The logic of the study is simple but the writing sometimes makes it convoluted and unnecessarily complicated. The worst example is the first paragraph of the Conclusions. Besides being unclear whether “stress state” refers to the state before or after the earthquake, or both, what is said here is incorrect at face value. In the linear system considered in this study, stress change is uniquely determined by fault stress drop, with no relation with the absolute state of stress before or after the event. If by “stress state” you actually mean "Coulomb stress", then nothing new is being said here, as Coulomb stress is always known to depend on pre-earthquake stress state. This paragraph should be deleted. There are similar situations in other parts of the paper. I leave it to the authors to check them out.
We have deleted the paragraph and revised other parts of the manuscript accordingly.

I think the discussion in Section 2.2 is unnecessarily detailed and thus distracting. Much of it (with figures) should be moved to the Supplement. The practical difference between the two definitions is very small, and I am not sure if the new definition has actual scientific advantage (see specific comment on line 230-231 below).
We have simplified the section 2 and reduced the number of figures. We have detailed some implications of the dynamic Coulomb wedge theory (DCWT) to better explain the advantage of our CFS calculations. Further details on the revision are provided in our response to the comment on line 230-231 below.
It is difficult to understand the finite element model results given the way they are presented. For example, in Fig. 9d, presumably the positive Coulomb stress in the lower crust in 200 – 280 km distance is for normal faulting, because Fig. 9a shows a steepening of sigma_1. But such steepening can also indicate less compressive stress instead of extensional stress. I am sure that the rigidity contrast across the model Moho is responsible for the lower-crust positive Coulomb stress, but I cannot tell how. It is not possible to understand the results based on the information shown in these plots (see specific comments on Fig. 9 below).
We have revised the presentation of the model results and now illustrate the stresses before and after the earthquake, as well as the coseismic incremental stress change, in terms of deviatoric stress and using stress crosses, similar to Figure 5 in Wang et al. (2019). The stresses in the models are always compressive; real tensional stresses do not occur. The lower crust in 200-280 km distance remains under deviatoric compression, as now shown by red stress crosses or red dashes in the previous version of the figure. It is right, that the rigidity contrast across the Moho influences the Coulomb failure stress change (DCFS). Without the rigidity contrast the DCFS is slightly negative in that region. We now address the rigidity contrast in the discussion section and provide supplementary models without the rigidity contrast (Figs. S9-S11). The new Table 1 summarizes how different model parameters affect the percentage of positively stressed aftershocks.

Other comments by line numbers:
58-- requires to account for -> requires accounting for
We have corrected the syntax.

79-- alpha is slope angle, not slope.
We have replaced “slope” with “slope angle” where necessary.

102-- I know that the author understands this but is trying to avoid bringing up more complications. Unfortunately, it cannot be avoided. Because Dahlen's lambda_b is a

small-taper approximation, so is his mu_b', and therefore his solution is not exact (as explained in Wang et al., 2006 GRL). Only if mu_b' is properly defined, will the solution be exact.
We have added this detail (lines 96-102).

138-- Because “dynamic weakening processes” is used to describe processes under high rate friction today, it is better to say “coseismic weakening” here.
We have changed “dynamic weakening” to “coseismic weakening”.

