We thank the reviewers and the associate editor for their thoughtful comments and suggestions. We have largely revised and rewritten the manuscript to address the key points raised by the two reviewers and highlighted by the associate editor. Figures have not been changed (beyond corrections requested by the reviewer), but the abstract, introduction, discussion, and conclusion have been rewritten, and the results section has been shortened and largely revised.

Line number corresponds to the MS with tracked changes.

Associate editor comments:

I have reviewed the revised manuscript alongside a second round of reviews from Al Neely (also reviewed the initial submission) and an anonymous reviewer #3. Based on these reviews and my own reading, it does not seem that the comments from the initial reviews were sufficiently addressed. Reviewer #3 also raised some useful points about the clarity of writing. I suggest that the authors revise their manuscript and resubmit it for additional consideration. In the revision, please do not consolidate the reviewer comments. Please address each individual comment separately so that we can more easily keep track of specific comments and responses.

Here I also provide a summary of the most important points that I think the authors need to address in a revision.

In revising your manuscript, please carefully consider the following:
1. Dr. Neely has asked for some important clarifications and information about the experimental setup and conditions. Please address these concerns and provide the requested information either in the main manuscript or the supplement, as possible.

We have clarified the experimental setup and conditions in response to Dr. Neely's comments. First, we have completely rewritten the introduction. In particular, the 2nd paragraph now clarifies the definition of abrasion, plucking, and macroabrasion as there is confusion in the community (see detailed comments to Dr Neely). And we specify that we adhere to the definition and classification proposed by K. Whipple and co-authors, in which plucking is not strictly speaking a purely hydraulic entrainment process, but also encompasses the processes of mechanical fatigue, weathering, and horizontal crack development between preexisting fractures due to bedload impact. The best example is this figure extracted from the latest version of the review paper by Whipple and co-authors (fig 4A, 2022), itself a copy of the seminal work of Whipple and co-authors in 2000:

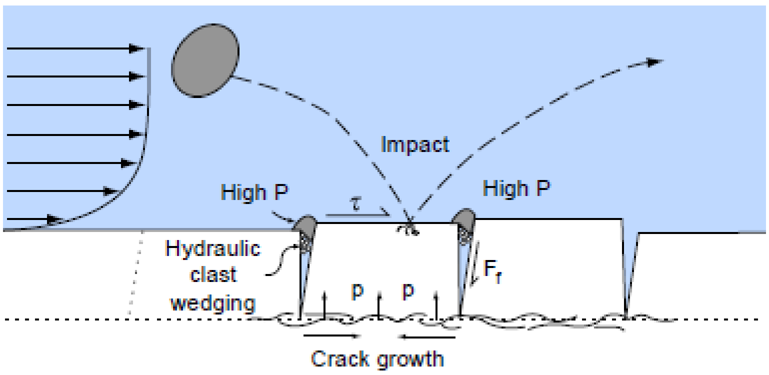

Figure R1. Description of plucking extracted from Whipple et al. (fig.4A, 2022)

The processes highlighted in this figure are actually very close to those occurring in our experiments, in particular the need for horizontal crack development for plucking to occur in relation to bedload transport. We now clearly explain in the introduction the 2 definitions that exist on plucking and clearly state that we adhere to K. Whipple and co-authors' definition. Along with K. Whipple's typical classification of incision processes, we stick to describing the two main categories of processes: plucking and abrasion and use it in our presentation of the results, before discussing the occurrence of macro-abrasion in our experiments (macro-abrasion is clearly defined in the introduction). We hope that by clarifying the definitions, the various readers will understand what are the processes that we are describing in the experiments.

Apart from this semantic point, we largely agree with Dr Neely's comment. We have thus added more details on the experimental protocol and the mechanical characteristics, in particular regarding the mechanical behavior of the BVOH, and illustrated as best as we could how weak the fractures in our disks are (see detailed comments to Dr Neely). We also provide photographs of two different disks (one without fractures and one with dip fractures) at 2 time-step intervals (before erosion and after ~60 min of erosion) in supplements (Fig. S2). A new section has been added in the discussion "4.2 Erosion processes in the experiments: empirical and physical analysis" in which we suggest that **macroabrasion** occurs in our experiments along with abrasion and plucking. This new section, along with the previous one 4.1 "Benefits and limitations of the new artificial fractured network" clarifies and answers many points raised by Dr Neely on the experimental setup and conditions.

