

Peer Review File

Manuscript Title: Phonon modes and electron-phonon coupling at the FeSe/SrTiO₃ interface

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referees' comments:

Referee #1 (Remarks to the Author):

The manuscript showcases a very nice application of vibrational EELS to determine and understand the structural origin of the electron-phonon coupling at the FeSe/STO interface. The scientific context is presented appropriately according to existing literature. The experimental data are of high quality and demonstrate the ability of this group to perform high-end and challenging EELS experiments. For that, the authors make use of dark-field EELS at atomic resolution, which has already been demonstrated several years ago and has been used by the same group of authors in previously published work, as this technique is now becoming more common among vibrational EELS specialists. Therefore, I would expect the novelty of this work to be about the new information brought on the FeSe/STO interface. Despite the high quality of experimental data, the discussion about contrast variations in atomically resolved phonon spectral maps is not as convincing as the authors imply. Furthermore, the absence of a proper match between experimental vibrational EEL spectra and the calculations (DOS) proposed in this study is all too evident to support some of the claims of the authors. This is limiting and detrimental to the interpretation of the spectra, and is therefore a major downside that prevents this manuscript from being convincing. I provide additional suggestions and comments below.

- Page 3, the authors write that “The Ti-L edges in Ti-O layer 0 show a small redshift of about 0.3 eV, indicating that a slight reduction possibly due to the oxygen vacancies.”. From figure 1d, it appears that the eg peaks are redshifted rather than the whole Ti-L edge.

Page 3, the authors write that “The similar near edge fine structures in the Ti and O core-loss suggest that the interfacial Ti and O are likely in a similar bonding environment as in bulk STO”. This statement brings some confusion, and the similarities in near edge structures and in environments are very relative. Despite the redshift, it seems reasonable to assert that the Ti-L edges indeed show similar features at layers 0, 1 and bulk STO. However, there are major changes to the O-K edge fine structures, making this edge quite dissimilar in layers 0 and 1, compared to bulk STO.

- The authors begin to interpret the spectral variations of phonons from the agreement between calculated DOS and experimental EEL spectra. However, as mentioned above and contrary to the claims of the authors, several DOS shown in figure 2 hardly reproduce the experiments. For instance, in figure 2b (FeSe) for which the intensity ratio of the peaks at 15 and 35 meV (DFT) is opposite to the experimental spectrum FeSe/STO (blackline). But also, and more importantly in figure 2e (Ti-O

layer 1) and 2f (Ti-O layer 0) where the agreement in terms of energy position and intensity is quite poor for most phonon modes and especially above 50 meV.

One could wonder about the meaningfulness of these DOS calculations, especially since the peak intensity/position are not discussed beyond the overall agreement between calculations and experiments, which can already be argued upon.

The need for relevant calculations is even more obvious from the absence of theoretical support for the OP vs IP signals. Theoretical spectra with the appropriate directionality and magnitude in q space appear necessary to ensure proper interpretation of the experimental anisotropy and draw solid conclusions about the FeSe/STO interface.

- What is the criterion to define a strong or weak peak, as indicated by the filled or empty arrows in figure 2? Visually, it is true that some peaks are stronger than others, qualitatively speaking. But in some cases, the choice for filled or empty arrows is not obvious. For instance, the optical phonon peak at ~ 99 meV in Fig. 2i “O, OP” is not much more intense than the ones for “O, IP” or “Ti, OP”.

- In the 45 meV OP phonon map in figure 3 the contrast actually appears vertical, which is opposite to the maps for the 66 and 99 meV modes (horizontal contrast). Could you comment on this feature? Same remark for the IP bulk phonons at 45 meV (horizontal contrast) vs 66 and 99 meV (vertical contrast).

- In Figures 3c-f, the authors show phonon maps at different energies and discuss the localization of phonon modes and their IP/OP selectivity. This figure is a central result of this study, from which the authors build their discussion. Overall, the OP phonon maps at 75 and 84 meV show more intensity at the interface, but at the same time this is also not very clear. Maybe intensity profiles could make this more obvious? Also, the authors could indicate the position of Ti-O layers 0, 1 next to each map to make their arguments more convincing.

Without the help of the ellipsoid and the arrows in figure 3e, the localization of phonons on specific atomic columns in the OP maps would be much more difficult to distinguish. Here as well, although there are intensity variations that seem to be correlated to the position of atomic columns, this is not very clear. For instance, there are also intensity variations among atomic columns in layer 0 in all maps of figures 3e and 3f, and especially in the 75 and 83 meV maps (IP signal). What do these variations correspond to? What distinguishes phonon variations from noise level?

- Page 7, the authors write “The interfacial double-TiO_x layer show strong vibrational signal near 55, 75, and 84 meV, as shown in Fig. 3(e) and (f).”

With the help of the arrows and ellipsoids, it is possible to distinguish phonon contrast in the two highest energy maps, even though qualifying the vibrational signal as “strong” is arguable. However, this statement is not very convincing in the 55 meV map, which actually looks somewhat similar to the 45 meV and 66 meV maps in the region of the interface. One could also argue that the contrast in STO (away from the interface) in the 55 meV map resembles the 45 meV map. Finally, the 55 meV mode does not seem to correspond to a peak in the spectra shown in Figure 2; how was the energy of this particular peak defined for bulk vs interface phonons?

- The authors mention the variations of intensity of the FeSe phonons in OP vs IP (Fig.3 c,d). How

influential is the extraction of the elastic peak on the intensity of the FeSe phonons at 15 meV?

- Page 7, the authors mention that “The interfacial double-TiOx layer show strong vibrational signal near 55, 75, and 84 meV, as shown in Fig. 3(e) and (f). This contrasts with the bulk STO phonon in Fig. 3(c) and (d), whose intensity decreases approaching the interface.”

This “contrast” between bulk and interfacial phonon is not convincing from the maps. Here as well, maybe intensity profiles could be of interest to support this claim.

- Several experimental results reported in figures are not discussed nor even mentioned in the manuscript. This is the case for the modulation of intensity of some IP/OP phonon modes (filled/empty arrows) as well as lower energy peaks in fig 2, but also some phonon maps in Fig. 3 (45 and 55 meV), or the phonon intensity variations across the interface in Fig. 5d (FeSe + Sr, 35 meV), etc. As a general remark, some results are shown but it is not entirely clear why.

- Figure 5d is lacking information and explanations: how are the phonon intensity plots generated from raw data? What do the red and grey shaded areas represent? Do the phonon intensity profiles correspond to OP or IP, or both?

- Caption of Figure 5d: “The large and small dFe-Ti are indicated by the red and black arrows.” Where are the arrows?

- Page 10, referring to figure 5c the authors mention that “In the 1uc FeSe/STO sample annealed at 500°C, we mostly observe the double-TiOx interface between 1uc FeSe and STO and thus a small dFe-Ti, as shown in Fig. 5(b).”. This claim is difficult to connect with Figure 5c, where two data points are represented for the 1uc sample annealed at 500C, one at small d(Fe-Ti) and one at large d(Fe-Ti). More generally, in Figure 5c, there are two data points for each 1 uc sample, one at low dFe-Ti (grey area) and one at high dFe-Ti (red area). Does it mean that for each 1uc sample and processing conditions, two areas have been investigated and yield two opposite results? More clarity would be expected from this plot and the description of these results.

