

## ARTICLE OPEN



# Common microscopic origin of the phase transitions in $\text{Ta}_2\text{NiS}_5$ and the excitonic insulator candidate $\text{Ta}_2\text{NiSe}_5$

Lukas Windgätter<sup>1</sup>, Malte Rösner<sup>2</sup>, Giacomo Mazza<sup>3</sup>, Hannes Hübener<sup>1</sup>, Antoine Georges<sup>3,4,5,6</sup>, Andrew J. Millis<sup>6,7</sup>, Simone Latini<sup>1</sup> and Angel Rubio<sup>1,6,8</sup>

The structural phase transition in  $\text{Ta}_2\text{NiSe}_5$  has been envisioned as driven by the formation of an excitonic insulating phase. However, the role of structural and electronic instabilities on crystal symmetry breaking has yet to be disentangled. Meanwhile, the phase transition in its complementary material  $\text{Ta}_2\text{NiS}_5$  does not show any experimental hints of an excitonic insulating phase. We present a microscopic investigation of the electronic and phononic effects involved in the structural phase transition in  $\text{Ta}_2\text{NiSe}_5$  and  $\text{Ta}_2\text{NiS}_5$  using extensive first-principles calculations. In both materials the crystal symmetries are broken by phonon instabilities, which in turn lead to changes in the electronic bandstructure also observed in the experiment. A total energy landscape analysis shows no tendency towards a purely electronic instability and we find that a sizeable lattice distortion is needed to open a bandgap. We conclude that an excitonic instability is not needed to explain the phase transition in both  $\text{Ta}_2\text{NiSe}_5$  and  $\text{Ta}_2\text{NiS}_5$ .

*npj Computational Materials* (2021)7:210; <https://doi.org/10.1038/s41524-021-00675-6>

## INTRODUCTION

The excitonic insulator phase has been theoretically proposed in the 1960s<sup>1–6</sup> and is predicted to appear in semiconductors (or semimetals) with excitonic binding energies larger than the bandgap (or bandoverlap) of their conventional groundstate. Under this condition, the groundstate becomes unstable against the spontaneous formation of bound electron hole pairs, i.e. excitons. Depending on the conventional phase being semiconducting or semimetallic, the transition is described by a Bose–Einstein condensation (BEC)- or a Bardeen–Cooper–Schrieffer (BCS)-like mechanism, respectively<sup>2,4,5</sup>.

A major challenge in the experimental detection of an excitonic insulator state in a crystal is that the excitonic transition is coupled to other degrees of freedom, such as lattice distortions, which make an unambiguous detection difficult. This is the reason why several materials such as  $1\text{T-TiSe}_2$ <sup>7,8</sup> or  $\text{TmSe}_{0.45}\text{Te}_{0.55}$ <sup>9</sup> have not been unambiguously confirmed to host an excitonic insulating (EI) groundstate. These are indirect bandgap or semimetallic materials that exhibit finite momentum ordering and in the case of  $1\text{T-TiSe}_2$ , for example, the existence of a charge density wave masks the possible presence of an EI phase.

Very recently, also two-dimensional materials have shown promises to host an excitonic groundstate. Experimentally, strong evidence for the existence of EI phases has been provided both for transition metal dichalcogenide bilayers<sup>10</sup> and monolayers, such as  $\text{WTe}_2$  and  $\text{MoS}_2$ <sup>11,12</sup>. An extensive theoretical study has identified several candidate material combinations for hetero-bilayers that could host an excitonic insulator by analysing the electronic properties of a wide range of materials<sup>13</sup>. Further low-dimensional excitonic insulator candidates are carbon nanotubes<sup>14</sup>, Sb nanoflakes<sup>15</sup>, double bilayer graphene<sup>16,17</sup> and topological systems<sup>18</sup>, such as  $\text{InAs/GaSb}$ <sup>19</sup>. In the bulk phase,

$\text{MoS}_2$  has recently been revealed as another candidate material since its bandgap can be tuned via pressure, yielding excitonic binding energies larger than the quasi-particle gap, which would allow for a transition into a possibly EI groundstate<sup>20</sup>.