2. All three reviewers (including the anonymous reviewer from the initial submission) have raised concerns about the applicability of these results to natural systems. I appreciate that the revised manuscript includes new subsections on application to natural systems and also other plucking studies, but those sections read more like a literature review than a discussion that integrates or synthesizes the new findings with the literature. Please follow the reviewer guidelines to revise these sections and consider how and if you can relate these findings to natural systems throughout the manuscript to convince the reader of the relevance.

In hindsight, we agree with the reviewers that we overextended the application of our experimental results to natural systems. To clarify these aspects and avoid mixing interpretations pertaining to the experiments and not to natural systems, we now have the following organization:

4.1 Benefits and limitations of the new artificial fractured network (**revised**)

4.2 Erosion processes in the experiments: empirical and physical analysis (**new section**)

4.3 The impact of fracture geometry on erosion rates in the experiments (**new section**)

4.4 Comparison with previous experimental studies on plucking and abrasion (**revised**)

4.5 Relevance to natural systems (**rewritten and shortened**)

We know offer a new explanation to the dominance of abrasion over plucking in our experiments: we suggest that the rate of erosion by plucking is limited by the depth and slow rate of horizontal fracture propagation between pre-existing vertical fractures (as in fig. R1) such that for the conditions explored, abrasion is systematically a dominant component (**abstract, discussion section 4.3 and conclusion**). However, we now clearly state **that we cannot generalize the fact that abrasion tends to dominate plucking in natural systems because**:

- we have not yet explored all dimensions of the parameter space in our experimental setup, in particular, grain size and sediment mass, flow velocity and intact concrete strength. We do not know how specific are our choices of grain size, mass and velocity in creating the dominance of abrasion vs plucking that we observe for this set of experiments (abstract, section 4.3, conclusion).

- Elementary processes are not directly comparable to natural systems in which other processes may actually favor fracture propagation such as hydraulic wedging and pressure fluctuations that are not present in our experimental setup (abstract, section 4.1 and 4.5, conclusion)

However, we demonstrate that our experiment is the first to actually combine abrasion, macro-abrasion and plucking (section 4.2), and is thus a valuable tool to advance our understanding on how these processes combine in relation to mechanical weaknesses and how they govern the morphodynamic evolution under bedload transport. As for Sklar and Dietrich (2001)'s abrasion mill experiments which had limitations (e.g., a strong vertical recirculation zone that we do not have), we view our experiment as an important step in guiding the development of a mechanistic theory that could then be applied to natural systems with the right set of parameters.

3. Reviewer #3 has provided additional suggestions on how to streamline the manuscript to make it more concise and to better frame the story for the reader. I recommend following the suggestions for the broad restructuring (e.g., cutting some intro/background, revising what is included to make it more relevant) and ensuring each paragraph is readable on its own. However, I will leave it to the authors to determine whether individual phrases should be revised or cut for brevity, as I think this is a matter of personal writing style and is not up to the reviewer.

The manuscript has been streamlined following Reviewer #3's suggestions. The introduction has been largely rewritten and shortened. We have moved some unnecessary material describing the concrete properties to supplementary material. The Results now present only the observations and have been significantly shortened while keeping the same organization and figures, while interpretations are now placed in the Discussion. As explained before, the discussion has largely been rewritten with existing sections being shortened (in particular the relevance to natural systems), and new sections 4.3 (erosion processes in the experiments) and 4.4 (impact of fractures in the experiments) added to provide better insights and understanding on the behavior of the experiments. The overall length of the MS is shorter, and we believe these changes have substantially improved the clarity and organization of the manuscript.

Dr Neely comments:

I have reviewed the revised submission and authors responses to the first round of reviews. There were some significant original suggestions from the first round of reviews that did not receive much clarification. Specifically, the material properties of the plastic material and the significance of a 3rd joint set or the experimental design with respect to plucking in natural channels.

I understand that these are experiments that were already conducted, and certain aspects of the experimental design cannot be changed. Also, I know that this is a project with limited duration. Science benefits most from patience and persistence. Hang in there!

I encourage the authors to address the points below or restructure the presentation of their results to clearly state limitations when extending these results to natural river channels. I am not an experimentalist or plucking expert, but the conditions that produce plucking are not controlled or present in these experiments. To me, it seems like the experiments are simulating macro-abrasion processes more than plucking? It might be a simple difference in wording, but this could confuse readers, myself included.

