Response to the Reviewer 1's Comments

Blue: Reviewer's comment,

Black: Response to the comment,

Red: Lines added or modified in the revised manuscript

While I think the general idea of the paper – to study how much lateral viscosity variations (LVV) due to temperature and strain rate dependence may help to explain the discrepancy between the (higher) amplitudes of dynamic topography inferred from mantle flow and the (lower) residual topography estimates based on observations is useful in that it addresses an unresolved problem, and I also appreciate the relatively simple setup, which should help with gaining a qualitative understanding, I think the current paper suffers from several shortcomings, which limit its usefulness.

Firstly, the parts without LVV add nothing new to what is already known. Of course, I realize that these are mainly meant for comparison with the later results with LVV. But that a contrast between low-viscosity mantle and high-viscosity lithosphere leads to increased dynamic topography, and the topography gets higher the stronger and/or thicker the lithosphere is, and that an asthenospheric low-viscosity channel leads to reduced topography can all be inferred from topography kernels (see e.g. my papers from 2001 doi:10.1016/S0012-821X(01)00229-1 Fig. 2 and 2016 doi:10.1093/gji/ggw040 Fig. 3), for a broad range of depths and size of anomalies (corresponding to spherical harmonic degree). In contrast, your results are just for particular anomaly depths and (rather small) size compared to what is seen in tomography.

We agree with the reviewer that the arguments at the beginning of the paper can be derived from radial stress or topography kernels for all wavelengths and depths (Hager and Clayton, 1989; Richards and Hager, 1989; Steinberger et al., 2001; Steinberger, 2016). As also pointed out by the reviewer, we imposed in our model a particular depth in the upper mantle and a relatively smaller anomaly size. It's true that one could follow a spherical harmonics approach in addressing dynamic topography. However, when the viscosity has lateral variations, spherical harmonic analysis becomes relatively hard to investigate analytically. When the mantle has non-linear rheology, all wavelengths of dynamic topography become coupled, and the degree of coupling should depend on how viscosity varies with temperature, pressure and strain-rate. Previous estimates using perturbation theory are insightful on understanding the impact of horizontal harmonic variations in viscosity on the dynamic topography as well the geoid (Richards and Hager, 1989), however these models assume that the lateral viscosity variations are *in phase* with the density anomalies, which is not necessarily the case in power-law rheology which introduces additional radial and lateral variations in viscosity as we show in our numerical models. Therefore, the variations in viscosity driven by a density anomaly in the upper mantle can reach well beyond the spatial dimensions of the embedded anomaly, and affects the mechanical lithospheric thickness. In that case, the dynamic topography varies considerably from the case where the mantle is isoviscous.

Based on above arguments, in our manuscript, we first introduced the Morgan (1965)'s analytical work on dynamic topography which uses radially layered viscosity model (up to two layers) for the Earth's interior. This helps us to easily compare it with more complex numerical models in the case for upper mantle density anomalies. Regarding the size of the anomaly, we explained our reasoning. A small radius minimizes the artefacts in the calculations and provides a better comparison with the analytical solutions carried out in an infinite half-space. The aim of this paper is not to predict the dynamic topography by using a density model derived from a seismic tomography (Steinberger et al., 2001, 2019; Flament et al., 2013), but to give insights on the first order changes in dynamic topography driven by non-linear rheology of the mantle. Having said that, we agree with the reviewer's point that the impact of larger anomalies, especially in the lower mantle, should be considered in future works.

I think for such small scales the effect of a low viscosity asthenosphere channel are stronger than for the larger scales seen by seismic tomography. E.g. in my 2016 paper Fig. 9a I find that one needs a very strong reduction in asthenosphere viscosity in order to get an appreciable reduction in topography, if anomalies are inferred from tomography. So I think the comparatively strong reductions in topography you show for a low-viscosity channel are partly misleading.

