Thanks, Simon, for the detailed comments. Our replies are given in *italics*.

General remarks

1. *We added another study by Lifton et al. (2009) that we previously overlooked and that investigates local lithological controls on valley width.*
2. *We added 'hydrology and flood dynamics' as an additional relevant topic at the end of the first paragraph, slightly revised the statements and bolstered up the references.*
3. *Regarding the meandering literature mentioned at several points by the AE – we were aware of many of the mentioned papers (Constantine et al. 2014 was actually cited in the introduction). We were not aware of the paper of Ielpi and Lapôtre 2019, so thanks for this suggestion. The crucial points for us are that (1) the controls on migration speed are debated and none of the mentioned papers provides a full picture. A problem here is that many of the potential drivers are cross-correlated. For example, Ielpi and Lapôtre use channel width to normalize, which is also correlated, for example, with slope and water discharge. As such, a purely empirical approach is not ideal for clearly discriminating control variables. We provided a comprehensive discussion of potential controls of migration speeds in the Bufe et al. 2019 paper, including several approaches for dimensional analysis, and a discussion of the experimental and field data available at the time of publication. (2) Within in the context of the valley width model, the relevant parameter is lateral transport capacity qL, not lateral migration speed, and the precise controls on qL / speed are not of first order relevance for the outcomes of model (as long as we agree that speed V and qL are related by qL = V*H+; see Bufe et al. 2019 for evidence and discussion).*
   We have now added a couple sentences after the introduction of eq. 3:
   *"*Both *V* and q_L are determined by hydraulic and sedimentological boundary conditions. The precise controls have not yet been completely resolved (e.g., Bufe et al., 2019; Constantine et al., 2014; Ielpi & Lapôtre, 2019; Wickert et al., 2013), but are not directly relevant for the remainder of the derivation."

16 Strange syntax. I don't think you need the "once".
Deleted as suggested.

43 keep this if you want, but I don't think it is needed.
Deleted 'grounds' as suggested.

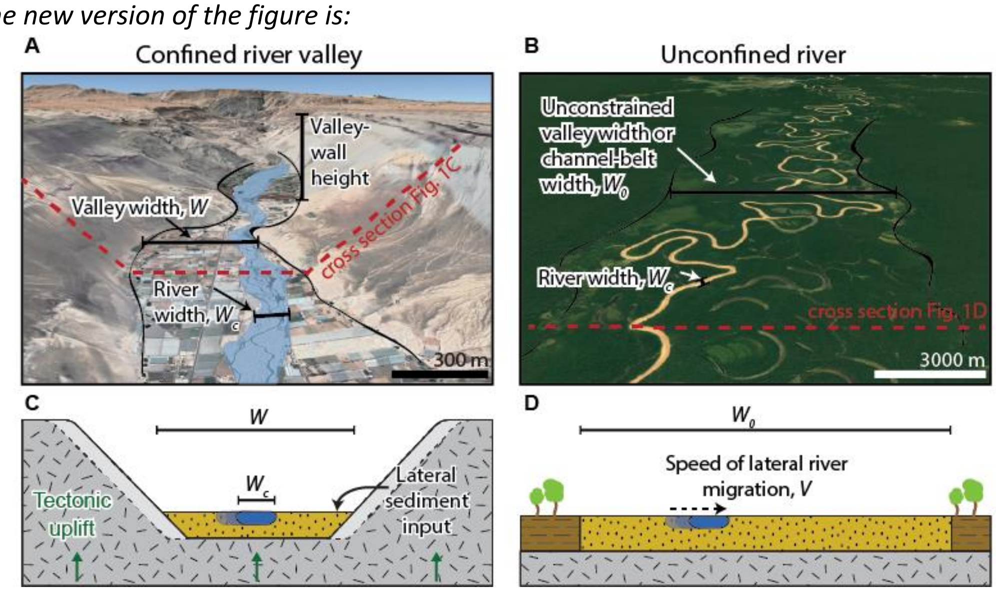
102 "forth and back" sounds very strange to a native English speaker. If you don't wish such persons to be confused I suggest "back and forth".
Changed as suggested (although, 'forth and back' seems more logical, because one needs to go forth before being able to come back…).