168-170. It is incorrect to use the absolute value operator here. For example, if tau flips direction such that tau_post = -tau_pre, this equation would incorrect yields a delta(tau) = 0 instead of 2*tau_pre. One just has to specify that tau is in the direction favouring slip as in King et al. (1994), then -tau will resist slip.
The calculation of DCFS differs for optimal failure planes and for faults of specified orientation. King et al. (1994) describe both approaches. The CFS on a failure plane is defined for the shear stress magnitude, i.e., for the absolute value (e.g., Oppenheimer et al., 1998; Reasenberg and Simpson, 1992; King et al., 1994; Harris, 1998). When calculating DCFS for optimal failure planes, as in our study, the use of the absolute value operator is correct, because the sign of the shear stress only reflects the sense of shear, that differs for the two failure planes. The change in tau in the example of the reviewer implies that the failure planes change their sense of shear (the one from sinistral to dextral, and the other from dextral to sinistral), but it would not bring the failure planes closure to failure. Thus, delta_tau should be zero. However, if DCFS is resolved on a fault of specified orientation, then the sign of the shear stress is defined as mentioned by the reviewer.
We understand that the latter calculation of DCFS is more common in the literature as it can be carried out without considering total stresses. Many readers may be more familiar with the approach, and the difference to our approach is essential. To acknowledge this situation, we now illustrate both approaches of calculating DCFS in section 2 (see revised Fig. 4) and briefly explain why the DCFS for optimal failure planes is the more appropriate choice for our study. In this regard, we have also detailed implications of the DCWT at the beginning of section 2. It should now become clear, that our DCFS values reflect the same tendencies as changes of mu_b in the mu_b-lambda space of the DCWT.

Fig. 4. The plots in (a) and (b) are switched by mistake and therefore contradict their headings at the top.
The figure has been removed.

230-231. Not a valid argument. The conventional definition does not require the knowledge of pre-existing weakness and stress anisotropy either.
The argument has been removed.

314-- I am curious how the code prevents numerical instability at large (e.g. 250 km) depths if gravity is applied as a body force. Because of the very large lithostatic stress, differences between principle stresses are beyond computer precision.
Abaqus allows using double-precision (with 64-bit word length) for the model execution, so we cross-checked and confirmed that the computer precision used for the models is sufficient.

385-389. I suspect the large mu_b'=0.2 adds much push against the upper plate. Without it, would mu_b' for the rest of the fault be larger than 0.015 to 0.022?
The mu'_b-pre values for the rest of the fault would not be larger, because the values are obtained by adding the Delta_mu'_b values (calculated from the mean stress drop) to the mu'_b-post values. The mu'_b-post are determined by finding the values required for deviatoric tension in areas of normal faulting.

Fig. 9a. I find it difficult to understand the model results because the display contains no information on shear stress and stress magnitude. Is it possible to plot stress crosses scaled with stress magnitude?
We now illustrate the stresses by stress crosses scaled with magnitude as in Wang et al. (2019).

Fig. 9c. Differential stress without direction misses important information. How do we know whether the change promotes compressive or extensional failure?
The type of promoted faulting depends on the state of stress. Normal faulting is promoted if the wedge is under deviatoric tension (plunge of s1 >45°); thrust faulting if the wedge is under deviatoric compression (plunge of s1 <45°). The plunge of s1 was indicated in panel a) and color coded (red: deviatoric compression, blue: deviatoric tension). We have now adjusted the illustration following Fig. 5 in Wang et al. (2019). Note that we also distinguish a plunge of s1 between 40-50°

Fig. 9e. Are the observed earthquakes in 200 – 280 km distance normal-faulting events? Presumably the positive Coulomb stress in the lower crust in this region shown in Fig. 9d is for normal faulting, because Fig. 9a shows a steepening of sigma_1. But such steepening can also indicate less compressive stress instead of extensional stress. I am sure that the rigidity contrast across the model Moho is responsible for the lower-crust positive Coulomb stress, but I cannot tell how.
The Coulomb failure stress changes are for optimal failure planes. Thus, it depends on the stress state and related plunge of s1 whether normal or thrust faulting is promoted (see previous response). The area in 200-280 km distance is under deviatoric compression such that thrust faulting is promoted. Note, however, that the plunge of s1 is close to ~40-45° in some areas, in which case both normal and thrust faulting may be supported. There are no Japan Meteorological Agency focal mechanism solutions for the earthquakes in 200-280 km distance. The catalogue of Yoshida et al (2012) show some reverse faulting near the coast, consistent with the stress state in our model.