See answer to AE N°1. Indeed, we believe that this is a difference of wording and semantics, and we have sought to clarify that in the new introduction (paragraph 2):

"Away from waterfalls and debris-flow-dominated reaches, erosion processes that drive river incision into rocks with mechanical discontinuities are classically sub-classified into two categories (Whipple et al., 2022): abrasion and plucking processes. Here, we only consider rivers transporting coarse bedload material, typical of actively eroding mountain belts, and in rock lithologies where dissolution is negligible. Abrasion refers to the progressive wear induced by the impact of sediments on the bedrock substrate (Hancock et al., 1998; Sklar and Dietrich, 1998; Whipple et al., 2000b). This process can operate at different scales from wear abrasion (i.e., grain-by-grain abrasion) to macroabrasion. Macroabrasion is a recent terminology encompassing a variety of mechanical processes: chipping of blocks by grains (Beer and Lamb, 2021; Hancock et al., 1998; Whipple, 2004); breaking of bedrock into blocks by the impact of large clasts (Scott and Wohl, 2019); progressive fracturing of the bedrock surface through repeated impacts creating rock fragments that can then be subsequently entrained (Chatanantavet and Parker, 2009, 2011). Macroabrasion does not necessarily require pre-existing rock mass discontinuities. On the contrary, plucking occurs and is described in the context of rocks having pre-existing discontinuities such as fractures, joints, or bedding (Whipple et al., 2022). A detailed review of the literature shows that the exact processes encompassing plucking differ among authors. Some consider it as the process of removing blocks that are already completely loose by hydraulic forces (Saha et al., 2021; Scott and Wohl, 2019), with a variety of processes involved, such as sliding (Dubinski and Wohl, 2013), vertical lifting (Wilkinson et al., 2018), and toppling (Lamb and Dietrich, 2009). Others describe plucking as a two-step process (Beer et al., 2017; Chatanantavet and Parker, 2009; Whipple, 2004; Whipple et al., 2000b). First, the formation and loosening of blocks susceptible to detachment (e.g., by weathering, macroabrasion, hydraulic jacking or wedging), and second, their entrainment by hydraulic forces (Beer and Lamb, 2021; Chatanantavet and Parker, 2009; Turowski et al., 2023; Whipple et al., 2000a, b). A typical example of the second definition is the progressive impact-driven creation of cracks between pre-existing fractures to create loose blocks that are subsequently entrained (e.g., fig. 4A in Whipple et al. 2022, itself derived from fig. 4 in Whipple et al. 2000b). Hydraulic wedging and jacking by transported grain can both help in crack development as well as entrainment (e.g., Hancock et al., 1998; Hartshorn et al., 2002). In this study, we adhere to this second definition of plucking. We note that glacial erosion by plucking (also called quarrying) is also defined as a similar 2 steps process encompassing a weathering/fracturing phase and an entrainment phase (Anderson and Anderson, 2010; Hallet, 1996; Scott and Wohl, 2019). Beyond semantics, we note that a key point is that a larger range of erosion processes encompassing wear, macroabrasion and plucking can occur in rocks with mechanical discontinuities, compared to homogeneous rocks. This makes quantifying the impact of discontinuities on erosion efficiency a challenging problem (Scott and Wohl, 2019) as the proportion and the rate of each process must be quantified." Lines 49 to 76

We choose, as we did before, to adhere to the definition of plucking put forward by K. Whipple (in 2022, but similar to his seminal 2000 paper with Hancock and Anderson), which includes horizontal crack propagation between existing vertical fractures as being part of the plucking process. That is why we still use plucking in the results part.

In my opinion, the current form of the manuscript still overextends the interpretation of the experimental results to field settings that are not comparable to the experimental design. This was brought up by both reviewers.

We completely agree with the reviewer and have largely changed the discussion to reflect this (see response to AE N°2)

The experiments were performed rigorously, and using novel techniques, so there is still much insight gained from the results and the experimental work. But, as it is presented now, the paper is describing flume experimental conditions that do not really reflect plucking that occurs in natural channels.