Miscellaneous comments

- It appears that the same overlapping orange spectrum is plotted in figure 2d,e,f. However, in the caption, it is indicated that “The overlapping orange spectra are from SrO or TiO2 layers in STO”. I understand that one would expect an overlapping SrO layer spectrum in fig 2d, and a TiO2 layer spectrum in fig 2e, f. This does not appear to be the case in the current figure.

- In figure 2g,h,i, the black curves look like filtered data plotted atop the experimental data points. Please provide more information.

- For ease of visualization, indicate the position of “Ti” and “O” columns of layers 0, 1, in some maps of figure 3c-f.

- The connection between the phonon bandstructure and the Eliashberg spectral function would be

easier to visualize if the energy scales were matching in figures 5b,c.

- Extended data Fig. 7: the EELS spectra have been acquired on axis, as mentioned in the main text. However, this could also be indicated directly in the figure or caption.
- Page 5: “Also shown in Fig. 2(b) is the DF-EEL spectrum of a bulk FeSe sample (blue curve)”. The color is inconsistent, replace “blue” by “orange”.
- Typos and grammar could be improved in the manuscript, and in SI. Here are some examples:

Manuscript:

Page 6: “by integrated within” -> by integration within ; “Depends on the” -> Depending on the
Caption of Figure 3: (b) and (c) -> (c) and (d) ; (d) and (e) -> (e) and (f)

Page 7: “at as such” -> at such

Page 8: “these band” -> bands ; “Fig. Sx2”

SI, page 1: “is consist”, “these... consists”,

SI, Page 2 “can be substantially influence” -> influenced

SI, caption figure S3: “Fromation” -> Formation

- The caption of Figure S3 does not match the embedded legend. Also, which side of the figure is Ti-rich / O-rich?

Referee #2 (Remarks to the Author):

The paper presents a highly ambitious measurement of atomically resolved phonons at an interface in order to tie into microscopic models of superconductivity enhancement at the FeSe/STO interface. The data appears to be cutting edge and begs a clearer discussion of the data. The plots in Figs. 2 and extended data of Fig. 5, especially, deserve a more careful discussion.

1) What are the black solid lines in Fig. 2? Please define.

2) The peaks in the black solid lines of Fig. 2 e through h aren't all readily apparent in the data, especially for Figure 1i on Ti-O layer 0. Please discuss discrepancy.

3) The PhDOS from DFT does not clearly match the data, there is a minimum in the DOS at ~ 70 meV where the data shows a clear peak. Some of the discrepancy between theoretical DOS and measured may be explained by the different structures presented in the supplement. Can the authors get an image along the [110] zone axis of the STO to locate the position of the Ti and oxygen and distinguish between different structures?

4) The term "interface phonon" should be defined earlier in the manuscript.

5) Are the statistics good enough to make definitive claims, such as "...show strong vibrational anisotropy" (bottom of page 5) and "a gradual softening of the O phonons" (bottom of page 8)? Is one supposed to see a shift in the peak position or height? What are the heights? How are they fit? The statistics on the momentum resolved data (especially for the 0 layer) are probably not good enough to make general statements on peak position. For layer 0, the only well defined peak is the Ti-OP peak in Fig. 2i. It would help the reader to not point out peaks that aren't clearly apparent.

Figure 3 shows good promise and is a nice figure but can be improved. Can the authors show theory derived images to highlight what the reader should be looking for in the noisy data?

Is the false color image in Fig. 3 proportional to I as suggested by the plot legend?

If so, how is $I \times E^2$ calculated? Is it from a fit, or is it the raw intensity?

The raw spectra should also be shown in the extended figures to show the elastic peak in order to display the magnitude of the signal relative to the background.

It is known FeSe T_c varies dramatically with sample preparation. The T_c of the samples could be measured (for example with resistivity measurements as a function of temperature) to verify T_c . Here, it seems measurements of sc are limited to STM measurements, mostly at fixed temperature (5 K).

Figure 4b shows many bands. How are the highlighted ones chosen to be the relevant ones?

This manuscript promises a high impact, especially if revisions of the discussion of the EELS data are made to improve the clarity of the manuscript.

Referee #3 (Remarks to the Author):

Using high-resolution electron energy-loss spectroscopy (EELS), the authors present a study of the phonon modes near the interface of monolayer FeSe on SrTiO₃. Surprisingly, they find the existence of an additional TiO₂ layer with new out-of-plane optical phonon modes in the energy range 75-99 meV. Based on DFT calculations it is argued that these modes have a strong electron-phonon coupling and can thus be related to the enhanced T_c in FeSe/SrTiO₃.

I think this work is highly original, as it uses state-of-the-art experimental methods to address a timely issue, namely the interfacial phonon couplings in FeSe/SrTiO₃. In general the paper is well written, provides a lot of interesting insights. There are minor English mistakes here and there. Their conclusions are robust and warrant further investigations in possible phonon modes at the interface of FeSe and STO.

So while I would like to see this paper published as soon as possible, I do have some questions that need to be answered, mainly with regard to the theoretical parts of the paper. The questions below need to be answered and the manuscript should be changed accordingly:

- What is T_c ? The authors report the gap in the dI/dV at 5 K but I cannot find any data on the actual T_c of their samples.
- About the DFT results: The extra TiO₂ layer "0" causes a lot of metallic Ti bands to cross the Fermi level (eg. Extended Data Figure 4). This is in contradiction to various ARPES results (for example Ref 2 or 19) that clearly see only one FeSe band crossing the Fermi level. How do the authors explain this?
- The DFT results of Fig. 4 are presented in an unclear way. The (c) panel is very difficult to read because the dashed line is, well, not a continuous line. Moreover, the vertical axis ticks of panel (c) do not align with the vertical axis ticks of panel (b), making it impossible to read at which energy there are peaks in α^2F . Finally, the peaks in α^2F seem to happen at 18, 65 and 79 meV, which is quite different from what the main text and the panel (a,b) are trying to convey.
- The text "The comparison of phonon dispersions with and without the AFD phase transition is discussed in Fig. Sx2." refers to a figure Sx2 but I can't find which figure that is supposed to refer to.
- In the introduction, a difference is made between FK phonons and "oxygen optical phonons". Since in either case we are dealing with optical phonons on the surface of STO, I don't really see why the authors emphasize there is a difference here.
- The key theory of why phonon coupling can enhance T_c in FeSe/STO is based on the concept of "forward scattering", ie, only small-momentum transfer between FeSe electrons and out-of-plane optical phonons in the substrate. The whole paper does not mention this concept, even though their findings are consistent with this theory. For example, in Fig. S5 in the supplementary they clearly show that the electron-phonon coupling is strongest at Γ (ie. small momentum). Early theories [for example Rademaker et al New J. Phys. 18 022001 (2016)] were based on the idea that the relevant phonon coupling is from an oxygen optical mode in the STO, not (as this paper suggests) from the new optical modes from the interfacial TiO₂ layer. This paper really needs to discuss the forward scattering mechanism; in particular: how the theory of forward scattering changes

considering the 'new' optical modes discovered here. Furthermore, estimates of α^2F are not enough to properly take into account the enhancement due to forward scattering, and therefore a momentum-dependent Eliashberg calculation might be warranted.