We investigate bulk  $\text{Ta}_2\text{NiSe}_5$  (TNSe), which is considered to be a very promising candidate to host an excitonic condensate<sup>21</sup>, with several experimental evidence suggesting an EI groundstate: angle-resolved photo-electron spectroscopy (ARPES) measurements have shown a characteristic band flattening near the  $\Gamma$ -point upon cooling below the critical temperature<sup>22–25</sup>, a dome-shaped bandgap–temperature phase diagram, similar to the theoretically predicted one, has been found<sup>26</sup>, and the opening of a gap has been measured in scanning tunnel spectroscopy (STS)<sup>27</sup>, optics<sup>26</sup> and ARPES<sup>24,28</sup> experiments below the critical temperature<sup>27</sup>. The bandgap in the low-temperature phase is reported to be between 160 meV in optics<sup>26</sup> and 300 meV in STS measurements<sup>27</sup>. The current literature is controversially reporting the high-temperature phase of TNSe to be either a direct bandgap semiconductor<sup>21,29</sup> or a small bandoverlap semimetal<sup>24</sup>. It is, however, agreed that there exists a structural phase transition, and possibly a transition into an excitonic groundstate, at the critical temperature  $T_c = 326\text{ K}$ <sup>26,30–32</sup>. The structural transition has been reported to be from a high-temperature orthorhombic to a low-temperature monoclinic phase<sup>30–32</sup>.

On the other hand  $\text{Ta}_2\text{NiS}_5$  (TNS), a material of the same structure of TNSe with Sulfur replacing selenium, does not exhibit the same excitonic signatures of TNSe, such as band flattening and bandgap opening. Indeed a structural transition at  $120\text{ K}$ <sup>33</sup> has been reported without the formation of flat bands or a metal to semiconductor transition<sup>34</sup>.

Due to the uncertainty on whether TNSe has a semimetallic or gapped high-temperature phase, from a theoretical perspective it is unclear which mechanism between BEC and BCS can

<sup>1</sup>Max Planck Institute for the Structure and Dynamics of Matter, Luruper Chaussee 149, 22761 Hamburg, Germany. <sup>2</sup>Radboud University, Institute for Molecules and Materials, Heijendaalseweg 135, 6525 AJ Nijmegen, Netherlands. <sup>3</sup>Department of Quantum Matter Physics, University of Geneva, Quai Ernest-Ansermet 24, 1211 Geneva, Switzerland. <sup>4</sup>Collège de France, 11 place Marcelin Berthelot, 75005 Paris, France. <sup>5</sup>Center for Computational Quantum Physics, Flatiron Institute, New York, NY 10010, USA. <sup>6</sup>CPHT, CNRS, Ecole Polytechnique, IP Paris, 91128 Palaiseau, France. <sup>7</sup>Department of Physics, Columbia University, New York, NY 10027, USA. <sup>8</sup>Nano-Bio Spectroscopy Group, Departamento de Física de Materiales, Universidad del País Vasco, 20018 San Sebastian, Spain. ✉email: lukas.windgaetter@mpsd.mpg.de; simone.latini@mpsd.mpg.de; angel.rubio@mpsd.mpg.de

be applied. In fact, experimentally there is no unambiguous evidence that excitons are involved in the phase transition at all. These facts indicate that, in relation to TNSe, the concept of excitonic insulator can be source of confusion, as it has been generally employed to refer to a phase showing a characteristic gap opening and band flattening<sup>22,23,25,27</sup> without direct observation of excitonic features and/or excitonic condensation.

A recent symmetry analysis of the possible electronic instabilities issuing from the electron–electron interactions has shown that the electronic phase transitions in the material should not be ascribed to a condensation phenomenon but rather to the electronic lowering of the discrete lattice symmetry<sup>35</sup>, which corresponds to the breaking of a continuous U(1) symmetry during the excitonic phase transition. This observation, together with other recent experimental and theoretical observations<sup>24,36</sup>, suggest a prominent role of structural instabilities in the system. In this work, we abandon the concept of excitonic condensation and show that the experimental features of TNSe and TNS mentioned above can be completely explained without explicitly taking the formation of excitons into account. We demonstrate how the electronic and structural properties of the two crystals can be rationalized in terms of electronic and intrinsic lattice instabilities that drive the system from the orthorhombic to monoclinic phase. The occurrence of such a lattice-driven instability agrees with a prior report of Alaska Subedi<sup>36</sup>.