We do believe that our experiment reproduces key elements of the plucking process as defined by Whipple et al. (see comment to AE N°1) and that the reviewer simply does not use the same definition of plucking than we do. Leaving aside the debate on the definition of plucking that we do not seek to solve in this MS, we have rewritten the discussion to better address the relevance of our experiments to natural systems (see response to AE N°2)

Specifically:

1) There is still not much clarity on the properties of the plastic material that was used to create heterogeneity in the erodibility of the substrate along the fracture planes. The authors suggest that this is difficult to quantify and do not provide further information. Because no information is given, I am providing suggestions, although I am not familiar with working with the material firsthand. So if these are impractical, I apologize, but this is a limitation of the study that is significant, and I was hoping the authors could provide further information or simple experimental results.
   The plastic used in our experiment is indeed challenging to analyze in terms of its strength as it interacts with the cement in complex way: while the plastic fully dissolved when it is immerged alone in water (same temperature and duration of an experiment), we do not observe such complete dissolution for a "fractured" disk during our experiments or in the tests performed, even in hot water. We clarify this in the methods and materials section (line 141) and in the discussion (4.1, line 381)
   Performing tensile strength analyses of this plastic material under varying humidity conditions would thus not be relevant. We understand the concerns of the reviewer and have added more empirical elements to give a sense of the limited resistance of the plastic network:
   "*After several tests of 3D printing material, we chose BVOH to print the synthetic fracture networks, a plastic material that fully dissolves when immersed in hot water. However, in our setup, the cement alters the chemical properties of BVOH and it does not dissolve, but is turned into a softened plastic remaining in the fracture plane of width 0.9 ± 0.1 mm. Consequently, water cannot flow through the fractures, and the blocks are not fully loose laterally. Although the tensile strength of intact BVOH is high and generally depends on orientation, ranging from 8.7 to 33.7 MPa for ZX (vertical) and XY (horizontal) planes, it is strongly reduced in the condition of our experiments. For instance, when the fractured disks are immersed in the column, they cannot be extracted again as they break instantaneously along the fracture planes. However, we note that their softness can create a coating on the fracture planes when they are exposed, which could dampen bedload impacts before being fully abraded.*"
   Even if this is qualitative, we believe it gives the reader the important information that the network as a very small mechanical strength.

   a. Can't observations of whether the plastic is remaining or gone throughout the experiment be useful for determining the relative material properties?
      We note that plastic exposed at the surface tends to be removed along with the detached blocks. We sporadically observe plastic protrusions, indicating that this behavior is not systematic. However, we do not have measurements that would allow us to quantify this process systematically.

b. I would expect that the solubility of the BOVC is constrained to some degree or could have been constrained by submerging the material at different temperatures and durations. This should be compared to the length of time for each experiment and the amount of water that is circulating through the flume (does the water reach saturation?)

The BVOH when not embedded in concrete dissolve in water in 20 minutes for hot water and in 30 minutes for cold water but did not dissolve when embedded in our concrete disks (even after 4 days in water). This is now stated in the MS in the methods and materials section (line 141) and in the discussion (4.1, line 381).

2) No pictures of the flume runs are provided. Topographic scans are different from photographs, which could show development of fractures in the concrete or differences in erodibility between the plastic and the concrete, which apparently is not straightforward and could evolve during the experiment (relating to 1a).

a. Why are these pictures still not provided? This makes me somewhat skeptical.

We are sorry, about this omission in the last review. We did not add pictures initially, as we do not think it would help the reader as they are not close shots of the discs: the only photos we can take are from the top of the erosion column used for SFM (Fig. 1). The plexiglass is also scratched by the saltating gravels and is not transparent enough for good pictures. There are two reasons why the discs need to remain fixed at the bottom of the erosion column throughout the experiment: (i) to ensure a stable 3D referencing, and (ii) to avoid breaking apart the disk if we try to extract it. For the same reason it is not possible to take a close shot outside of the column at the end of the experiment.

We have now included the requested photographs (two different disks, one without fractures and one with dip fractures, at 2 time-steps intervals: before erosion and after ~60 min of erosion (Fig. S2 in the supplements). We cannot do better than that, and unfortunately cannot rerun experiments in the limited timeframe of this PhD project.

3) There is no 3D connectivity of the fracture network. Is this really a fully developed fracture network that is representative of fractured bedrock?

a. Most prior studies of plucking use already detached blocks, because most channels are presumably eroding through already detached blocks (at least at some depth scale, the fractures connect or there is a weakness in the 3rd dimension (i.e. bedding plane) to generate a fracture. As I mentioned before, it is very hard to see fractures underneath a riverbed. They are usually buried if present! But at least some cliffs/spires/waterfalls where plucking is active show subhorizontal discontinuities which would create a 3d network of fractures. There is no discontinuity in this experiment to replicate this condition. Qualitative --- but I have visited the channel in this study of plucking by Snyder et al., 2003, and there are detached blocks near small waterfall lips throughout the channel, and the blocks have geometries matching the thickness of bedding planes.