Author Rebuttals to Initial Comments:

Point by point response

(Referee remarks in black, our response in blue.)

Referee #1:

The manuscript showcases a very nice application of vibrational EELS to determine and understand the structural origin of the electron-phonon coupling at the FeSe/STO interface. The scientific context is presented appropriately according to existing literature. The experimental data are of high quality and demonstrate the ability of this group to perform high-end and challenging EELS experiments. For that, the authors make use of dark-field EELS at atomic resolution, which has already been demonstrated several years ago and has been used by the same group of authors in previously published work, as this technique is now becoming more common among vibrational EELS specialists. Therefore, I would expect the novelty of this work to be about the new information brought on the FeSe/STO interface. Despite the high quality of experimental data, the discussion about contrast variations in atomically resolved phonon spectral maps is not as convincing as the authors imply. Furthermore, the absence of a proper match between experimental vibrational EEL spectra and the calculations (DOS) proposed in this study is all too evident to support some of the claims of the authors. This is limiting and detrimental to the interpretation of the spectra, and is therefore a major downside that prevents this manuscript from being convincing. I provide additional suggestions and comments below.

Thank you for the positive remarks and raising the important questions. We have conducted more refined experiments and calculations, and have substantially revised the manuscript to address these concerns. Please see our detailed response below.

1. Page 3, the authors write that “The Ti-L edges in Ti-O layer 0 show a small redshift of about 0.3 eV, indicating that a slight reduction possibly due to the oxygen vacancies.”. From figure 1d, it appears that the e_g peaks are redshifted rather than the whole Ti-L edge.

Thanks for pointing this out, we agree that the t_{2g} peaks redshift by a smaller amount compared to the e_g peaks. Therefore, we have corrected the description for this part and added labels for the t_{2g} and e_g peak in the core-loss EELS plot.

Page 3, the authors write that “The similar near edge fine structures in the Ti and O core-loss suggest that the interfacial Ti and O are likely in a similar bonding environment as in bulk STO”. This statement brings some confusion, and the similarities in near edge structures and in environments are very relative. Despite the redshift, it seems reasonable to assert that the Ti-L edges indeed show similar features at layers 0, 1 and bulk STO. However, there are major changes to the O-K edge fine structures, making this edge quite dissimilar in layers 0 and 1, compared to bulk STO.

Thank you for pointing out the difference in O-K edge. We have also clarified the statement regarding the bonding environment and difference in O-K edge in the same paragraph. The difference in O-K edge is mainly in the second peak, which is due to the empty Sr 4d states. This also reflect the fact that the surface double TiO_x termination.

2. The authors begin to interpret the spectral variations of phonons from the agreement between calculated DOS and experimental EEL spectra. However, as mentioned above and contrary to the claims of the authors, several DOS shown in figure 2 hardly reproduce the experiments. For instance, in figure 2b (FeSe) for which the intensity ratio of the peaks at 15 and 35 meV (DFT) is opposite to the experimental spectrum FeSe/STO (blackline).

We agree with the referee that the agreement between vibrational EELS and DFT PhDOS in Fig. 2 is less ideal. The discrepancies can be attributed to several factors such as multiple scattering, electron channeling, and the influence of phonon polaritons in the experiment, as well as the limited model size in the calculations.

For Fig. 2b, the EELS of FeSe film (black line) includes contributions from the substrate phonon polaritons (PhPs). The energy overlap between FeSe phonon and one of the STO PhPs alters the overall spectral shape. In the revised manuscript, we instead show the phonon energy loss in the main text, and provide the detailed analysis regarding PhPs in DF-EELS in the supporting information.

But also, and more importantly in figure 2e (Ti-O layer 1) and 2f (Ti-O layer 0) where the agreement in terms of energy position and intensity is quite poor for most phonon modes and especially above 50 meV.

To improve the agreement between theory and experiment for Fig. 2e and f (now Fig. 1f and g), we have performed additional DFT calculations considering structural models with varying Ti/O ratios in the Ti-O layer 0. Figure S4 now shows a comparison between phonon EELS and PhDOS from various structural models. Newly calculated PhDOS of a Ti_{1.75}O₂ interface structural model shows improved agreement with experiments. This indeed suggests that the vibrational properties of the double TiO_x layers are strongly influenced by the interface structure and bonding (Ti to O stoichiometry).

One could wonder about the meaningfulness of these DOS calculations, especially since the peak intensity/position are not discussed beyond the overall agreement between calculations and experiments, which can already be argued upon.

The need for relevant calculations is even more obvious from the absence of theoretical support for the OP vs IP signals. Theoretical spectra with the appropriate directionality and magnitude in q space appear necessary to ensure proper interpretation of the experimental anisotropy and draw solid conclusions about the FeSe/STO interface.

We thank the referee for this comment, which inspired us to perform additional calculations and analyses. We agree that it is important to consider the EELS collection in simulation. Therefore, in the revised manuscript, we adopt a new approach for the interpretation of the vibrational EELS data. We have performed FRFPMS simulation with DFT data as input. The results take into account both the PhDOS and the direction + magnitude in q -transfer. The simulated phonon spectral images are now directly comparable with experimental ones in the new Fig. 2 (Fig. 3 in previous version). This approach provides more detailed comparison between experiments and simulations as it allows visualizing the vibrating atomic columns at each phonon energy window.

3. What is the criterion to define a strong or weak peak, as indicated by the filled or empty arrows in figure 2? Visually, it is true that some peaks are stronger than others, qualitatively speaking. But in some cases, the choice for filled or empty arrows is not obvious. For instance, the optical phonon peak at ~ 99 meV in Fig. 2i “O, OP” is not much more intense than the ones for “O, IP” or “Ti, OP”.

The relative intensity variation of the high energy optical phonons originates from the vibrational anisotropy of the oxygen atoms. We agree that this variation is not as obvious in some EEL spectra. This is especially true for the “Ti” columns, which involve both oxygen and titanium atoms along the beam direction. Therefore, we have clarified the discussion by only showing the O column results in the main text. We have also performed additional EELS experiments to improve the statistics of the EEL spectra and phonon spectral images. With the improved EELS data, the strong and weak peak intensity variations are now more clearly presented. The interpretation of the atomic column resolved spectra by combining them together with the phonon maps in Fig. 2.

4. In the 45 meV OP phonon map in figure 3 the contrast actually appears vertical, which is opposite to the maps for the 66 and 99 meV modes (horizontal contrast). Could you comment on this feature? Same remark for the IP bulk phonons at 45 meV (horizontal contrast) vs 66 and 99 meV (vertical contrast).