104 The basic premise is that the upper value of valley width is determined by the channel belt width, when the channel is unconfined. The model then goes on to discuss lateral velocity of the river in such a setting. I find it very strange that there is no reference at this point to the river meandering literature. Surely the information about lateral migration rates from that literature is relevant here?
See general comment #3.

119 This is a strange way of saying something that seems trivial. Maybe I have misunderstood. You are just saying there are waiting times longer than the mean waiting time? Is this not obvious? I suggest altering the wording here so it is clear what you are trying to say.
Maybe we are making things a bit more complicated than necessary. The point is that in a fully stochastic version of the model, switching time is obviously not constant, and the steady state valley width sets to some 'effective' or 'characteristic' switching time. This characteristic switching time should scale with the mean switching time, but cannot be expected to be exactly equal to it. The parameter c is meant to account for this. The sentence was revised to:
"As such, the valley sets its width to some effective lateral migration time that can be expected to scale with the mean waiting time, but is not necessarily exactly equal to it."

126 Support (or not) for this assumption can be found in the meander literature. It would be useful to say how dependent the results are on this assumption.
See general comment #3.

131 I am a bit confused by this image. The "unconfined" image shows a braided river. This images does delineate the floodplain as part of the channel belt. The confined river shows a weakly braided river with a floodplain being included. I don't think these two images are comparing the same thing.
I think panel b would be more useful if you showed a relative elevation raster derived from lidar, which makes it clear that the river migrates back and forth along the floodplain (many examples are shown on social media, if you are into that sort of thing).
This comment sparked quite some discussion amongst us about the definitions of channel belt, floodbelt, floodplain and so on. We have resolved to not go into that in the present paper, and have instead revised the figure to focus on the terms that are actually relevant for the model.
The new version of the figure is:



141 Again, some mention of the meander literature would be useful here. Do you not think it is relevant? If not, why not? There have been several studies on meander migration rates that link meander velocity to sediment fluxes (not lateral...from upstream) e.g. constantine et al (2014) Nature geosciences
See general comment #3.

166 This paper: https://www.nature.com/articles/s41561-019-0491-7 says that h is a power law function of width, and velocity is also a power law function of width.
I don't expect a new model but some reference to this literature would be useful.
Thanks, we had not seen this paper.
The geometrical relationship between channel width, flow depth, flow velocity, and discharge is known as hydraulic geometry and has been widely investigated since the 1950ies (the relationships were first described by Leopold and Maddock, cited elsewhere in our manuscript). The relationship between width and depth is related to this concept. The mean behavior suggests that the width-to-depth ratio is weakly dependent on discharge, with an exponent of ~0.1 when moving downstream and ~0.14 at a given cross-section. Please note that the power law exponents (and correspondingly the aspect ratio) vary widely between river systems (see our discussion on this issue in section 4.3 and 5.2, and the compilations of Park 1977 and Rhodes 1978 cited there). Further, this is a mean behavior, which can locally change due to instabilities, bank erosion and so on. Even in the paper suggested by the AE (Ielpi and Lapôtre, 2019), flow depth varies about an order of magnitude for a given width (see Fig. 3c). Further, the paper is not concerned with the channel switching direction of migration, for which we develop an argument at this particular point.
See also general comment #3.

185 As long as the forcing is constant.
Yes, that is correct. It is clearly stated in line 183, two sentences prior, in the opening sentence for the paragraph and section. The model framework would also allow to deal with non-uniform incision, but we do not explore this in the present paper. No changes.

203 I think it would be helpful to the reader to spell this out: "as can be expected based on straight, entrenched rivers in rapidly uplifting landscapes." (or choose some statement as to why this is expected).
Changed to "as can be expected for entrenched rivers in rapidly uplifting landscapes".