Snyder, N.P., Whipple, K.X., Tucker, G.E. and Merritts, D.J., 2003. Importance of a stochastic distribution of floods and erosion thresholds in the bedrock river incision problem. Journal of Geophysical Research: Solid Earth, 108(B2). https://doi.org/10.1029/2001JB001655

See response to AE N° 2: we adhere to the definition of K. Whipple in which the ongoing development of a horizontal crack between preexisting fractures is part of the plucking process (Whipple et al. 2022). The reviewer considers plucking only as the entrainment of already loose blocs, which can sometimes be the rate-limiting process of plucking (as in Snyder et al., 2003), but, to our experience in active mountain belts (Taiwan, New Zealand) is rarely the rate-limiting process for plucking to be efficient. That is another debate that we did not want to start in the MS to keep it concise. The problem stems from the coexistence of two different interpretations of plucking in the literature, and the community does not facilitate our job or the reviewer's! That is why we frequently use "entrainment" rather than "plucking" in the MS when we discuss specifically the removal of blocks, as this has a closer correspondence with the physics of sediment transport (in particular in section 4.4, comparing our results with previous "plucking" experiments with loose blocks).

But we also agree that it would be interesting to explore the role of a third set of fractures. We initially tried, but never found a correct solution. We now specify this for readers who would like to test similar conditions (section material and methods, L 161).

b. Still, the experiment is looking at the influence of spatially heterogenous bedrock properties on the spatial pattern of erosion, which is an important problem, but to me, the experiment is quantifying a process that is more similar to macroabrasion than plucking? This is shown by the continuum of "block" sizes in figure 7 (maybe this could also be presented as a histogram of plucked areas?). If plucking was really occurring along fracture planes, I would think there would not be a continuum of block sizes. There would be a higher frequency of particles detached that match that of the fracture spacing.

We agree with the reviewer, and we use the observation that erosion patches smaller than the fracture spacing likely correspond to macro-abrasion exploiting mechanical weaknesses in the vicinity of fractures. This is part of the new discussion section 4.2, addressing the nature of erosion processes occurring in our experiments. Because a large fraction of erosion patches corresponds to the fracture spacing (but also measured debris in S4), in particular at high fracture density, we state that plucking (as defined by Whipple et al., 2000) occurs in our experiments. We qualitatively discuss cases where macro-abrasion might actually dominate plucking (at low fracture density, 50 m-1, Fig. 4a), but also cases where plucking largely dominates macroabrasion (highest fracture spacing, 145 m-1, Fig. 4a) as plucked blocks have a size very similar to the fracture spacing.

Macroabrasion is still an important erosion process! but it is not clear to me that these experiments can be equated with plucking? which could be confusing to many readers.

We agree with the confusion, which is preexisting in the literature and we show it in the introduction (see comment to AE N°2). We now state that understanding the nature of the processes occurring in the experiments is also an objective of the MS. We have tried to find solutions in which we did not use the terminology abrasion and plucking in the presentation of the results until we discuss the nature of the processes occurring in the experiments, but did not find a solution and wording that did not make the MS longer and clearer.

Also still, it is misleading to state that plucking does not significantly increase erosional efficiency, implying that this is true everywhere (i.e. without a statement clarifying that this is just in the experiments, which do not necessarily have conditions that cause or perpetuate plucking?).

We have removed this conclusion for natural systems. See answer to the associate editor (N°2).

I am trying to provide these comments to be constructive (although critical)!

We thank the reviewer for the time taken and the quality of his review. This has challenged us and led to an improved version of the MS.

Line by line comments (line number refers to line in revised manuscript):

Line 66: "two main steps" – the first step may be completed prior to any interaction with bed sediment?