The assumptions about lithospheric thickness and radial viscosity may be effective in concluding the above mentioned argument. At the length-scales of our work (<1,000 km), the viscosity contrast as well reduction in the mechanical thickness of the lithosphere (via viscosity) affects the amplitude of dynamic topography. However, at longer scales (e.g. degree 2) with only radial variation in viscosity and imposed lithospheric thickness, one may need strong reduction in asthenospheric channel viscosity in order to get appreciably lesser amplitudes of dynamic topography. Furthermore, extra caution is necessary when comparing global seismic tomography models having horizontal resolutions of hundreds of km which is appreciably higher than the resolution of our numerical models (6-10 km's).

Also, in my Tectonophysics paper (doi:10.1016/j.tecto.2017.11.032) I find that the largest discrepancy by more than a factor 2 is at spherical harmonic degree two, whereas the discrepancy is much smaller at higher spherical harmonic degrees (i.e. smaller scales). It seems that your results could mainly explain a discrepancy at small scales, whereas the real discrepancy is at very large scales, and your results cannot explain this.

We agree that our model setup is not the best option to address the discrepancy at longest wavelengths or lowest harmonic degrees. We mentioned on the wavelengths being investigated in our numerical models by adding the following on Lines 20-22 in the revised text: "*Our experiments show that, at relatively short wavelengths (<1,000 km), the amplitude of dynamic topography, in the case of non-Newtonian mantle rheology, is reduced by a factor of ~2 compared to isoviscous rheology.*" However, we note that in non-Newtonian rheology, all wavelengths are coupled (Richards and Hager, 1989) and the dynamic topography at spherical harmonic degree 2, to a certain extent, will be influenced. In that regard, we added the following (on Lines 25-27 in the revised text): "*Although our results are strictly valid for flow wavelengths less than 1,000 km, we note that in non-Newtonian rheology all wavelengths are coupled, and the dynamic topography will be influenced.*" We would like to note that we seek to further investigate the effect of mantle viscosity on dynamic topography at longer wavelengths in the future.

Secondly, I think the usefulness of the models with LVV is severely limited because of the limitation of viscosities to the interval 10**19 Pas to 10**22 Pas. In this way, model 4b is almost the same as model 1 with constant viscosity (and giving very similar amplitude), model 4a approximately corresponds to the 2-layer model with topography accordingly increased, and model 5b has a low-viscosity channel with topography accordingly reduced. What I am puzzled about, though is that case 5a gives almost the same topography as 4a although it is also two-layer (although with thinner lithosphere). I think this limitation kind of beats the purpose of introducing a realistic rheology, because models essentially turn out to be a more complicated implementation of the easier models without LVV above. Also, I expect that without a cutoff, lowering activation enthalpy would not only lead to overall reducing viscosity, but also reducing viscosity contrasts. So, in contrast to your results I would expect a weaker lithosphere-asthenosphere contrast, and hence reduced dynamic topography for the lower activation enthalpy.

Exp. 5a and Exp 4a gives not similar topographies although their viscosity fields look like two-layer and similar to each other. As the reviewer pointed, the thickness of the lithosphere in both experiments are different, as of Exp4a having higher lithospheric thickness and therefore larger amplitude of dynamic topography. In Figure 7a of the manuscript, we show that, in Exp. 5a, the thickness of the lithosphere (in terms of viscosity) is reduced by about 30 to 45 km in comparison to Experiment 4a, which delivered a dynamic topography of ~149 m (in Exp. 5a) with the same viscosity contrast (Figure 7b,c). This experiment confirms that different mantle rheology results in different lithospheric thickness through local changes in viscosity. This is one of the key ideas of this paper.