In the revised version, the phonon maps have been explained in more detail. The 45 meV image shows the Ti+O columns of STO. The 40-70 energy range indeed shows

horizontal and vertical contrast because of the rotational modes of the TiO_6 octahedra. The 99 meV images are from the O breathing modes. These details are now discussed in Fig. 2 and Extended data Fig. 4.

5. In Figures 3c-f, the authors show phonon maps at different energies and discuss the localization of phonon modes and their IP/OP selectivity. This figure is a central result of this study, from which the authors build their discussion. Overall, the OP phonon maps at 75 and 84 meV show more intensity at the interface, but at the same time this is also not very clear. Maybe intensity profiles could make this more obvious? Also, the authors could indicate the position of Ti-O layers 0, 1 next to each map to make their arguments more convincing.

To further improve the contrast and statistics of the phonon maps, we have performed additional EELS experiments, as mentioned in the answer to question 3. The phonon spectral images shown in the new Fig. 2 and extended data Fig. 4 clearly show that the 75-84 meV vibrations are stronger at the interface. Positions of Ti-O layers 0, 1 are labeled next to each map. Intensity profiles of 75 and 99 meV phonons across the interface are shown in Fig. 4. Thank you for the suggestion.

Without the help of the ellipsoid and the arrows in figure 3e, the localization of phonons on specific atomic columns in the OP maps would be much more difficult to distinguish. Here as well, although there are intensity variations that seem to be correlated to the position of atomic columns, this is not very clear. For instance, there are also intensity variations among atomic columns in layer 0 in all maps of figures 3e and 3f, and especially in the 75 and 83 meV maps (IP signal). What do these variations correspond to? What distinguishes phonon variations from noise level?

Thank you for the question. With the additional EELS experiments, we further improved signal-to-noise ratio of our EELS data, we believe that the phonon signals can be clearly distinguished now in Fig. 2 and extended data Fig. 4 even without the ellipsoids. The contrast variation on the atomic columns can also be better observed from the revised Fig. 2.

6. Page 7, the authors write “The interfacial double- TiO_x layer show strong vibrational signal near 55, 75, and 84 meV, as shown in Fig. 3(e) and (f).”

With the help of the arrows and ellipsoids, it is possible to distinguish phonon contrast in the two highest energy maps, even though qualifying the vibrational signal as “strong” is arguable. However, this statement is not very convincing in the 55 meV map, which actually looks somewhat similar to the 45 meV and 66 meV maps in the region of the interface. One could also argue that the contrast in STO (away from the interface) in the

55 meV map resembles the 45 meV map. Finally, the 55 meV mode does not seem to correspond to a peak in the spectra shown in Figure 2; how was the energy of this particular peak defined for bulk vs interface phonons?

With the further improved EELS data, the energy of the interfacial vibrations are now clear. The OP interfacial phonons are observed between 75-88 meV, whereas the IP interfacial phonons are found at 55 meV and 75-88 meV, as shown in the new extended data Fig. 4. Previous fig. 3 and the corresponding main text have been revised accordingly. This observation is supported by our new FRFPMS simulations.

Regarding the 50-55 meV map, it is true that the interface/bulk contrast is not as strong. We have therefore performed additional experiments and verified that this energy indeed exhibit enhanced interface vibration, especially in the IP EELS. This is also supported by the difference between bulk and interface vibrational EEL spectra, as shown in Fig. 1(f, g).

7. The authors mention the variations of intensity of the FeSe phonons in OP vs IP (Fig.3 c,d). How influential is the extraction of the elastic peak on the intensity of the FeSe phonons at 15 meV?

Upon examining the OP and IP EELS of FeSe, we find that the ZLP height is indeed different in the two configurations. This is due to the difference in diffraction condition of FeSe, which has stronger diffraction intensity in its IP directions. This also leads to stronger phonon signal. However, since the phonon anisotropy of 2d FeSe is not directly related to the EPC, this discussion has been removed in the revised version.

8. Page 7, the authors mention that “The interfacial double-TiOx layer show strong vibrational signal near 55, 75, and 84 meV, as shown in Fig. 3(e) and (f). This contrasts with the bulk STO phonon in Fig. 3(c) and (d), whose intensity decreases approaching the interface.”

This “contrast” between bulk and interfacial phonon is not convincing from the maps. Here as well, maybe intensity profiles could be of interest to support this claim.

With the improved phonon spectral maps in the revised manuscript, the contrast between bulk and interface is now clearer. Since some vibrational modes of STO and double TiOx layers overlap in energy, the experimental map of interface modes still contains STO contributions. Intensity profiles of 75 and 99 meV phonons across the interface are plotted in Fig. 4. Thank you for the suggestion.

9. Several experimental results reported in figures are not discussed nor even mentioned in the manuscript. This is the case for the modulation of intensity of some

IP/OP phonon modes (filled/empty arrows) as well as lower energy peaks in fig 2, but also some phonon maps in Fig. 3 (45 and 55 meV), or the phonon intensity variations across the interface in Fig. 5d (FeSe + Sr, 35 meV), etc. As a general remark, some results are shown but it is not entirely clear why.

Thanks for pointing this out. In the revised version, we have rearranged the fig. 1, 2, 3 to make sure all results are discussed clearly. Additional data and description are now presented in the extended data and SI, as mentioned in answers to previous questions. For example, the 40-70 meV images are now integrated into one in the main text Fig. 2. The detailed images with finer energy windows are shown in the new extended data Fig. 4.

10. Figure 5d is lacking information and explanations: how are the phonon intensity plots generated from raw data? What do the red and grey shaded areas represent? Do the phonon intensity profiles correspond to OP or IP, or both?

We have revised Fig. 5 extensively to make the information clearer. The phonon intensity profiles are vertical slices of phonon spectral images in Fig. 2, with additional data for the large d interface. The red and grey shaded regions represent the location of 1uc FeSe for the large and small d interfaces. Only OP EELS results are shown because of the relevance to the EPC.

11. Caption of Figure 5d: "The large and small dFe-Ti are indicated by the red and black arrows." Where are the arrows?

The arrows and text are added in the revised figure (now Fig. 4).

12. Page 10, referring to figure 5c the authors mention that "In the 1uc FeSe/STO sample annealed at 500°C, we mostly observe the double-TiOx interface between 1uc FeSe and STO and thus a small dFe-Ti, as shown in Fig. 5(b)". This claim is difficult to connect with Figure 5c, where two data points are represented for the 1uc sample annealed at 500C, one at small d(Fe-Ti) and one at large d(Fe-Ti).

More generally, in Figure 5c, there are two data points for each 1 uc sample, one at low dFe-Ti (grey area) and one at high dFe-Ti (red area). Does it mean that for each 1uc sample and processing conditions, two areas have been investigated and yield two opposite results? More clarity would be expected from this plot and the description of these results.

This is indeed an unexpected observation. We investigated multiple regions for each 1uc samples by preparing multiple FIB specimens. We indeed observe both large and small dFe-Ti at local regions from each 1uc sample. These results do not contradict

each other but rather suggest spatial inhomogeneity in the d spacing within the samples. We have revised the discussion for this figure to improve clarity.