Line 88: "However, at long time scale, the landscape should equilibrate with the tectonic forces so that these variations in erosion rates might only be transient"
But the point is that landscapes would then show a fluvial relief (or hillslope/colluvial channel relief) that is lower, like the analyses of DiBiase et al., (2018) – Figure 3c) and then Neely & DiBiase, (2023) – Figure 3, Figure 5a), which look at interaction between the fracture density and fluvial, colluvial, and hillslope relief.
Also, these landscapes might have a different response timescale to upstream propagation of baselevel signals i.e. an acceleration in uplift at a range bounding fault?
To me, a change in river erosion for a given channel geometry due to a process such as plucking is pretty significant in these contexts. River profiles are often used to interpret tectonic histories with models that do not consider differences in erodibility due to changes in erosion processes (i.e. Fox et al., 2014 – and similar studies) – and then the timing or occurrence of a change in tectonics is misinterpreted – (these inversion models can still be very useful starting hypotheses to identify more complex interactions)
Fox, M., Goren, L., May, D.A. and Willett, S.D., 2014. Inversion of fluvial channels for paleorock uplift rates in Taiwan. Journal of Geophysical Research: Earth Surface, 119(9), pp.1853-1875. https://doi.org/10.1002/2014JF003196

Neely, A.B. and DiBiase, R.A., 2023. Sediment controls on the transition from debris flow to fluvial channels in steep mountain ranges. Earth Surface Processes and Landforms, 48(7), pp.1342-1361. https://doi.org/10.1002/esp.5553
We recognize it was not pertinent for the paper and removed it.

Line 101: In most cases? What cases? This line seems to contradict the concerns that both reviewers brought up about not having a 3D network of connected joints?

We removed this unclear sentence.

Line 105: Hydraulic jacking as mentioned in the past review involves smaller clasts that wedge into fracture planes or plug flowpaths under the clast, creating lift. [Done]

Line 107: This definition seems a bit too general. The same definition could describe abrasion, which is a completely different process from plucking?

We added clarity by better define the processes in the introduction (line 52).

Line 148: "joint" – this spelling correction occurs a few times throughout the manuscript [Done]

Line 507: "constant sediment input"? Over time more sediment is added? Or there is sediment added in the beginning of the experiment? Sediment is also produced from the eroded concrete. Minor, but this is a bit confusing to understand.

We removed this unclear sentence.

Line 508: "identical hydraulic conditions over time" – maybe the forcing is constant from the propellor, but the bed surface topography and fluid that is circulating is changing due to erosion?

We removed this unclear sentence.

Line 517: "at the scale of our experimental setup" I think this kind of cautionary language and limiting the scope of the results to the experiment is needed in much of the discussion of the results. Extending these results to natural landscapes is challenging due to points raised in the review. Indeed, we overinterpreted our results in the discussion. We have now rewritten it (see associate editor comment N°2).

Line 519-521: but the point is exactly this! The slope of the channel or the width of the channel may be different because the plucking process is effectively removing material to balance the tectonic uplift rate (or sediment supply rate) with a different channel geometry. There is no citation in that point to any field studies suggesting that plucking does NOT enhance erosional efficiency, so I do not see a basis for this argument or extending the observations to field settings?
We removed the conclusions or interpretations related to channel morphology in the MS.

Line 531-532: Minor detail, but this is not always true. Neely & DiBiase 2023 (Figure 5b) show that channels have similar width, or slightly wider channels with sparser fracture spacing. Why? I do not know (perhaps jamming effects due to large clasts sizes relative to narrow channels?). But that is what was there.
As above, we have removed these elements of the MS

Lines 543-549: This paragraph illustrates the challenges of applying the results from the annular flume to a river. I think readers will be confused if they are left with this point. The experiments do not capture these feedbacks, but these feedbacks are thought to be important for sustaining plucking in the field? This paragraph has been revised, see answer to associate editor (N°2, 3).

Reviewer #3 comments:

I found this paper to contain intriguing experimental results, attractive figures and thoughtful consideration of experimental design confounding variables. The experiments are an important contribution to the field of experimental approaches to bedrock erosion, and the authors were thorough in explaining the rationale of experimental design and the technical aspects of the experimental setup, and they provide a meticulous accounting of the results. However, the manuscript must correct two major issues before it can be considered suitable for publication.

