not the case of the andesite. What would be the elements to be taken into account in the relation (vs Ab) proposed by the authors to account for the results of Kodama?-

Thanks for the comment. The main rebuttal here is this: the new work by Pal et al. (2021), which uses direct numerical simulation of impact fragmentation via the DEM method, demonstrates unequivocally that the dynamics of chipping and the transition to fragmentation are NOT case specific, but are general. Our findings are very much in line with those simulation results, and no calibration or tuning was done in either study to produce the results found.

Now, other details – we are well aware of the study of Kodama (which Jerolmack has cited in numerous previous works on pebble rounding and attrition). First of all, ALL results are dependent to some degree on the experimental setup. Sklar and Dietrich (2001) and then Attal and Lave (2009) built ever more sophisticated devices to more properly simulate bed load erosion of a bedrock bed. But, these experiments do not allow one to isolate the damage mechanism. Moreover, particle burial and exhumation in a natural river means that any result of attrition rate as a function of distance/collisions/etc. in the lab cannot be translated directly the field anyway. Back to Kodama, though. First, the size effect of tensile/compressive strength is well known. Larger particles have a higher probability of having a large crack, and so in axial tests for most natural rocks one observes that strength goes down with increasing size. In this manner our experiments suffer from the same drawback as all experiments on attrition; the strength of the particle is surely changing as it loses mass, but one cannot measure it through an experiment because mechanical tests are destructive. Second: Kodama only examines two *different materials, subjects mixtures of particles to complex transport/collisions, and some of the results are confusing or even contradictory. There is no clear explanation why attrition rates are larger for one material when particle sizes are larger, but larger for the other when particles are smaller. It is for this reason that we have moved to simpler, single (in this case) or binary (Miller and Jerolmack, 2020) collisions where we can track a single particle and see what it does. Third: the conclusion of Kodama that one rock type fragments and the other chips is actually consistent with our findings here. We suspect that it is due to the difference in material strength between chert and andesite in their samples. Fourth: Miller and Jerolmack (2020) tested a wide range of materials and showed how Ab determines attrition rate.*

Lack of rigor

The basic mathematical rigor is absent: in the legend of figure 8, it is written that $k = A_b.C_1$ and on the other hand $k = 0.026A_b$, so that anyone would propose C_1=0.026, but here the authors conclude instead that C_1= 1/0.026 ! Also the estimated coefficient for figure 6c is wrong. Even if it could be a matter of carelessness, one can unfortunately then doubt the whole treatment of the results in all the calculations which are not explained.

Thanks for pointing out the error in C1. It was also pointed out, and addressed, in response to Reviewer 1. As for the other calculations, without pointing out specifics we cannot change things. We have made numerous corrections and clarifications throughout the paper, and hope these address this general concern.

More broadly, our reanalysis of the data – involving re-calculating the Ab parameter as recommended by this reviewer – bolsters the case that there is actually a bi-partite relation between mass loss parameter (k) and the attrition number (Ab). While a linear relation can be force through the data, it is not appropriate because the mechanism of attrition changes and so should the functional form. Our new Discussion section lays this out, compares to both the Miller and Jerolmack (2020) results and also the new numerical work of Pal et al. (2021), and concludes that there are two distinct relations that represent chipping and fatigue failure, respectively.

How can the authors contrast their results with those of Sklar and Dietrich knowing that S&D use the tensile failure threshold, while they use a rough approximation of the compressive strength?

We have removed this comparison because it is a red herring. It distracts from the main point that we observe a continuous transition from chipping to fragmentation that follows the new simulation results from Pal et al. (2021).

How can they be surprised (and without being able to explain it!) by a factor of 1000 on the value of C_1 and that of Miller and Jerolmack knowing that A_b is not defined in the same way in this paper? Considering a factor ~10 between σ_C and σ_t for mortar (e.g. Bu et al., 2017; Singh et al., 2015; Chen et al., 2013), and the fact that σ_C is overestimated by σ_U (ultimate strength instead of the elastic/plastic transition) and Y is underestimated (slope is greater than sigmaU/strain value), there is no difficulty in explaining the 3 orders of magnitude observed for C1 between these two studies.