In Page 7, “After HAADF imaging of a series of FeSe/STO samples, we find that the local interface structure is characterized by either a large $d_{\text{Fe-Ti}}$ of over 6 Å, or a small $d_{\text{Fe-Ti}}$ of nearly 4 Å, as summarized in Fig. 4(c). In thick (10 uc) FeSe films grown on STO substrate, we observe exclusively large $d_{\text{Fe-Ti}}$ interfaces. In each of the annealed 1uc FeSe/STO samples, we observe coexistence of the small and large $d_{\text{Fe-Ti}}$ interfaces. The small $d_{\text{Fe-Ti}}$, double-TiO_x structure is the predominate type of interface for the 500°C annealed 1uc FeSe/STO sample, as shown in Fig. 1(b) and 4(b). Whereas in the as-grown and lower temperature (460°C) annealed 1uc FeSe/STO samples, the large $d_{\text{Fe-Ti}}$, Se-TiO_x type interface was observed more frequently.”

Miscellaneous comments

- It appears that the same overlapping orange spectrum is plotted in figure 2d,e,f. However, in the caption, it is indicated that “The overlapping orange spectra are from SrO or TiO₂ layers in STO”. I understand that one would expect an overlapping SrO layer spectrum in fig 2d, and a TiO₂ layer spectrum in fig 2e, f. This does not appear to be the case in the current figure.

Thank you for pointing this out. The TiO₂ layer is now compared in Fig. 1. The vibrational EELS of different atomic columns of STO is shown in Extended data figure x. EEL spectra of SrO and TiO₂ layers in STO are shown in supporting information Fig. Sx

- In figure 2g,h,i, the black curves look like filtered data plotted atop the experimental data points. Please provide more information.

The black lines are spline fit of the EELS data points. The details are added in the Method section.

- For ease of visualization, indicate the position of “Ti” and “O” columns of layers 0, 1, in some maps of figure 3c-f.

Thank you for the suggestion. We added arrows to indicate the in the Figure above.

- The connection between the phonon bandstructure and the Eliashberg spectral function would be easier to visualize if the energy scales were matching in figures 5b,c.

The phonon bands and Eliashberg spectral function have been revised to match the vertical scales.

- Extended data Fig. 7: the EELS spectra have been acquired on axis, as mentioned in the main text. However, this could also be indicated directly in the figure or caption.

The EELS spectra in extended data Fig. 7 are dark-field EELS. This has been clarified in the main text.

- Page 5: “Also shown in Fig. 2(b) is the DF-EEL spectrum of a bulk FeSe sample (blue curve)”. The color is inconsistent, replace “blue” by “orange”.

We have made sure that the colors are consistent in the revised manuscript.

- Typos and grammar could be improved in the manuscript, and in SI. Here are some examples:

Manuscript:

Page 6: “by integrated within” -> by integration within ; “Depends on the” -> Depending on the

Caption of Figure 3: (b) and (c) -> (c) and (d) ; (d) and (e) -> (e) and (f)

Page 7: “at as such” -> at such

Page 8: “these band” -> bands ; “Fig. Sx2”

SI, page 1: “is consist”, “these... consists”,

SI, Page 2 “can be substantially influence” -> influenced

SI, caption figure S3: “Fromation” -> Formation

Thank you for the helpful suggestions. These has been corrected.

- The caption of Figure S3 does not match the embedded legend. Also, which side of the figure is Ti-rich / O-rich?

This has also been corrected. Thank you for the careful review.

Referee #2 (Remarks to the Author):

The paper presents a highly ambitious measurement of atomically resolved phonons at an interface in order to tie into microscopic models of superconductivity enhancement at the FeSe/STO interface. The data appears to be cutting edge and begs a clearer discussion of the data. The plots in Figs. 2 and extended data of Fig. 5, especially, deserve a more careful discussion.

Thank you for the positive comments. We have substantially revised the manuscript to further improve the statistics of EELS results and the clarity of data interpretation. Please see detailed answers below.

1. What are the black solid lines in Fig. 2? Please define.

The black lines are spline fit of the EELS data points. The details are added in the Method section.

2. The peaks in the black solid lines of Fig. 2 e through h aren't all readily apparent in the data, especially for Figure 1i on Ti-O layer 0. Please discuss discrepancy.

The column resolved EEL spectra are integrated from very local regions around the atomic columns labeled. This sets a stringent limit on the frames of data that can be summed, therefore a relatively high noise. In the revised manuscript, we have performed additional EELS experiments to improve the statistics of the EEL spectra. With the improved EELS data, the energy loss peaks in the column resolved EELS (new Fig. 2 (c, d, e)) have become more apparent. Ti-O layer 0 has lower symmetry than other layers (i.e., TiO_2 layers in STO), thus show more complex phonon bands. The low atomic density of this layer also tends to lead to relatively weak phonon loss peaks.

3. The PhDOS from DFT does not clearly match the data, there is a minimum in the DOS at ~70 meV where the data shows a clear peak. Some of the discrepancy between theoretical DOS and measured may be explained by the different structures presented in the supplement. Can the authors get an image along the [110] zone axis of the STO to locate the position of the Ti and oxygen and distinguish between different structures?

Thanks for pointing out the discrepancy. With additional calculations considering more possibility for the interface structure, we find that the agreement between experiments and simulations can be improved by changing the Ti/O ratio in the Ti-O layer 0. The newly calculated PhDOS of a $\text{Ti}_{1.75}\text{O}_2$ model (supporting information Fig. S4) shows PhDOS near 74 meV, which provides better agreement with EELS data.

We have also performed new experiments on thinner TEM specimens to reduce multiple scattering, which can be another source of the discrepancies. The new EELS of Ti-O layer 0 (Fig. 1(g)) from thinner sample shows reduced intensity at 100 meV and enhanced intensity near 85 meV, which matches better with the PhDOS.

To get a better understanding of the interface structure, we have also performed more detailed HAADF image simulations (revised Fig.S2) and STEM imaging experiment along the [110] axis of STO. The result suggests that while the RT13 reconstruction model may likely be able to describe the interface structure. However, the large unit cell of this model prevents calculations of electron-phonon coupling or even phonon bands.

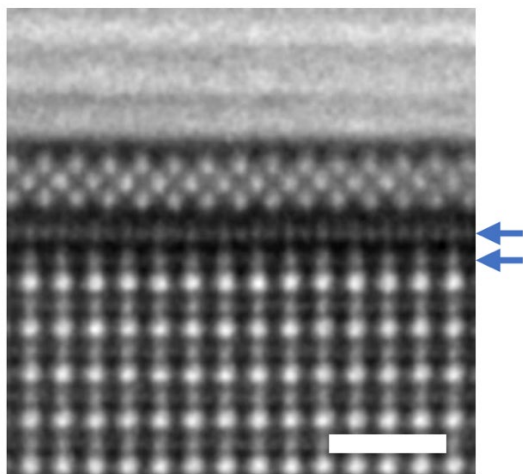


Figure R1. ADF-STEM image of 1uc FeSe/STO along the [110] axis of STO. The double TiO_x layers are indicated by blue arrows. Scalebar: 1nm.