First, the manuscript feels long and wordy which limits reader comprehension. At the macro level, this stems from disorganization in the results and discussion, as well as an introduction that describes at length systems that are not relevant to the experiments at hand. These three sections require more focus to tighten them to keep the reader's mind from wandering – at times there is commentary in Results and new methods introduced in the Discussion. The results section should succinctly describe what was observed, and then the Discussion section should contextualize those findings within previous experimental results and field observations. My advice to address this is to revise the manuscript such that every paragraph begins with a topic sentence stating either a fact or a knowledge gap which will then frame the remainder of the paragraph text. The Introduction especially feels like paragraphs begin and end at random. Ideally, a reader should be able to skim a manuscript by reading the first sentence of each paragraph and get the gist of the motivation for the study, its results, and implications. Topic sentences in the Results should state a result. At the micro level, many sentences in the Results contain extraneous phrases (e.g. "we now explore…" "Another way to document these patterns is…", Lines 311-313) that should be removed_to sharpen the focus on the results of the study. I provide some line-by-line advice for both macro and micro fixes.

We thank the reviewer for his detailed advice. As explained in the response to the AE (N°2 and 3), we have significantly reorganized the MS, and largely rewritten it to streamline it, and remove unnecessary working. Results and interpretations are now clearly separated. See answer to associate editor (3) for more details.

Second, the paper fails to convincingly connect their experiments to the "real world." Part of this stems from spending too many words diving into the details of the experimental results without justifying to the reader early or often enough that these experiments are close enough to how real rivers behave that it is worth peering so closely and deeply into these experimental results (see above). But there are too few words in the Discussion that connect specific results to observations about plucking and abrasion observed in the field – it seems like the authors take for granted that the reader will believe that their blocks and spiraling water behave enough like real rocks and rivers to implicitly connect their results to natural phenomena, but if one is not convinced, large swaths of the Discussion are challenging to accept. I point out these instances in line-by-line advice below.

See answer to associate editor (2) as to the connection with natural systems. We note that abrasion mill are not new, have limits in terms of hydraulics, but that does not mean that they have not contributed major advances in our understanding of the tool and cover effect for instance, or the relationship between erosion efficiency and rock mechanics, even if the rocks were not always real rocks, but concrete or foam (e.g., Sklar and Dietrich, 2004 and Scheingross et al., 2014). We actually state this in this introduction (first paragraph and third paragraph).

Because these are the first experiments which are able to obtain abrasion, macro-abrasion and plucking simultaneously, we believe it is important to detail the nature of the processes (new specific discussion section 4.2), and why in the experiments abrasion systematically dominates plucking and macro-abrasion (new specific discussion section 4.3). We agree it takes some space, but we believe it is essential to link elementary processes to macroscopic behavior (e.g., average erosion rate, proportion of plucking) at least in the experimental device.

Line 152: Is this the proper citation format for this resource?

We modified the citation format for a more appropriate one

Figure 4b and c: it feels like maybe the 125 m-1 fracture density sample is a bit of an outlier, as is the particularly high erosion rate sample with 150 m-1 fracture density, but some interpretation is made of these patterns. I would recommend being more cautious interpreting any trends with fracture density (e.g. interpretation that erosion has a max at 150 m-1 when it really might just increase with increasing fracture density if that one point is an outlier). We do not consider the 150 m^{-1} experiment (10_20) as an outlier, since all experiments with this configuration exhibited similar behavior. The 125 m^{-1} experiment (10_35), however, was tested only once, and we agree that it lacks reproducibility. Its lower erosion rate may reflect experimental variability or the limited occurrence of plucking due to the larger fracture spacing (35 mm), as no plucking was observed in the 40_40 mm configuration, which shows a similar spacing. We discuss this aspect in the new section 4.3

Line 299: Cite "period of weakening" precedence? [Done]

Line 326: Well, it's not monotonic, but more of a step function, and really not much of a trend at all if the 160-degree sum sample is an outlier.

The 157° sample shown in Figure 6 is not an outlier. A second experiment conducted under the same conditions eroded too rapidly (losing about 400 g after only 16 minutes) to be properly analyzed with our method. We now further use this info, with a new approach that account for the asymmetric plucking pattern (which is thus not directly comparable to vertical fracture) to show that this specific combination of dip angles generates the largest proportion of plucking (estimated at about 40 %), as well as the highest average erosion rate of all our experiments. See the new section 4.3 on the role of fracture patterns on erosion rates. We added a mention of this experiment in the results section (Lines 326 to 329) and in the supplementary material (Fig. S7).

Lines 518-521: This sentence perplexed me; your experimental "river" does not adjust its width or slope to maintain any form. Instead I think your experiment showed that plucking versus not plucking does not ultimately affect erosional efficiency, so I don't know how that result can connect to rivers adjusting to account for plucking – your experiments seem to imply that real rivers shouldn't need to account for plucking at all because plucking doesn't change erosional efficiency. Am I misunderstanding something?