In response to the reviewer's comment on the viscosity window, we run models with higher viscosity range. We increased the permitted viscosity range by two order of magnitude from [10^{19} Pa·s - 10^{22} Pa·s] to [10^{18} Pa·s - 10^{23} Pa·s]. In order to have better accuracy in the calculations, we kept the penalty value for convergence of the solution as low as before (0.03), however, due to our limitation on the resolution, the solutions had second order variations that we had to further smooth out the topography curves before picking their peak values. In the following supplementary figure 1, we plot the ratio of the viscosity field in the supplementary models to the viscosity field of the models in the manuscript, for each rheology and activation energy, along the same 2D cross-section. We give the change in dynamic topography, in percentage, also the minimum and maximum variation in viscosity across the entire numerical model. We also plot the viscosity ratio profile at the centre of the model (Fig. S2). In all rheologies, the lithosphere tends to have higher viscosity, if allowed, and the asthenosphere viscosity between the anomaly and the base of mechanical lithosphere doesn't show increase as much as the lithosphere. This results in an overall increase in viscosity contrast, therefore promotes an increase in dynamic topography. However, the magnitude of such increase was not notable because viscosity variations also affect the lithospheric mechanical thickness and/or the thickness of the low viscosity asthenospheric channel, which play a role in determining the amplitude of dynamic topography. In the case with dry olivine with low activation energy (Fig. S1-a), the viscosities inside the density anomaly show decrease, but the viscosities in the lithosphere show an overall increase. This results in a gradual increase in viscosity contrast between the lithosphere and asthenosphere, but also with modulation of the topography by the lower crust, limiting the increase in dynamic topography to 7.5 m. With dry olivine rheology with higher activation energy (Fig. S1-b) results in an increase in viscosity nearly across the entire cross-section, giving gradual increase in viscosity contrast between the mechanical lithosphere and asthenosphere above the anomaly. However, the thickness of the low viscosity asthenospheric channel is also increased, which is dominating the dynamic topography and resulting in gradual decrease by about 5 metres. The response of the dynamic topography to varying rheologies is similar in wet olivine rheology. These indicate that the viscosity window we used in the models doesn't severely limit the usefulness of them. However, we mentioned this variation (~5%) in the revised version of the manuscript by adding the following on Lines 329-333 in the revised text: "*Using a larger viscosity range in the models* ($10^{18} Pa \cdot s \leq \eta(P,T,\dot{\varepsilon}) \leq 10^{23}\ Pa \cdot s$) *resulted in ~5% variation in the amplitude of dynamic topography, indicating that the effects of non-linear rheology are reasonably captured in our models with smaller viscosity range* ($10^{19} Pa \cdot s \leq \eta(P,T,\dot{\varepsilon}) \leq 10^{22}\ Pa \cdot s$)."