With these new results, we have revised the manuscript to address the discrepancy. The vibrational spectra of Ti-O layer 1 and 0 and now compared with TiO_2 layers in STO (Revised Fig. 1 (f, g)). We highlight additional vibrational modes at this region by taking the difference between the spectra. The PhDOS are provided in supporting information Fig. S4 to show the PhDOS of both Ti_2O_2 and $\text{Ti}_{1.75}\text{O}_2$ models.

4. The term "interface phonon" should be defined earlier in the manuscript.

We now provide the definition of interface phonon when as soon as we start the discussion about the vibrational spectra in Fig. 1.

Page 3 "From the difference between the two EEL spectra, we identify additional vibrational modes of the interface (interface phonons) near 53 meV and between 75-88 meV (red shade)."

5. Are the statistics good enough to make definitive claims, such as "...show strong vibrational anisotropy" (bottom of page 5) and "a gradual softening of the O phonons" (bottom of page 8)? Is one supposed to see a shift in the peak position or height? What are the heights? How are they fit? The statistics on the momentum resolved data (especially for the 0 layer) are probably not good enough to make general statements on peak position. For layer 0, the only well defined peak is the Ti-OP peak in Fig. 2i. It would help the reader to not point out peaks that aren't clearly apparent.

The refined EELS data have better statistics, ensuring that the main spectral features are observable in both the data (scatter points) and the spline fit (black lines). The vibrational anisotropy is also improved. To clarify the discussion on column resolved EELS, we have rearranged spectra and combined them with the phonon maps in Fig. 2. The discussions have been revised accordingly. We have also removed the minor issue regarding the softening of O phonon in the revised manuscript.

6. Figure 3 shows good promise and is a nice figure but can be improved. Can the authors show theory derived images to highlight what the reader should be looking for in the noisy data?

Thank you for the positive comments and great question. We added simulated phonon maps in the revised manuscript to strengthen the interpretation of the vibrational EELS data. The FRFPMS simulation with DFT data as input takes into account both the PhDOS and direction+magnitude in q-transfer. The simulated phonon spectral images are now directly compared with experimental ones in the new Fig. 2 (Fig. 3 in previous version).

We have also performed the EELS experiments with further improved conditions to acquire data with better statistics. The stronger interface phonon intensity and vibrational anisotropy at the interface are now clearer.

Is the false color image in Fig. 3 proportional to I as suggested by the plot legend? If so, how is IxE^2 calculated? Is it from a fit, or is it the raw intensity?

In this context, I represents the raw EELS intensity and E is the energy in EELS. The product of IxE^2 provides a method for removing the elastic background without introducing artifacts from curve fitting, and has been used in other vibrational EELS studies. This part has been clarified in the Method section.

7. The raw spectra should also be shown in the extended figures to show the elastic peak in order to display the magnitude of the signal relative to the background.

The raw EELS data is shown in Extended data Figure 3.

8. It is known FeSe T_c varies dramatically with sample preparation. The T_c of the samples could be measured (for example with resistivity measurements as a function of temperature) to verify T_c . Here, it seems measurements of sc are limited to STM measurements, mostly at fixed temperature (5 K).

The main reason for choosing STM for superconductivity measurement is that it allows microscopic characterization, which can be better integrated with our STEM-EELS study. Due to the extreme sensitivity of the 1uc FeSe/STO surface, the gap opening temperature (T_g) measured by in-situ ARPES and STS has been used to estimate the T_c (Nat. Mater. 12, 605–610 (2013), Nat. Mater. 12, 634–640 (2013), Nat. Commun. 12, 2840 (2021)). Furthermore, the variation of T_c on annealing condition for 1uc FeSe/STO has been well documented in literature. We therefore adopted the same method and

observed a similar increase in superconducting gap with annealing temperature. Given that the ratio between superconducting gap and T_c for 1uc FeSe/STO has also been studied, here we use STM/STS as our superconductivity characterization. The largest gap observed here is 18 meV, among the largest reported for 1uc FeSe/STO. It should be acknowledged that transport measurement on the same samples would be ideal. However, the air exposure during ex-situ experiments would make the EELS study even more challenging, if not impossible.

9. Figure 4b shows many bands. How are the highlighted ones chosen to be the relevant ones?

Thank you for raising this important question. Figure 4b shows the phonon dispersions of the FeSe/Ti₂O₂/STO (3uc, AFD phase) model, which presents complex phonon dispersions due to the large number of atoms in the unit cell. We find that the interface phonon energies depend on the thickness of STO layers in the ultrathin regime, which is why we include these complex dispersions in the main text for comparison of peak positions with experiment EELS data.

To analyze their contributions to superconductivity, it is important to directly calculate the EPC of each phonon band. Unfortunately, it is still formidable to directly calculate the EPC through DFT for large unit cells as FeSe/Ti₂O₂/STO (3uc). Therefore, we used the smaller FeSe/Ti₂O₂/STO (1uc) model for this purpose and now discuss results from this model in the main text (now Fig. 3). Nevertheless, we find large similarities between the vibration motions of the interface phonon modes and the electronic band structures of these two models. Therefore, we highlighted the phonon modes that may have significant contribution to EPC according to the atomic displacement patterns in the FeSe/Ti₂O₂/STO (3uc) model. There are indeed many other phonon modes at nearby energy to the highlighted ones. However, their atomic displacement patterns are quite different. The other phonon modes mainly involve in-plane vibrations. More detailed analysis of the phonon modes is provided in Fig. 3 and the new extended data Fig. 5.

This manuscript promises a high impact, especially if revisions of the discussion of the EELS data are made to improve the clarity of the manuscript.

Again, we thank the referee for the positive assessment and insightful suggestions. We hope our revisions and explanations are satisfactory.

Referee #3:

Using high-resolution electron energy-loss spectroscopy (EELS), the authors present a study of the phonon modes near the interface of monolayer FeSe on SrTiO₃. Surprisingly, they find the existence of an additional TiO₂ layer with new out-of-plane optical phonon modes in the energy range 75-99 meV. Based on DFT calculations it is argued that these modes have a strong electron-phonon coupling and can thus be related to the enhanced T_c in FeSe/SrTiO₃.

I think this work is highly original, as it uses state-of-the-art experimental methods to address a timely issue, namely the interfacial phonon couplings in FeSe/SrTiO₃. In general the paper is well written, provides a lot of interesting insights. There are minor English mistakes here and there. Their conclusions are robust and warrant further investigations in possible phonon modes at the interface of FeSe and STO.

So while I would like to see this paper published as soon as possible, I do have some questions that need to be answered, mainly with regard to the theoretical parts of the paper. The questions below need to be answered and the manuscript should be changed accordingly:

We thank this referee for careful assessment of our work and the positive remarks.

1. What is T_c? The authors report the gap in the dI/dV at 5 K but I cannot find any data on the actual T_c of their samples.