We present a comprehensive *ab initio* study using density functional theory (DFT) for the transition from the orthorhombic to the monoclinic phase beyond standard generalized gradient approximation functionals. In section “Structure relaxation,” we start by investigating the structural stability of the system and show that at low temperature a monoclinic groundstate is energetically favoured to the orthorhombic one, which suggests the existence of a lattice instability. A similar analysis applied to TNS highlights an equivalent structural transition mechanism from an orthorhombic to a *C2/c* symmetric monoclinic cell.

In section “Electronic bandstructure,” we then investigate the electronic bandstructure of both the high- and low-temperature phase of TNSe and TNS using different exchange correlation functionals with increasing accuracy. For TNSe, we can establish the orthorhombic groundstate to be metallic and observe a sizeable gap opening as well as a band flattening with the phase transition. This shows that for this material the structural distortion is essential for the bandgap opening observed in optics and STS measurements and cannot be explained considering only the electronic degrees of freedom. In section “ $G_0W_0$  calculations,” we show that a similar finding results from our many-body perturbation theory-based calculations carried out at the  $G_0W_0$  level. Performing an equivalent DFT analysis for TNS shows that both the orthorhombic and monoclinic phases are gapped systems with parabolic electronic dispersions at  $\Gamma$ . This explains why a metal to semiconductor transition and band flattening, which have been characterized as a signature of an excitonic insulator in TNSe, has not been observed for TNS<sup>34</sup>. We then show that electronic heating enhances the valence band flattening and investigate in the following section “Origin of the structural instability” the origin of the instability observed in TNSe. We find no evidence for an electronically driven spontaneous symmetry breaking in the TNSe charge density upon infinitesimal lattice distortions. In the symmetry broken lattice structure, the total energy is, however, clearly lowered, which renders the electron–phonon interaction the main driving mechanism for the structural transition. In section “Phonon dispersion,” we discuss the phononic bandstructure and single out the phononic instabilities of the orthorhombic phase, the ones that could possibly drive the structural transition. A detailed discussion of the obtained phonon modes in relation with the most recent Raman measurements shows that we are able to reproduce all phonon peaks, except for the first two  $B_{2g}$  modes, which exhibit a strong

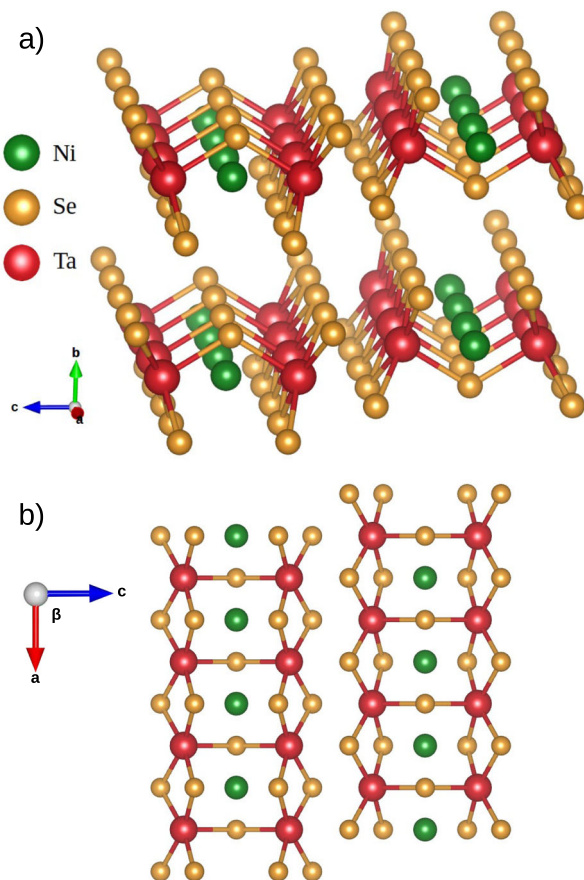
broadening, which hints to a strong coupling of these modes to other degrees of freedom. Finally, in section “Electron–phonon coupling,” we show that the electronic dispersion is strongly modulated by different phonon modes. In particular, it is expected that thermal occupation of the phonon mode responsible for the structural transition leads to a further bandgap opening.

