Dear reviewer,
thank you for your constructive criticism of our work. Taking into account your suggestions and feedback has certainly improved our manuscript. We hope that the new version is more nuanced and easier to follow. Particularly, we hope we could clarify what we want to achieve with this method and what is outside of its scope.
The original comments appear in regular font, while our responses are included in italics.

Anonymous Referee #1
Dear editor and authors: I have read this manuscript and am under the impression that it requires a major revision. Here are some comments/questions that I hope can be useful to you:

R1.MA1: What is the benefit of using "vote maps" over tomography models? Density does not necessarily need to be "scaled" from an existing tomography model: the scaled tomography model for example could be used as a target model towards which your inversion could be "regularized". This could allow for density/velocity decorrelation where required by the data. A similar philosophy was followed e.g. by Simmons et al. to derive their "Gypsum" model (2010). Why not try this approach as well?

There are two reasons for using votemaps. First, votemaps already exist and are the subject of ongoing research. Votemaps offer a way to distil a collection of seismic tomographies into two maps and have been used to judge the robustness of seismological determinations. If the areas of consensus do indeed correspond to real features, and if these features are able to explain the gravity field using reasonable density values, then that increases our confidence that the areas of consensus have a consistent density. Second, our goal is a direct inversion for mantle density, and votemaps provide adequate constraints for this. Note that this approach could very easily fail if the votemaps do not contain anything which is compatible with the observed gravity field. In that sense the fact that we ***can*** *explain the gravity field with the votemaps indicates that they contain something realistic.*
There are different ways to determine density without directly scaling it from velocity, such as the approach by Simmons et al. (2010). The advantage of our votemap approach is that it requires little to no additional assumptions from mineral physics and can work with only seismic data (which you need anyway) and gravity (which is readily available). In any case, since the research interest into thermochemical variations and decoupling of velocity and density will only increase, it seems timely to present our approach to the community.
That being said, we have investigated the question of velocity to density scaling in more depth. The density inversion results (with no constraints on the density values) were compared to the seismic velocities from SMEAN2, (section 5.2). This has led to some interesting observations, which either point towards highly variable composition or unmodelled dynamic effects in the mantle below the transition zone. Additionally, we developed an approach that regularizes toward a seismically derived density distribution but allows decorrelation if required. This proved less insightful than we hoped, because velocity/density decorrelation is essentially necessary to fit the gravity field with votemap constraints. Thus, it does not make a big difference, whether one regularizes towards a seismic-derived density, because it gets completely overwhelmed anyway. We have included these tests as supplementary information.

R1.MA2: In relation to the above: at lines around _50 the authors imply that "converting seismic velocity anomalies into density anomalies implies a purely thermal origin". I don't think that this is true. Many studies combine P and S velocity heterogeneities to estimate both thermal and compositional effects.

This was an oversimplified statement and was meant to refer specifically to scaling shear wave velocities to densities. In addition, a purely thermal anomaly leads to correlation between shear wave velocity and density, but this does not mean the converse is true, because some compositional anomalies might also lead to a correlation between shear wave velocity and density. We have rewritten lines 50-:

R1.MA3: lines 55-60. I disagree that there are "few compatible direct observations of core-mantle boundary deformation". See e.g. Soldati et al. GJI 2012, Soldati et al. G3 2013 (And a lot of literature before and after). Incidentally, you seem to imply here that "the gravity effect of any deformed boundary [...] can simply be calculated and removed"... but is this what you do? You seem to neglect the deltarho caused by the CMB in your inversion, to then compute it a posteriori? And how about the 660 and other internal discontinuities?

The gravity effect of any deformed boundary ***can*** *simply be calculated and removed – provided that the variation of the boundary and its density jump are known. For the surface topography this is clearly not a major problem. For the CMB it is a problem because there is no consensus regarding the pattern or magnitude of CMB topography as estimated with seismological techniques. There are thus two options, either calculate the CMB deformation using mantle convection or account for possible CMB deformation a posteriori. Since one of our main goals for this approach was to decouple density and viscosity inversion, we decided to use the second option. Our posteriori modelling of CMB topography is simplistic, because we disregarded lateral viscosity variations and the special*