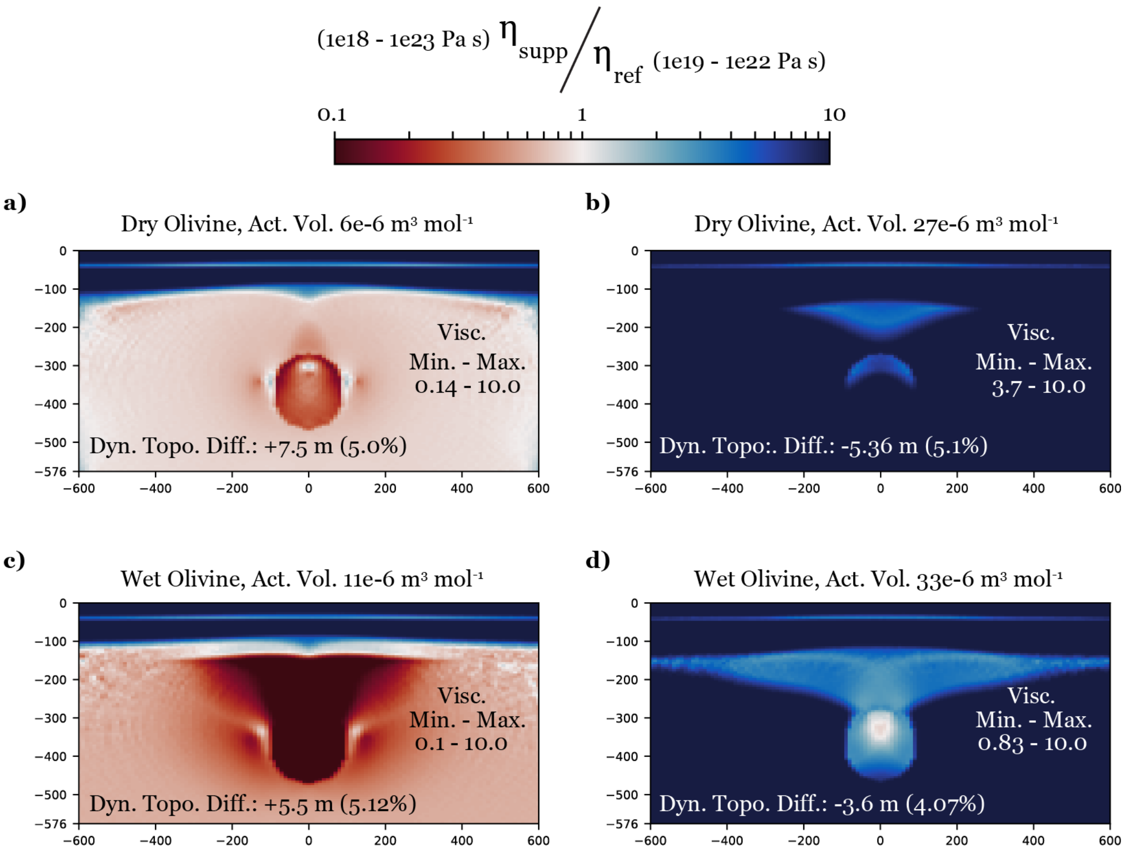

Figure S1: The ratio of viscosity fields of the supplementary models having wider viscosity window of 10^{18} Pa·s to 10^{23} Pa·s, to the models in the manuscript using a relatively narrower viscosity window of 10^{19} Pa·s to 10^{22} Pa·s. The change in dynamic topography and variation in min. and max. viscosities are given in the lower-left and middle-right for each model.

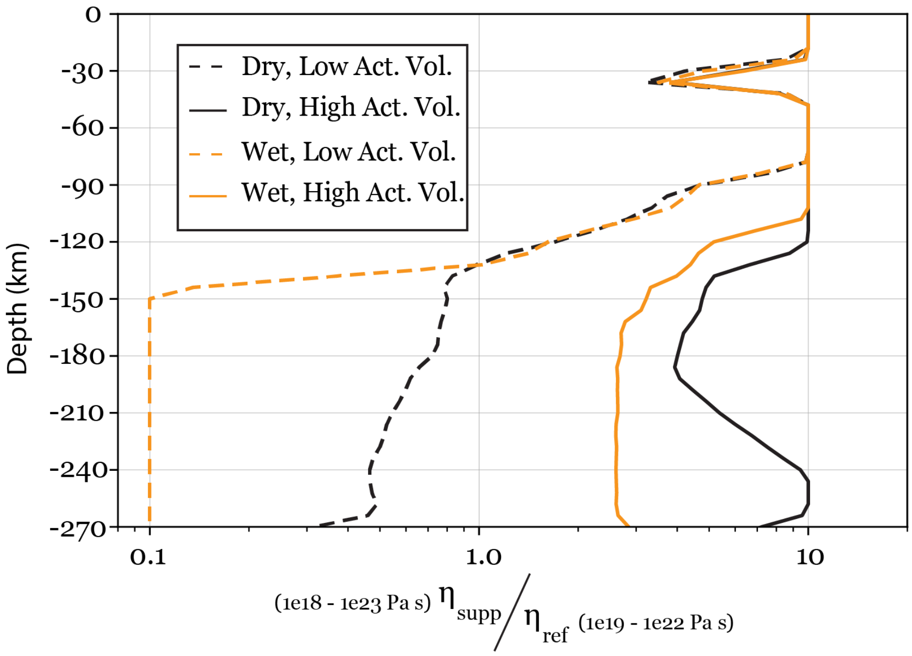

Figure S2: Variation of viscosities (ratios) at the centre of the numerical models.