415-- If it was extensional also before 2011 as said later in the text, it should be explained here. The figure only shows events after 2011. Nakamura et al. (2016) showed a mixture of reverse and normal events before 2011 in this area.
We have added this information (lines 406-409): "The stress state in the forearc along the Iwaki cross section is heterogenous before the earthquake (Fig. 8a). Most of the submarine forearc is in a neutral stress state (plunge of s_1 40-50°) or under deviatoric tension, which is compatible with mixed reverse and thrust faulting reported for the years before the Tohoku-Oki earthquake (e.g., Hasegawa et al., 2012; Nakamura et al., 2016; Yoshida et al., 2012)."

427-- The small mu_b-pre values used here may be needed to keep the stress drop low so that is does not exceed the values shown in Fig. 10b. However, there was large shallow afterslip in this area, and the total stress drop responsible for the aftershocks is larger than the coseismic stress drop shown in Fig. 10b. Iinuma used GPS over a much shorter time window that reflects mostly coseismic change, but the aftershocks are affected by the afterslip which continues to relieve stress on the megathrust over a longer time.

The mu_b-pre are obtained by adding Delta-mu_b to the mu_b-post values as explained in section 3.2. Thus, the mu_b-pre values do not need to be low because of the stress drop. The values are low because the stress drop is small and because the mu_b-post values need to be low to cause deviatoric tension in areas of normal faulting (especially near Iwaki). We agree that stress release due to afterslip may bias the inferred extent of deviatoric tension. We now address this aspect at the end of section 3.2 (lines 335-342). "… it should be noted that most of the post-mainshock focal mechanisms indicate normal faulting, some of which may have been caused by afterslip and aftershocks on the megathrust in the postseismic period (e.g., Bedford et al., 2016; Nakamura et al., 2016; Sun et al., 2014). Such events may record stress release on the megathrust in addition to the coseismic stress drop and influence our assessment of the post-seismic stress state. We expect this potential effect to be small on our calculations because the normal faulting started soon after the mainshocks (e.g., Farías et al., 2011; Lange et al., 2012; Yoshida et al., 2012; Japan Meteorological Agency) and affected always the same forearc areas in the first postseismic year (Figs. S3-S4 in Supplement)."

511-512. The second point is not very useful. More fundamental is the plunge which shows tension vs. compression.

We agree and have simplified the entire paragraph.

514 onward. Poor writing. Reverse the sequence by first saying flat surface can only allow compression, both before and after an earthquake…

The sentence has been removed.

521-522. can promote … only if …

The sentence has been removed.

540-- There should be a distinction between pervasive and local failure. See discussion in Section 4.4 of Wang et al. (2019). In my view, the lack of recognition of potentially very large, multi-scale heterogeneity is the biggest shortcoming of Coulomb stress analysis as is commonly conducted. I do not ask the authors to solve this problem in this work, but some qualitative discussion will be useful.

We have added some qualitative arguments (lines 516-524): "The requirement of weak faults for failure also implies that their absence may cause tectonic quiescence even though the Coulomb failure stress increases. High pore fluid overpressures may be difficult to sustain through time and over large areas such that only small fractions of the forearc lithosphere may be close to failure. Stress changes caused by megathrust earthquakes may therefore preferentially drive small earthquakes (Wang et al., 2019). Consistently, the vast majority of the earthquakes investigated in this study have low magnitudes of about 2.5-3.5 and record local failure on small faults. However, the aftershock seismicity of both mainshocks also included damaging earthquakes with magnitudes of 6.6-7.0 inland Japan near Iwaki (Fig. 5a) and in the coastal region near Pichilemu, Chile (Fig. 6a). The large magnitude aftershocks occurred in earthquake clusters affecting the entire crust down to 20 km (Iwaki) and 35 km (Pichilemu) depth, which shows that megathrust earthquakes can cause pervasive failure in the interior of forearcs"

576-- cause -> would necessitate
We have changed “cause” to “would necessitate”.

589-590. I have a hard time finding what mu value was used for all the other models. Was it 0.7? It should be prominently stated somewhere, and the reader should be reminded here again.
All main model results are obtained for mu = 0.7. We have added this information to the figure caption and state it more prominently in the first paragraph of section 4, which now provides basic information on the presented model results.