## RESULTS

### Structure relaxation

TNSe is a layered material bound by van-der Waals forces. It consists of parallel Tantalum and Nickel chains and is thus referred to as a quasi-one-dimensional material (see Fig. 1)<sup>32</sup>. Experimentally, the high-temperature phase exhibits an orthorhombic structure, which is distorted to a monoclinic structure in the low-temperature phase. The distortion, however, is subtle with a change in the  $\beta$  angle (see Fig. 2b) of just  $0.53^\circ$ <sup>32</sup> to  $0.69^\circ$ <sup>30</sup>.

We have performed a first-principles structure relaxation of atomic coordinates, cell shape and cell volume for different functionals. In all cases, the relaxation resulted in a triclinic structure. The resulting lattice parameters are shown in Supplementary Tables 1 and 5: the two functionals vdW-optB88 and vdW-optPBE<sup>37,38</sup>, which include van der Waals corrections, result in relaxed structures that agree well with the experimentally measured values, while the Perdew–Burke–Ernzerhof (PBE) functionals overestimates the experimentally measured interlayer lattice parameter  $b$  by 10.5%<sup>30</sup> (see Supplementary Tables 1 and 5).



**Fig. 1**  $\text{Ta}_2\text{NiSe}_5$  structure. Layered structure of the orthorhombic phase of TNSe. The Nickel and Tantalum atoms form parallel chains along the  $a$ -axis. **a** shows a 3D view of TNSe. **b** shows the projection onto the  $a$ - $c$  plane. For TNS, the Se atoms are simply replaced with Sulfur atoms.

As expected, this shows the importance of including van der Waals forces when describing layered structures, such as TNSe.

The angles of the relaxed triclinic structure are  $\alpha = 90.005^\circ$ ,  $\beta = 90.644^\circ$  and  $\gamma = 89.948^\circ$ . As the two angles  $\alpha$  and  $\gamma$  are almost rectangular, the triclinic cell and the corresponding monoclinic cell are almost degenerate. Furthermore, we have checked that the small triclinic distortion does not modify the electronic properties compared to the exact monoclinic structure. Thus, in the following, we refer to the fully relaxed structure as monoclinic. The monoclinic angle  $\beta = 90.644^\circ$  agrees well with the experimentally measured values  $90.693^\circ$ <sup>30</sup> and  $90.53^\circ$ <sup>32</sup>.

To obtain the relaxed geometry for the high-temperature orthorhombic cell, we have performed the same relaxation using the vdW-optB88 functional enforcing the orthorhombic symmetry. The resulting lattice parameters are shown in Supplementary Tables 2 and 4. As the relaxed structure predicted by the vdW-optB88 functional has the best agreement with the experimentally measured values for both the monoclinic and orthorhombic phase, we have chosen it for all following structural calculations.

We have performed similar relaxations for the Sulfur compound: a full relaxation using the vdW-optB88 functional, again, results in a triclinic geometry that is almost degenerate with the monoclinic cell. The monoclinic angle is  $\beta = 90.50^\circ$ . For the orthorhombic phase we have performed a further relaxation with fixed symmetry. The lattice parameters for both structures are shown in Supplementary Table 3 and agree very well with the experimentally measured values from X-ray diffraction. This agrees with a recent Raman study by Ye et al., which reports an orthorhombic high-temperature phase and a structural transition to a monoclinic phase at  $T = 120\text{ K}$ <sup>33</sup>.

Therefore, we find that in both TNSe and TNS a monoclinic geometry is energetically favoured, which shows that a similar lattice instability occurs in both materials. In section “Phonon dispersion,” we extensively discuss the phononic properties of the two materials and identify the structural instability with an unstable  $B_{2g}$  phonon. Although the structural change between monoclinic and orthorhombic phase is small, the modifications for the electronic bandstructure with the phase transition is significant (see section “Electronic bandstructure”). This hints to a strong coupling between electronic and lattice degrees of freedom (see section “Electron–phonon coupling”).