Furthermore, reduction in activation enthalpy does result in overall reduction in viscosity, but the viscosity function also depends non-linearly on the strain rate, which strongly modulates this reduction, and introduces viscosity contrasts between different regions in the upper mantle. However, this effect could be lessened in diffusion creep.

In the following are a few more consecutive comments: l.54: as said, this large discrepancy is at the very largest scales, much larger than your model.

Thanks, we modified that line of the text accordingly by replacing the line stating: "*The problem is that dynamic topography predictions derived from mantle convection models are generally larger by a factor of two than estimates from residual topography (Cowie and Kusznir, 2018; Flament et al., 2013).*" with the following in the revised text (Lines 55-58) *"The problem is that dynamic topography predictions derived from mantle convection models are generally larger by a factor of two (more significant at the very large scales) than estimates from residual topography*

(Hoggard et al., 2016; Cowie and Kusznir, 2018; Davies et al., 2019; Steinberger et al., 2019).”

l. 220 this equation could actually be quite simplified. Because grain-size exponent p=0 the factor d**(p/n) is equal to 1 and therefore disappears. In each case, A**(-1/n)*f_H2O**(-r/n) is just a given number so you could simplify the equation in this way.

Thanks, we simplified that equation in Line 229 in the revised text. We kept other terms as it is because we use both dry and wet rheologies and different power-law exponent (crust vs. mantle).

l. 223 should be "volume and energy" (i.g. the other way round)

Thanks for the correction. We modified that sentence accordingly (Line 231 in the revised text).

l. 273 should be "wet olivine" (remove "dry").

Thanks for the correction. We modified it (Line 292 in the revised text).

l. 321 I don’t know where I would have said in that paper that the misfit demands a scaling factor ~0.35, It it true that one needs to downscale shallow seismic anomalies, but I believe this has nothing to do with viscosity structure; it is rather because the thermal anomalies and corresponding seismic anomalies in the lithosphere are largely compensated by chemical anomalies, with a much smaller seismic signature.

We should have cited only (Cowie and Kusznir, 2018) because this argument is only made there. We corrected that sentence accordingly in the Lines 336-339 in the revised manuscript. We find the reviewer's point about chemical anomalies and viscosity structure at shallow depths speculative to a certain extent. Our numerical model do show that viscosity variations at shallow depths determine the effective lithospheric thickness which can strongly affect the amplitude of dynamic topography.

Fig 1 a: Why the results for Morgan Hard Sphere and Molnar Hard Sphere are different? I think they are both analytical results, so they should be identical.

Thanks for the question. It mainly results from the higher order term that Molnar et al. (2015) takes into account which is not considered in Morgan (1965). The term of interest is $\varepsilon^2 = \frac{a^2}{D^2}$ where a is the radius of the spherical density anomaly and D is the depth of the centre of the anomaly that is mentioned in Appendix A3 in Molnar et al. (2015). We briefly mentioned on this in Lines 95-98 in the revised manuscript. Now it reads as "*For the same problem, Molnar et al., (2015) provided a solution by considering a higher order term resulting in a slight difference with Morgan (1965a)'s solution (see Appendix A3 in Molnar et al. (2015)) allowing to consider density anomalies of finite viscosity (η_{sphere}) (Eq. 3)*".

Fig. 7c: Viscosity 10**20 Pas or 10**21 Pas at the lithosphere-asthenosphere boundary both seems much too low to me.