Note that we have removed discussion of this comparison for the reasons discussed above. First, Miller and Jerolmack's (2020) study was purely in the chipping regime. Our study transits from chipping to fragmentation, so it does not make sense to compare their single linear relation to our bi-partite data. Second, we suspect now that in our experiments there is present but much weaker relation between mass loss and Ab in the chipping regime, but that our admittedly noisy material properties obscure this trend. But, once the transition to fragmentation happens, there is an explosive growth in mass loss with Ab that we are able to resolve. While drawing this conclusion from the data alone might be sketchy, the new simulations of Pal et al. (2021) support this interpretation.

The authors insist on the one hand that they observe a transition between two distinct domains dominated by chipping and fragmentation respectively. But on the other hand, they try to fit their data with a single law (fig.8a, 8b, fig.11). This is paradoxical: if they can explain their data with one and the same law, there is continuity of processes and not a transition. If there are two distinct domains, then two distinct relations must be adjusted.

See all of the explanations above. The data more clearly now reflect two distinct regimes after the reviewer's suggested re-analysis. This is supported by the Pal et al. (2021) simulations.

Comments on the form.

- Several symbols or calculations are only presented in figure captions (k_{cm}, C_1 calculation, etc). Some of them (k_{cm}) are not discussed nor used further in the text, despite the fact that a relation should be proposed between graphical derivation of k_{cm} and "k" estimation. Nowhere, the authors indicate the value of the velocity v_i that was used to compute "k" …
    - *We have drastically revised and streamlined the Intro and Discussions thanks to both reviewers. We actually removed the equation for volumetric Sternberg's law in the intro since it was not necessary or relevant, and eliminated other unused variables as suggested. There is now only one "k", clearly defined for the first time in the Discussion section. Additionally, the impact velocity v_i is mentioned and the calculation for k and C1 are explicitly discussed in the text.*
- It would be necessary to add in sup info a table with all the data.
    - *The data are available from figshare, an open online repository that follows open data standards. The url for the data is provided in the "Data availability statement".*
- Why talk about rotation when it corresponds to a series of more or less identical drops at each rotation ending with an impact on a steel plate? It would be necessary to give

the characteristics of this drop and then more adequate to speak about "number of impacts" rather than “number of rotation”.

- *The number of rotations is equal to the number of drops (since the steel plate in the rotating drum ensures that the particle collides with the side of the drum each rotation). Language has been altered to refer to “number of impacts” rather than number of rotations.*

Other comments:

- P2-L5 (=page 1 on line 5): “most models implicitly assume … governed by fragmentation”. I am not sure which models the authors have in mind, but I would say it is the contrary. Most models whatever they consider long river size evolution (Sternberg; Parker; Attal and co-authors; Sklar and co-authors), Landscape evolution model (Carretier and co-authors) or theoretical models on shape evolution (Domokos and co-authors) consider progressive and continuous wear of the pebble, i.e. implicitly chipping rather fragmentation.
  - *This text has been removed and the Discussion changed. The most important point is that models don’t account for different attrition mechanisms, and the different mass loss relations that would go with them. New majorly revised Discussion focuses on this point.*
- P2-L8 and figure 2 caption: I am not sure to follow the logic behind this statement. Hertzian cones will produce fractures that are expected to be at ~40° from the surface of contact. How would it be explaining fractures that are parallel to the surface?
  - *This is a good point. Compression cracks are more generic than Hertzian, and form for a wide range of materials. In the majorly revised Intro we have removed much talk of Hertzian fracture and keep it generic in terms of compression cracks.*
- P3-L11: as far as I remember, I don’t think that Attal and Lave (2009) are dealing with particle shape in their study _ remove that reference or replace by Kuenen or Krumbein. In contrast, they proposed some transition based on pebble velocity or size between dominant abrasion and dominant fragmentation, so that this reference would more adequate on line 15.
  - *The Introduction has been revised and no longer uses the Attal & Lave (2009) study in this context.*
- P3-L16: “This study …” is ambiguous. Does it mean Novak-Szabo study? But in that case this study does not utilize laboratory experiments. Does it mean the present study? But in that case this sentence is out of place: it would sound like a sentence at the end of an introductive section to announce what will be done in the paper. But similar sentence is proposed again on page 5 (line 18 and further).
  - *We are referring to our study. We have edited the text to be clear.*
- P3-L34: this transition seems quite odd. What is the relation between the shape evolution and the controversy on the origin of fining by attrition vs sorting?