## Electronic bandstructure

In this section, we present a complete and systematic DFT study of both TNSe and TNS using standard and more accurate functionals. While we find that the fine electronic properties are sensitive to the details of exchange and correlation, the underlying mechanism of bandgap opening due to the structural phase transition is independent of the functional choice for both TNSe and TNS. We investigate the bandstructure of the orthorhombic and monoclinic geometries using different exchange correlation functionals, which include varying degrees of exact exchange. These are ranked in chemical accuracy with the PBE functional<sup>39</sup> being the most inaccurate, but computationally least demanding, followed by the modified Becke–Johnson (mBJ) functional<sup>40–42</sup> and finally the range separated hybrid functional HSE03<sup>43</sup> being the most accurate one.

The results are shown in Fig. 2. Results for the HSE06 functional are reported in Supplementary Fig. 9 and the definition of the  $M$ – $\Gamma$ – $X$  path is given in Supplementary Note 1.

Figure 2a, b display the obtained bandstructure of TNSe: the high-temperature orthorhombic structure exhibits a metallic groundstate independently of the functional presented here, while the monoclinic structure is always semiconducting. The bandgap opening is sizeable for all the investigated functionals and its position and value changes with the amount of exchange and correlation included. With the PBE functional, we obtain an

indirect bandgap value of 40 meV, with the mBJ functional a direct bandgap of 101 meV and with the HSE03 functional a direct bandgap of 183 meV. Thus, the bandgap grows with an increasing accuracy of the exchange and correlation functional. From Fig. 2b, we observe that, in the monoclinic phase, TNSe has a rather flat dispersion around  $\Gamma$  due to the valence band maxima being off  $\Gamma$  for both the PBE and mBJ functional.

By increasing the exchange contribution included in the functional, e.g. by tuning the mixing parameter in the hybrid functional or the  $c$  parameter in the mBJ functional, it is possible to further increase the bandgap. We observed that even in the orthorhombic phase increasing the exchange contribution can lead to a gap opening (see Supplementary Note 4 and Supplementary Figs. 5 and 6), which is consistent with prior studies<sup>44</sup>. The bandgap opening resulting from HSE03 agrees well with the experimentally measured bandgap of 160 meV<sup>26</sup>. In STS measurements, a gap of 300 meV has been reported<sup>27</sup>. However, we notice that the structure in ref. <sup>27</sup> shows a monoclinic distortion with an angle  $\beta \approx 92.5^\circ$ , i.e. much larger than the angle reported in previous studies and the angle obtained in our calculation ( $\beta \approx 90.6^\circ$ ). Correcting for this difference, we find that both the optics and scanning tunneling microscopic results are in good agreement with our HSE03 results (see Supplementary Note 6 for more details).