In our study, we choose STM/STS for superconductivity characterization, as it allows for microscopic characterization, which can be better integrated with our STEM-EELS experiments. The main reason for this is that the gap opening temperature (T_g) measured by in-situ ARPES and STS can be used to estimate the T_c (Nat. Mater. 12, 605–610 (2013), Nat. Mater. 12, 634–640 (2013), Nat. Commun. 12, 2840 (2021)). Furthermore, the variation of T_c on annealing condition for 1uc FeSe/STO has been well documented in literature. We therefore adopted the same method and observed a similar increase in superconducting gap with annealing temperature. Given that the ratio between superconducting gap and T_c for 1uc FeSe/STO has also been studied, here we use STM/STS as our superconductivity characterization. The largest gap observed here is 18 meV, among the largest reported for 1uc FeSe/STO. It should be acknowledged that transport measurement on the same samples would be ideal. However, the air exposure during ex-situ experiments would make the EELS study even more challenging, if not impossible.

2. About the DFT results: The extra TiO₂ layer "0" causes a lot of metallic Ti bands to cross the Fermi level (eg. Extended Data Figure 4). This is in contradiction to various ARPES results (for example Ref 2 or 19) that clearly see only one FeSe band crossing the Fermi level. How do the authors explain this?

We believe that the difference is mainly because our DFT calculation considered a model with relatively high symmetry in the Ti-O layer 0. In the revised manuscript, we have further considered a Ti_{1.75}O₂ model, in which the Ti vacancy lowers the symmetry. The result (shown in Fig. S7) shows diminished intensity of the Ti-contributed bands at the center of the Brillouin zone of the normal cell. And since ARPES intensity is proportional to the unfolded spectral function, the FeSe/TiO_x/STO with lower symmetry at the additional Ti-O layer should align with previous ARPES measurements that only observed electron pockets at the M points.

Another reason for the discrepancy could be the limited probing depth of ARPES. The Ti-O layer 0 is beneath the FeSe film. The exponential decay of photoelectron intensity can make the detection more difficult than the Fe d bands in ARPES measurements. In core-loss EELS of the cross-section samples, we indeed detected a small portion of Ti³⁺ at Ti-O layer 0. Therefore, the presence of Ti 3d bands crossing the E_F is not entirely ruled out and may warrant further investigation with sufficient probing depth and sensitivity.

3. The DFT results of Fig. 4 are presented in an unclear way. The (c) panel is very difficult to read because the dashed line is, well, not a continuous line. Moreover, the vertical axis ticks of panel (c) do not align with the vertical axis ticks of panel (b), making it impossible to read at which energy there are peaks in α^2F . Finally, the peaks in α^2F seem to happen at 18, 65 and 79 meV, which is quite different from what the main text and the panel (a,b) are trying to convey.

We thank the referee for raising this issue. Previous Fig. 4 displayed the phonon dispersions of the FeSe/Ti₂O₂/STO (3uc, AFD phase) model, which exhibit complex phonon bands due to the large number of atoms in the unit cell. We observed that the interface phonon energies depend on the thickness of STO layers in the ultrathin regime, which is why we included these complex dispersions in the main text for comparison of peak positions with experiment EELS data. Unfortunately, it remains a formidable challenge to directly calculate the EPC through DFT for large unit cells as FeSe/Ti₂O₂/STO (3uc), so we used results of α^2F from a smaller FeSe/Ti₂O₂/STO (1uc) model. This might have caused confusion due to the energy mismatch, as pointed out by the referee.

In the revised version, we have replotted the DFT results in Fig. 4 (now Fig. 3) with phonon dispersions and EPC for the same FeSe/Ti₂O₂/STO (1uc) model. We removed

the dashed line and aligned the vertical axes of the phonon dispersion and the α^2F plots. The peaks in α^2F are now clearly indicated at 65.7 and 78.0 meV, aligning with the interface phonon bands that exhibit large EPC, as indicated by their large linewidths. The 18 meV peak previously shown was due to an imaginary mode and have now disappeared in our new calculations, which utilize a larger supercell and denser k- and q-meshed. The lattice structures and the atomic displacements for these interface modes are shown in new extended data Fig. 5.

Phonon dispersions calculated for the FeSe/Ti₂O₂/STO (3uc, in the antiferrodistorive phase) are shown in the new Fig. S8. The energies of two interface modes in FeSe/Ti₂O₂/STO (3uc) in Fig. S8 are 75.2 and 91.2 meV, which are about 10 meV higher compared to those in FeSe/Ti₂O₂/STO (1uc). These results highlight the sensitivity of phonon energies to the thickness in the ultrathin regime due to the changes in boundary conditions. Nevertheless, there is a notable similarity in the atomic displacements of the interface phonon modes between the two models, demonstrating consistency in their vibrational characteristics and EPC.

4. The text "The comparison of phonon dispersions with and without the AFD phase transition is discussed in Fig. Sx2." refers to a figure Sx2 but I can't find which figure that is supposed to refer to.

Thanks for pointing this out. We have corrected this in the revised manuscript. The comparison of phonon dispersions with and without the AFD phase transition is now shown in the new Supporting Information Section 3.2 Antiferrodistorive phase of SrTiO₃.

5. In the introduction, a difference is made between FK phonons and "oxygen optical phonons". Since in either case we are dealing with optical phonons on the surface of STO, I don't really see why the authors emphasize there is a difference here.

Indeed, both terms have been used in literature when referring to the phonons are possibly involved in the EPC of 1uc FeSe/STO. However, we believe there are important differences between them. The F-K phonons are normal modes resulting from the interaction between the electromagnetic field and the lattice vibrations at the surface of ionic crystals. HREELS can probe F-K phonons due to the extended field of these phonons. For this reason, it was believed that the F-K modes of STO might mediate Cooper pairing in FeSe over a large distance. In contrast, the "oxygen optical phonons" stem from motion of oxygen octahedra and affect the electronic structure via local structural deformation. They were first linked to the Tc enhancement in 1uc FeSe/STO by Ref. [2]. These modes were believed to originate the highest energy longitudinal optical phonons of STO (the LO₃), which involve only the vibration of oxygen atoms. In our studies, we found that the EPC is more related to the atomic vibration as the

superconducting gap is sensitive to $d_{\text{Fe-Ti}}$. Therefore, we want to distinguish between the F-K and oxygen optical phonons as they represent different mechanisms for superconductivity enhancement. The original work on F-K phonon has been referenced in the introduction to enhance clarity.

6. The key theory of why phonon coupling can enhance T_c in FeSe/STO is based on the concept of "forward scattering", ie, only small-momentum transfer between FeSe electrons and out-of-plane optical phonons in the substrate. The whole paper does not mention this concept, even though their findings are consistent with this theory. For example, in Fig. S5 in the supplementary they clearly show that the electron-phonon coupling is strongest at Γ (ie. small momentum). Early theories [for example Rademaker et al New J. Phys. 18 022001 (2016)] were based on the idea that the relevant phonon coupling is from an oxygen optical mode in the STO, not (as this paper suggests) from the new optical modes from the interfacial TiO_2 layer. This paper really needs to discuss the forward scattering mechanism; in particular: how the theory of forward scattering changes considering the 'new' optical modes discovered here. Furthermore, estimates of α^2F are not enough to properly take into account the enhancement due to forward scattering, and therefore a momentum-dependent Eliashberg calculation might be warranted.