- *We have removed the discussion of chipping vs. sorting. This topic has been covered in many other places, and it seems it distracts from the main novelty of the present study since we don't directly address that "controversy" anyway (though our findings may further inform it).*

- P4-L4: "this mas loss is proportional …" I would rather say " … is presumed to be proportional…".
    - *This edit has been implemented in the text.*
- P4-L23: "Ab" as defined in Miller and Jerolmack, or implicitly proposed by Sklar and Dietrich involves the tensile strength, not an arbitrary yield strength (that could be in flexure, compression, traction, etc). This ambiguity is largely responsible for the observed difference between the value of C1 estimated in this study and that in the Miller and Jerolmack study.
    - *See points above. There are many possible reasons for differences in C1, but we no longer make the comparison since C1 applies only to the pure chipping regime and our data transit to fragmentation and exhibit two different relations.*
- P4-L29 to 35: This type of deformation discussed in this digression (already documented/discussed in Miller and Jerolmack) is no longer discussed in the rest of the paper. Therefore, I do not see its usefulness. I suggest that it be deleted.
    - *The section mentioned by the reviewer has been removed.*
- P6-L17: "… every rotation…": from what I observed on the youtube video, rotation is not a fundamental variable. The pertinent one is the number of free fall at each rotation. Similarly, rotation speed is of limited interest. In contrast it would be necessary to document the height of fall and consequently the estimated velocity of terminal impact (probably around 3m/s if the height of fall is around 40-50cm according to the device).
    - *Language has been altered to refer to "number of impacts" rather than number of rotations, the terminal impact velocity is also added to the text.*
- P6-L27: which model of Instron UT system?
    - *The model number (Instron Model 4206) has been added to the text.*
- P6-L28: I don't understand this sentence. It must be clarified. Did the authors put cubic specimen (what they call "particles") for the compressive test? Why this choice? Why didn't they make cylinders? In any case, I think that there are tables of correspondence in the literature to transform a yield stress obtain on a cube toward classical cylinder used for UCS measurements. It must be clarified also if the particles for the drum and the one for the strength tests were prepared from the same mixture (i.e. involving the same amount of water/cement ratio), or in two different batches.
    - *Details have been added regarding preparation of samples for the Instron vs drum.*

- P7-L5: Given that the measured parameter is not a Young modulus (the true Young modulus measured from the slope of the stress/strain curve in the elastic deformation domain should always be larger than the parameter measured here, because the slope is larger than the stress/strain ratio between 0 and the ultimate strength), I would suggest naming it by a different name and a different symbol (Y* for example)
    - *As the reviewer has suggested, we refer to this measurement/parameter as modified Young's Modulus (Y*) in the text.*
- P8-L5-6: what is the convention in e-surf? 0e6 or $3.0x10^6$?
    - *We have changed to the latter (typical scientific notation).*
- P8-L18: 2700 kg/m3 for the 3rd mixture is not in the range +20%
    - *This is true – we corrected the text to indicate that the range was approximately +/- 35%*
- P8-L20: what kind of heresy is this? The samples (probably both the specimen for the strength test and the particles introduced in the drum) display large dispersions, there is not theoretical model to justify a linear fit (so that replacing the data by a value derived from this linear fit introduced an extra uncertainty), and this simplistic procedure would reduce the errors? It makes no sense.
    - *Heresy removed. See above points where we have removed the linear fits from determining Ab.*
- P8-L28: I don't see on fig.6 that mass is reduced more rapidly at the beginning of each experiment. To the contrary, the average curve (black line) do not appear to depart from Sternberg's law.
    - *This is more apparent in linear space, where mass is lost more rapidly at the very beginning. However, this is only an initial effect and it is small; after that the curves do appear to fit with Sternberg's Law. We have removed discussion of this initial part as it's not relevant to the rest of our findings.*
- P9-L5 to 15: this section should be rewritten. Sklar and Dietrich's relation is rejected on the basis of fig.8 , but fig. 8 involves the variable "k", which is defined later in this section, so that the arguments of that section are not very clear.
    - *We removed all comparison to Sklar and Dietrich as this was a distraction.*
- P9-L6: "…ultimate strength proposed by Sklar and Dietrich…". No! S&D propose to use the tensile strength (estimated through Brazilian test), not the compressive strength. As far I remember, they even claim that they tried correlations with tensile ad compressive strengths and observed a much better correlation with tensile than with compressive strength. See also general major comments.
    - *This was addressed based on earlier comments from this reviewer.*