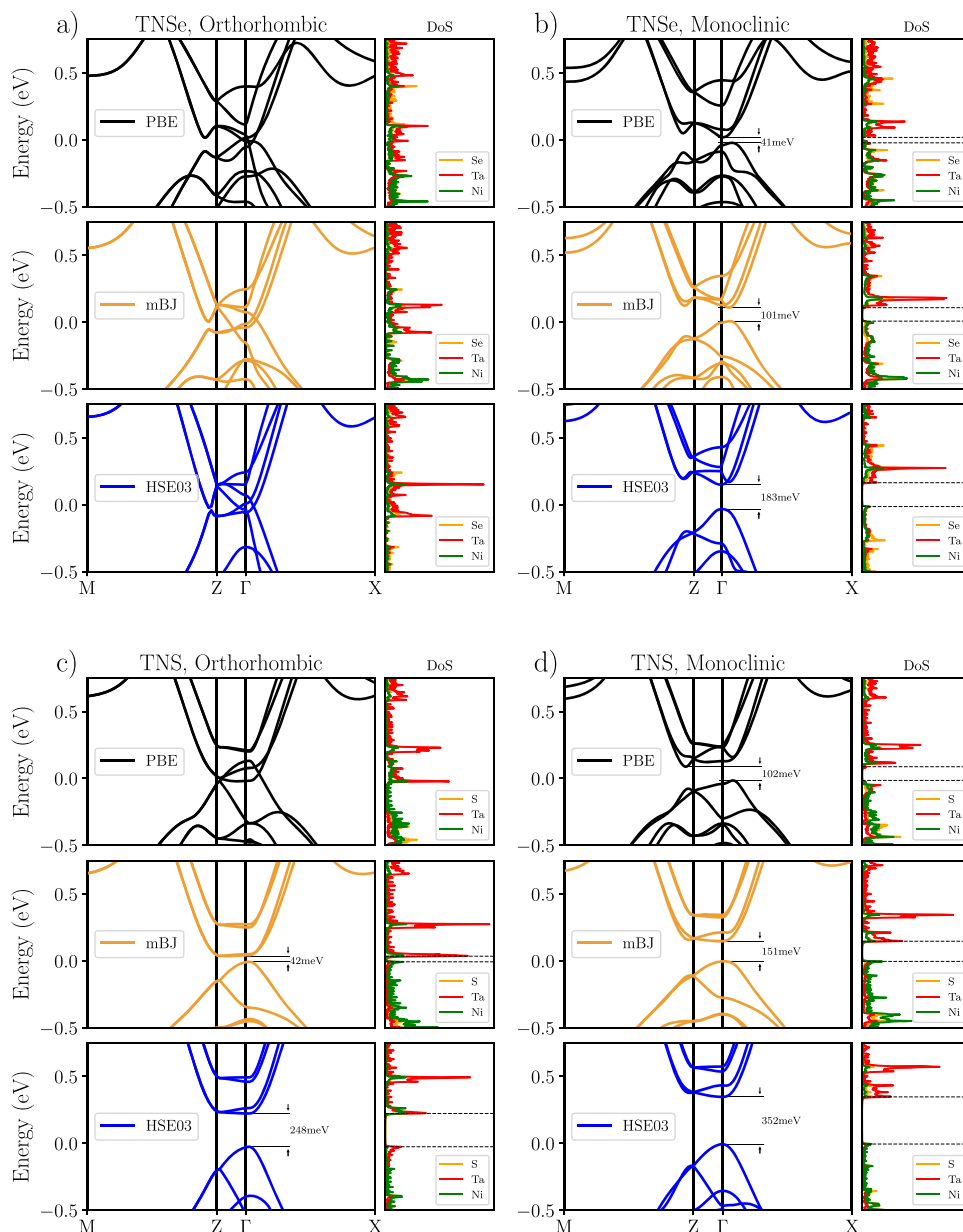
Performing similar calculations for TNS, Fig. 2c, d, we identify a major difference in the electronic behaviour of the Selenium and Sulfur compounds. We see that, using either mBJ or HSE03 hybrid functionals, the calculations predict a semiconducting groundstate even for the orthorhombic geometry. The structural transition to the monoclinic cell then leads to a further bandgap opening, with bandgap values being around twice the size of the TNSe ones. Therefore, in contrast to TNSe, no metal-to-insulator transition is observed. The bandgap is direct at the  $\Gamma$ -point with a parabolic dispersion towards the  $X$  direction. These findings agree well with the theoretical and experimental reports: ARPES as well as optical conductivity data show that TNS has a parabolic dispersion around the  $\Gamma$ -point<sup>34,44</sup> with a robust bandgap of 200 meV<sup>34,44,45</sup> and 250 meV in optics<sup>46</sup>.

We stress that, while both compounds have a similar lattice instability from orthorhombic to monoclinic geometry, in the case of TNSe the structural instability is accompanied by a metal-to-insulator transition, which is absent in the case of TNS. These observations match all known experimental trends for these two systems. In the case of the Selenium compound, the metal–insulator transition accompanied by the flattening of the valence band around  $\Gamma$  appears to be the intrinsic outcome of the structural instability of the system and not the distinctive signature of the excitonic insulator phase, as it has been interpreted so far. We support this statement in the next sections by explicitly accounting for the role of the electronic instability and comparing the excitonic binding energy with the bandgap value.

Given that states near the band gap have significant contributions from Ta, we have checked the validity of our results by also including spin–orbit coupling effects and reported the related band structure calculations in Supplementary Figs. 2–4, Supplementary Note 3 and Supplementary Table 6. We find that, besides a band splitting in the  $M$ – $Z$  direction and a minor renormalization of the bandgap, all our claims remain unchanged.