We completely remove any reference to width or slope adjustment as this not relevant to our study.

Line 65 "…thereby promoting the detachment and mobilization of blocks by plucking." Does this need a citation?

We reformulate this part to add clarity (Line 60-74).

Paragraph starting with Line 77: It's not clear to me that hillslope fracturing is relevant to this study. Would be clarified with a topic sentence that perhaps uses quantifying hillslope fracture density etc. as an example of attempts in the past to measure what this study can experimentally control, but I'm not sure it's an appropriate direction for the introduction.

We shorten the introduction, and removed what was not relevant for our study.

Line 87-88: Can you elaborate/be more specific about this timescale?

We removed this unclear sentence.

Lines 125-130: The geometry description would benefit from also having a cartoon/diagram. Right now the concrete disk as sketched lacks detail to show the different geometries you've set up.

We modified the diagram for more clarity and added a picture of the top of a disk in Fig. 1.

Line 150: "…but it can also be different." Clarify/elaborate?

It was unclear, so we rewrote it.

Line 177: "parts of the disk can be removed" by people? Sediment? Can be? Did occur? Rephrase for clarity here.

We clarified this sentence (line 186).

Line 204-205: what does this sentence mean? Data? Uncertainty? Clarify.

We removed this sentence because it was unnecessary to understand our study.

Figure 3: I don't like the bold lines being connected to disks, it is more intuitive that the bold lines represent some mean (which I think you should show for each fracture density). You can just annotate specific data lines with text to connect back to Figure 2. Instead I would use bold lines as averages.

We removed the bold lines on the figures. However, to keep readability on the figure, we did not add the mean.

Paragraph beginning with 385: this paragraph and the one following it mostly feels like Results. You are just making slightly different plots of your Results. The same goes for Section 4.3, and honestly both of these subheadings would make appropriate Results subheadings, and your Results subheadings would be better suited for either topic sentences in the Results or as topic sentences in the Discussion.

See answer to associate editor these paragraphs are now reorganized more properly for readability.

Lines 415-416: How similar is the impact and fracture scale to real rivers? If this were in the Results it would be early for interpretations, but since we're in the Discussion, please contextualize your results with field observations so that the reader can know if these results are relevant to the real world.

The discussion part about the relevance for natural setting is rewritten more clearly, specifically we acknowledge that our fracture patterns are simplified version of real systems. Also, we do not seek an exact scaling with nature or a specific site but a similarity of processes, and not necessarily rates. We thus do not believe that field observations (e.g., pictures) would actually help. We refer more generally to processes and articles that have observed processes, but rather to develop an understanding of how the fracture density impact the experiments dynamics to evaluate, if similar characteristics.

Lines 425-455: Just re-emphasizing that this is all Results, or 80% of it is results that should then be referenced in the Discussion when providing the interpretations.
We reorganized the results and discussion section, See answer to associate editor (N°2, 3).

Lines 469-473: This sentence needs more support from the results. How do we know? Show me or refer to specific results that support your interpretation.

The discussion part has been rewritten.

We thank the reviewer for their remarks and suggestions below. We have addressed all line-by-line comments where relevant in the revised manuscript.

Line 40 and throughout: I also love Anderson and Anderson, but is there a review paper or similar that might be a better citation than a textbook throughout this manuscript? [Done]

Line 40 and throughout: "well-known" does not need a hyphen. There are numerous instances where punctuation and phrasing are a little off (e.g. word choice of "maximum" in Line 44, comma after "turn" in Line 61). Checking with an English language editor will clear some of these stumbles. [Done]

Line 107-108: An odd place for the first definition of plucking in the introduction; move earlier. [Done]

Line 111-112: This is the first mention of the role of fracture network geometry specifically (not just density). This topic was not laid out in the Introduction despite it being the focus of the experiments. [Done] (Line 82)

Distill the existing Introduction into half its current length and then add more background on fracture network geometry to reflect the remaining focus of the paper. [Done]

Line 195: watch phrasing "10 times this protocol" [Done]

Lines 311-313: example of phrase to be cut to streamline. [Done]

Line 354: Rephrase "we have no clue" [Done]

Line 363-364: Check phrasing/word choice here. [Done]

Lines 383-384: Cut [Done]

Line 488: Once again, please refer to specific results/figures that support your interpretation. [Done]