- P9-L12-13: “we anticipate…” . I don’t follow the rationale behind this statement. I don’t find a linear relation so consistent with the data drawn on fig.8b. First the propose red line does not go through the origin; second the line is hand-drawn (a classical linear regression would be steeper); third a fit by a square root relationship would fit much better the two points at Ab=0.25 and 0.5; forth a graph with the logarithmic scale for the x-axis would show that the linear fit is a really poor predictor of the mass loss parameter for Ab<0.2.
    - *Addressed in comments above.*
- P9-L15: “mass loss is independent of material strength”. Or maybe the estimate of the strength based on a linear fit is just erroneous for the strongest materials!
    - *Linear fit has been removed. Addressed by comments above*
- P12-L2: the phrasing seems to me ambiguous. A limit on what? Do the authors mean that there is a limit on strength (or on Ab) beyond which particles achieve similar circularity?
    - *Yes, we mean a limit on strength, beyond which particles achieve similar circularity. The text has been significantly revised and no longer features this statement.*
- P14-L10&11: “particle composition”? Do the authors mean the lithology? Even if it is the case, Sklar and Dietrich did not explore shape evolution in their paper. I do not understand both the sentence and the reference. What does mean the “high composition of sand”? Do the authors mean the “high proportion of sand”?
    - *We mean the “high proportion of sand”, and this has been changed. The discussion has also been revised and the new version is hopefully more clear.*
- P14-L12: I do not see any bimodal distribution on fig.7. They seem to me unimodal on this graph (???).
    - *Rather, wide dispersion in size. Text changed.*
- P14-L16: I disagree with this conclusion “a relevant material grouping”. In log scale, for Ab<0.2, i.e. for 70% of the tested material this linear fit is a really poor predictor of the mass loss. On fig.8b , the linear relation is in fact based graphically on a single point (Ab=2, on the right part of the graph).
    - *Addressed by comments above.*
- P14-L18: “…not entirely understood”: see general comments
    - *Addressed above.*
- P15-L2: this conclusion should be discussed in a more extended way, particularly by checking if results of former studies verify or not this limit. Fragmentation of cherts [Kodama , 1994] or of limestone [Attal and Lave, 2009] seem to occur for Ab values

much lower than 0.014s2m-2 . So I would tend to say that this relationship deduced from artificial particles does not apply to natural pebbles.

- *See comments above direct translation of our findings to other situations (field, experiments) is not advised. Ab is a good parameter to explain RELATIVE differences of behavior within a system, but not as an absolute quantity to compare between systems (see ambiguity in C1 discussion).*

- Figure 6: along y-axis, fractional mass loss must be replaced by mass fraction. For the x-axis, it would be much better to display the true number of rotation rather than the normalised value. What is the "total number of rotation"? Are the 5 particle ran in the same experiment or in five different ones? In the second case, I don't understand what is the "total number of rotation". Normalization does not help in comparison at all. This normalization does not make more sense to derive the relation M=exp(-kCm*Nr). How applicable is this relation if we don't know the total number of rotation? In addition the proposed value of kCm is erroneous for the figure 6c.
    - *Yes, the y-axis should read mass fraction and this has been corrected. 5 particles were not rotated simultaneously in the drum, but particles of the same composition are compared against each other in the graph. The total number of rotations for each material strength is given in the caption; the reason it is graphically presented as normalized rotations is because weaker materials were rotated for ~20 drops, while stronger materials were rotated for ~2000 drops – so this is to make the graphs visually comparable. Furthermore, the coefficient in Figure 6c has been corrected. The graph is mostly used for illustrative purposes to indicate that Sternberg's Law applies to the various particles, rather than deriving useful values for the diminution coefficient.*

- Figure 7: It would be better to propose two graphs: one with linear scale and one with log-scale instead of this figure with two domains of distinct unit size. This would be in particular a more neutral presentation to highlight or not the presence of a transition.
    - *Re-calculating Ab using measured values results in all violin plots being able to fit in one panel.*
- Figure 8_ Caption: the C1 value must be corrected
    - *The C1 value in the Figure 8 caption has been corrected (although C1 = 0.14 after re-analysis).*
- Figure 7, 8, 11: given the large uncertainties on the ultimate strength, on the "Young" modulus and consequently on the parameter "Ab", it would be necessary to add error bars on these variables on those figures, and to include them in fitting and R2 calculations.
    - *Since Ab is calculated directly from measured values, we do not include error bars in Figure 7 or 8. Also, old Figure 11 has been removed from the revised version.*