convection dynamic of the LLSVPs themselves. However, these simple models implied that the main effect of CMB deformation is to compensate the gravity effect of the LLVPs, such that they can be included a posteriori. As reviewer #2 pointed out, there are scenarios, where the density structure inside the LLVPs decouples from the CMB deformation, in which case our assumptions would be incorrect. We have updated the discussion to include the uncertainties associated with this and a comparison with published estimates of CMB topography.
Other discontinuities such as the 660 km are not explicitly included in our inversion approach, but they should be contained in the votemaps if they cause a velocity change. Thus, the inversion should be able to pick up the signal of these discontinuity variations. However, they will be mixed up with the density variation due to the temperature or compositional changes that cause the depth variation of the continuity in the first place. A phase change without a strong seismic discontinuity will be more difficult to detect, but again the temperature and compositional variations that would lead to depth variations should be reflected in the vote maps.
In the calculations of dynamic topography kernels, we have not included latent heat release or buoyancy effects from phase transitions, because we only run one timestep of mantle convection with a given density model. Buoyancy effects from equilibrium displacement could also be included in instantaneous calculations. But they should be correlated to the temperature variations. Since we leave the densities unconstrained, it does not really matter whether density anomalies are due to temperature variations or the phase boundary displacements they cause.

R1.MA4: does this paper help us in any way to understand the earth better than we already do? The density maps (Figures 5 and 7) does not reveal anything new, or does it? The CMB map does not add anything to the many earlier observations and models (which at least to some extent would be useful to discuss, starting with Morelli and Dziewonski 1987).

We modified central assumptions that are typically made to determine density in the Earth:

- *Instead of (partly) correlating velocity to density, we used seismology as a geometry constraint. Density is a free parameter.*
- *We used actual topography instead of dynamic topography*
- *Still the topographical predictions of our approach are at least self-consistent, if a two-layer viscosity model is assumed.*
- *We highlight the impact of crust and upper mantle on density estimations in the deeper mantle*

With our new set of assumptions, we could partly detect expected features and have put our most reliable finding in the title of our paper. We also discuss the limitations of the proposed approach and possible improvements. A brief comparison with CMB topography models has been included in section 5.4.

R1.MA5: In summary, constraining the density structure of the mantle is an old problem that motivated a lot of work. I appreciate the authors' effort to contribute, but I have the impression that a major revision - expansion of their current work is needed for their contribution to be really helpful.

We do not agree that expansion of our work would be necessary for it to be interesting for the community. This paper is supposed to show an alternative approach of using the gravity field to learn something about the mantle. We are aware that this approach has its drawbacks, and we mention possible directions for future improvements in the conclusions.

some minor comments

R1.MI1: lines 70-71 "we stay close to the original source of the data" –> not sure what this means. –

We refer to downward continuing satellite gravity data to a lower height in order to improve the sensitivity (e.g. Uieda et al. (2017) used a similar data set at a height of 50 km). We instead choose to stay at a height of 225 km, the lowest measurement height of the gravity satellite mission GOCE. We have rewritten the paragraph for clarity

R1.MI2: page 6: the inversion algorithm could be described in a more transparent fashion. A clear notation should be adopted; for instance, either all vectors should be boldface, or none; it should be stated explicitly that g is a vector (if I understand correctly) whose i-th entry corresponds to the point r_i lat_i lon_i within the earth, etc. When you say "Aij is the gravity effect of potential density anomaly j on measurement point i", do you perhaps really mean that A_ij rho_j is the effect in question? Is A really a "design" matrix? What is a "design" matrix? Does the formula for the entries of A follow from eq. (4)? Could you then write explicitly a mathematical expression for A_ij? Or, if I misunderstood, explain explicitly how in practice A_ij is calculated?

We have extended the description of the inversion algorithm, hoping that it is now clearer for the reader what we actually do. We have dropped the term "design matrix" in favour of simply calling it a "matrix". An explicit formula for A_ij is given in terms of equation (4).

In what sense are L curves "ad hoc"? (I think you are right that there is a problem with L curves, but I don't understand what you mean by "ad hoc".)

We have reworded this section and added a reference to Farquharson and Oldenburg (2004), where L curve and cross-validation are compared.

R1.M3: line 358: "We assume a negative S-wave velocity deviation of 2 per cent"... After so much work to improve estimates of density anomaly, why is it OK to use such a rough estimate for velocity? Also, it has often be suggested that tomography might be systematically underestimating heterogeneity amplitude. Sincerely

Our inversion approach treats the entire LLVP (at each depth) as a large block with a constant density. Thus, we think it is appropriate to compare our density values with the average value of SMEAN2 tomography in the LLVP region as an indication of the average velocity of the LLVPs. Smaller scale variations like the ultra-low velocity zones or a potential change with depth inside the LLVPs are not resolved by our inversion. We have updated the section to better explain our choice of 2 per cent velocity deviation.

R1.T:
We have included your typographical corrections, thank you.
- line 114. "anomalous regions that is each" –> that are each... their... their ... unknown density valueS –
– caption of fig. 3 "opography"
- line 205 "are the similar"