We agree with the referee that the T_c enhancement in FeSe/STO results from forward scattering mechanism. Our calculations indeed indicate the electron-phonon coupling is strongly peaked around $q=0$, as now shown in the revised Fig. 3. However, our work highlights that the key factor for EPC is the overlap of phonon and electron wavefunctions in the interface region. In the revised version, we emphasize this new finding and also mention the forward scattering nature of EPC in the context of the phonon modes discovered here.

Reviewer Reports on the First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

I thank the authors for providing additional EELS data of higher quality than in the original version of the manuscript, and high-end simulations that appear very appropriate to support this type of experiments. The agreement between experiments and simulations is much improved, which supports the interpretation of the results. The strategic selection of results in the main text makes the manuscript easier to follow and therefore more impactful.

I appreciate that the authors considered my comments carefully and their effort to modify the manuscript extensively. I acknowledge that most of my comments and concerns have been addressed appropriately. Nevertheless, given that the manuscript has been almost entirely rewritten and all figures have been modified, I still have some minor comments on this revised version, which I hope the authors would be willing to address.

I recommend this manuscript for publication.

- Page 3, the statement “indicating the presence of Ti^{3+} at the interface” sounds very certain, however without proper support for this claim. It could be carefully reworded with “suggesting” instead of “indicating”.

- In the FRFPMS maps in Fig. 2, some atomic columns in a (horizontal) atomic plane appear with different intensity although one would expect them to be similar. This is the case for layer 0 (and FeSe) in the OP 14 meV map, and for layer 0 (and some O-columns in STO) in the 45-70 meV map. This apparent discrepancy should be fixed if possible, or at least its origin mentioned.

- Most energy values/ranges match very well for experiments and simulations in Fig. 2, except for the IP map at 77 meV (experiments) and 82 meV (simulation) despite an overlapping range of errors. What is the reason? The same remark holds for most energy values of the simulations in the Extended data Fig. 4 (e, f).

- In Fig. 2, the lookup table should be labelled I^*E^2 , assuming that the maps are indeed displayed similarly to Fig. 3 in the original manuscript.

- The manuscript reads well, but typos remain (first principal calculations -> first principles... ; infromation -> information, etc.).

Referee #2 (Remarks to the Author):

I looked at the revised manuscript. The authors have made a good faith effort to address the comments. Figure R1 is critical for constraining the interface structure and must be included and described in the manuscript. It's important because the structure will play a determinative role in the vibration spectrum. Otherwise, the manuscript is acceptable for publication.

Referee #3 (Remarks to the Author):

In the revised manuscript the data is presented in a much clearer way, and I think the authors for that. The updated DFT results are now clearly consistent with the experimental results. Following up on my earlier enthusiasm about these results, I can recommend publication.

Author Rebuttals to First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

I thank the authors for providing additional EELS data of higher quality than in the original version of the manuscript, and high-end simulations that appear very appropriate to support this type of experiments. The agreement between experiments and simulations is much improved, which supports the interpretation of the results. The strategic selection of results in the main text makes the manuscript easier to follow and therefore more impactful.

I appreciate that the authors considered my comments carefully and their effort to modify the manuscript extensively. I acknowledge that most of my comments and concerns have been addressed appropriately. Nevertheless, given that the manuscript has been almost entirely rewritten and all figures have been modified, I still have some minor comments on this revised version, which I hope the authors would be willing to address.

I recommend this manuscript for publication.

Thank you for the positive assessment and additional comments.

- Page 3, the statement “indicating the presence of Ti^{3+} at the interface” sounds very certain, however without proper support for this claim. It could be carefully reworded with “suggesting” instead of “indicating”.

We agree that the structure and bonding of the topmost TiO_x layer deserves more in-depth study. This has been changed to “suggesting the presence of Ti^{3+} at the interface”.

- In the FRFPMS maps in Fig.2, some atomic columns in a (horizontal) atomic plane appear with different intensity although one would expect them to be similar. This is the case for layer 0 (and FeSe) in the OP 14 meV map, and for layer 0 (and some O-columns in STO) in the 45-70 meV map. This apparent discrepancy should be fixed if possible, or at least its origin mentioned.

Thank you for the careful review. The intensity variation along the interface is indeed not expected, at least for the STO substrate region. This minor artifact may originate from the

2x2 unit cell used in our DFT simulation. In addition, the number of snapshots of displacement patterns in our FRFPMS simulation may also contribute to this. To improve clarity, we have fixed the minor intensity variation in the 100meV OP phonon image with a symmetrization procedure in the revised manuscript.

- Most energy values/ranges match very well for experiments and simulations in Fig. 2, except for the IP map at 77 meV (experiments) and 82 meV (simulation) despite an overlapping range of errors. What is the reason? The same remark holds for most energy values of the simulations in the Extended data Fig. 4 (e, f).

When comparing experiments with simulations, our criterion prioritizes the matching the atomic displacement pattern over energy. The reason for this is that the uncertainty in the structure of Ti-O layer 0 may influence the energy of some phonon bands, resulting in mismatch between EELS and simulations if we choose the exact same energy. While most energies show good agreement, we find that the simulated image at 82 meV matches best with EELS at 77 meV, likely due to the energy mismatch between some phonon bands.

- In Fig. 2, the lookup table should be labelled $I \cdot E^2$, assuming that the maps are indeed displayed similarly to Fig. 3 in the original manuscript.

The intensity shown is $I \cdot E^2$, which is now added for the color bar label. We also find that the exponent of energy has little influence on the contrast of individual image when the energy window is sufficiently small. These details are discussed in the Methods section.

- The manuscript reads well, but typos remain (first principal calculations -> first principles... ; infomation -> information, etc.).

Thank you for the careful review. These have been corrected.

Referee #2 (Remarks to the Author):

I looked at the revised manuscript. The authors have made a good faith effort to address the comments. Figure R1 is critical for constraining the interface structure and must be included and described in the manuscript. It's important because the structure will play a determinative role in the vibration spectrum. Otherwise, the manuscript is acceptable for publication.

Thank you for the positive assessment. We agree about the importance of the interface structure. Figure R1 is now added in the Supplementary information, together with the HAADF image simulation to discuss about the interface structure. Although the combined

ADF-STEM imaging from [100] and [110] projections may not be enough to resolve the structure completely, it provides valuable information on the interface structure that future work may resolve with better confidence.

Referee #3 (Remarks to the Author):

In the revised manuscript the data is presented in a much clearer way, and I think the authors for that. The updated DFT results are now clearly consistent with the experimental results. Following up on my earlier enthusiasm about these results, I can recommend publication.

Thank you for the positive assessment.