209 I don't expect any changes here, but in the meander literature the sediment supply controls the meander rate, so it would be interesting to consider if higher sediment fluxes from upstream would modulate the rate. Do you think it plays a role?
Yes, sediment supply is one of the controls on lateral transport capacity q_L. The precise controls on q_L do not affect the outcomes in the way we have been using the model in the present paper. See also general comment #3.

213 I would write out "dimensionless" here.
Changed to "Here, *P* [-] is a dimensionless parameter denoting the fraction of time…"

251 I think the fitting procedure should have more detail here.
W_0 and W_C are dependent on drainage area, but it seems like from the results below lump valleys with different areas together. Can you please explain how/why this is done and justify it. And give a precursor to the later section specifically dealing with this issue.
A justification is given in the subsequent sentences, lines 252-255 in the old manuscript, lines 260-263 (final manuscript) and lines 262-265 (tracked changes) in the revised version. Please let us know if this is not sufficient.

252 delete 'fit'
Deleted.

341 It isn't clear here if the references specifically said that you could linearly fit erosion rate and k_sn in the himalaya, or if these are general statements. It should say so.
Revised to:
"Even though relationships between k_{sn} and erosion rate are commonly fit with non-linear power laws, the scatter in most data sets make a linear fit equally appropriate, both in general (e.g., Kirby & Whipple, 2012; Lague, 2014) and for the Himalaya specifically (e.g., Lague, 2014; Scherler et al., 2014)."

346 Why is the exponent here different to the stated units of k_sn on the previous page?
Typo, corrected now.

354 I think this would be easier to understand if the second sentence started with: "These two limits are defined by..."
In its current state it takes some time to parse (for me at least).
We have revised the sentences to:
"At large values of the mobility-uplift parameter M_U, corresponding to large values of the lateral transport capacity q_L or small values of uplift rate U, the model predicts an asymptotic approach to the unconfined channel-belt width W_0 for zero hillslope sediment supply q_H (eq. 16). Conversely, when uplift rate is high or lateral transport capacity is low (small values of M_U), the equation levels off at the channel width W_C."

378 This seems quite wide. How does it compare to satellite imagery?
A direct comparison is hard to make, because (i) likely flood width is relevant, and (ii) the fit value is an average of channel widths of channels at different drainage areas. Yet, the value seems reasonable. For example, the flood width at Manas (basin with the largest drainage area of 5541 km^2 from the northern group) can be estimated to up to 700 m, with about 350 m seeming a reasonable estimate within the valley.
Coordinates of all locations are given in Table 2.

392 What an amazing result.
It is, isn't it?

435 Nice result.
Thanks!

508 The meandering literature has data on this and should be cited here.
*The statement here is within the context of our model. We added: "*In our model" *to the sentence. See also general comment #3.*
Generally, there are a lot of as yet untested predictions and assumptions in the model. Suitable data to directly test them are rare, and test construction is not so straight-forward for field settings in particular (for example, which flow depths is relevant in a channel with varying discharge?). For the particular statement made, the proportionality of channel belt width and flow depth, we are planning to look at the experimental data by Limaye (2020) in a future publication.

613 Not sure if I agree with this: the rate of tectonic convergence will be correlated across the range. Or is this not what you are arguing?

This is a valid point. Yet, in tectonically active settings, uplift is often not continuous, but occurs along faults, and there may be opposing effects on basins draining, for example, to the hanging wall and the footwall. There is also typically a variation of tectonic activity across the strike of major structures. The key point we want to make is that it is unreasonable to expect uniform forcing over a region as large as the Himalayas. We have revised to give a more careful statement:

*"*Case (ii) can occur if the relative base level uplift rate. Further studies are necessary to investigate whether the spatial or temporal variations in uplift rate for catchments in the Himalaya are consistent enough to cause major base-level changes and sediment aggradation of V-shaped valleys that are correlated across the range."