We inferred the lithosphere-asthenosphere *transition zone* from the numerical models. We understand that using lithosphere-asthenosphere *boundary* could be misleading because of the viscosities cited, however, it's also true that defining a lithosphere-asthenosphere *boundary* is helpful and necessary when comparing lithospheric thickness between different numerical models. Fig. 7a shows that the viscosities show strong decrease from 10^{22} Pa·s to 10^{19} Pa·s in a few tens of km, which is a typical lithosphere-asthenosphere transition profile with an exponential decay constant of 5-12 km (Conrad and Molnar, 1997). However, it's also possible that with the higher viscosity window, this transition zone could include higher viscosities as the reviewer would expect. We mentioned on this by adding the following in Lines 276-278: "*We infer the mechanical thickness of the lithosphere from the viscosity profiles plotted in Figure 7a, along which the lithosphere-asthenosphere transition zone shows a rapid decrease in viscosity (Conrad and Molnar, 1997).*"

Table 1: For better comparison with text and eq. 5, you could also include the symbols (in those cases where you have defined them) in another column. I think the units for the pre-exponential factor should be MPa**-ns**-n (not -1)

We added the symbols in another column in Table 1. We also mentioned on the brittle deformation law we had defined and density of rocks at depth in the numerical models, an information that was missing in the main text. We confirm that the units for the pre-exponential factor is $MPa^{-n}\ s^{-1}$.

l. 28: write "from the surface"

Thanks, we modified that sentence (Line 31 in the revised manuscript).

l. 65: better "dependence ... on" ?

Yes, that sounded much better. We modified the text accordingly (Lines 67-68 in the revised text).

References

Conrad, C. P. and Molnar, P.: The growth of Rayleigh-Taylor-type instabilities in the lithosphere for various rheological and density structures, , 95–112, 1997.

Cowie, L. and Kusznir, N.: Renormalisation of global mantle dynamic topography predictions using residual topography measurements for "normal" oceanic crust, Earth Planet. Sci. Lett., 499, 145–156, doi:10.1016/j.epsl.2018.07.018, 2018.
Davies, D. R., Valentine, A. P., Kramer, S. C., Rawlinson, N., Hoggard, M. J., Eakin, C. M. and Wilson, C. R.: Earth's multi-scale topographic response to global mantle flow, Nat. Geosci., 12, 845–850, 2019.
Flament, N., Gurnis, M. and Müller, R. D.: A review of observations and models of dynamic topography, Lithosphere, 5(2), 189–210, doi:10.1130/L245.1, 2013.
Hager, B. H. and Clayton, R. W.: Constraints on the structure of mantle convection using seismic observations, flow models, and the geoid, 1989.
Hoggard, M. J., White, N. and Al-Attar, D.: Global dynamic topography observations reveal limited influence of large-scale mantle flow, Nat. Geosci., 9(6), 456–463, doi:10.1038/ngeo2709, 2016.
Molnar, P., England, P. C. and Jones, C. H.: Mantle dynamics, isostasy, and the support of high terrain, J. Geophys. Res. Solid Earth, 1–26, doi:10.1002/2014JB011724, 2015.
Morgan, W. J.: Gravity anomalies and convection currents: 1. A sphere and cylinder sinking beneath the surface of a viscous fluid, J. Geophys. Res., 70(24), 6175–6187, 1965.
Richards, M. A. and Hager, B. H.: Effects of lateral viscosity variations on long-wavelength geoid anomalies and topography, J. Geophys. Res. Solid Earth, 94(B8), 10299–10313, 1989.
Steinberger, B.: Topography caused by mantle density variations: observation-based estimates and models derived from tomography and lithosphere thickness, Geophys. Suppl. to Mon. Not. R. Astron. Soc., 205(1), 604–621, 2016.
Steinberger, B., Schmeling, H. and Marquart, G.: Large-scale lithospheric stress field and topography induced by global mantle circulation, Earth Planet. Sci. Lett., 186(1), 75–91, doi:10.1016/S0012-821X(01)00229-1, 2001.
Steinberger, B., Conrad, C. P., Osei Tutu, A. and Hoggard, M. J.: On the amplitude of dynamic topography at spherical harmonic degree two, Tectonophysics, 760(November 2017), 221–228, doi:10.1016/j.tecto.2017.11.032, 2019.