## $G_0W_0$ calculations

While DFT calculations generally provide a good starting point for discussing material properties, they often lack quantitative agreement with experimental results. It is known that one needs to add many-body corrections to obtain a quantitative agreement with the experimental bandgaps. The standard approach that provides accurate bandgaps is the GW method<sup>47–50</sup>. The latter is a

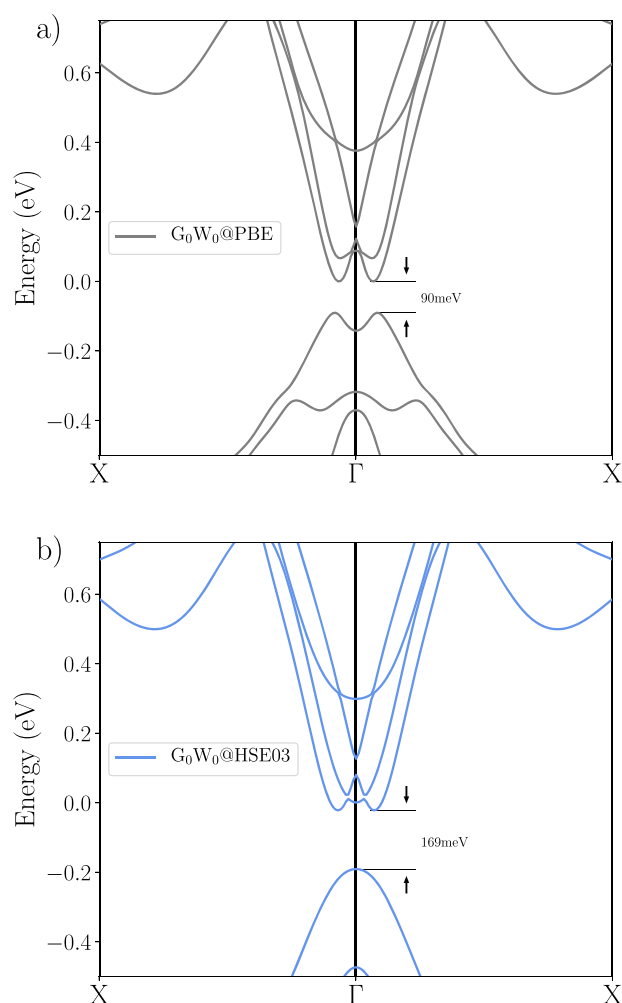


**Fig. 2 Bandstructure of  $\text{Ta}_2\text{NiSe}_5$  and  $\text{Ta}_2\text{NiS}_5$ .** Band structure of TNSe and TNS for the orthorhombic and monoclinic phase. For all the investigated exchange correlation functionals, we obtain a semimetallic groundstate for the high-temperature phase of TNSe. Only after the structural transition to the monoclinic geometry, a bandgap of between 40 and 183 meV opens. In TNS, already the orthorhombic phase exhibits a positive bandgap for the modified Becke–Johnson and HSE03 functional, which increases during the structural transition to the monoclinic phase. Therefore, the metal-to-insulator transition observed in TNSe is lacking for TNS. In all plots, the Fermi energy is set to 0 eV. The self-consistently calculated  $c$ -value for TNSe is  $c = 1.25$  and  $c = 1.26$  for TNS in both geometries.

many-body theory approach that consists of solving Hedin's equations iteratively<sup>51</sup> while neglecting vertex corrections. There exist various flavours of the GW approximation with the  $G_0W_0$  being the most commonly used. In  $G_0W_0$ , one approximates the starting wavefunctions to be the DFT ones and performs only one cycle of Hedin's equation to compute the self-energy and correct the electronic state energies. As such, the  $G_0W_0$  approximation relies on the DFT wavefunctions to provide a good description of the full many-body wavefunction and it is crucial to investigate the starting point dependence of the  $G_0W_0$  approximation. We have performed calculations using the wavefunctions and electronic dispersion resulting from PBE and HSE03 functional calculations (more hybrid functional starting points are shown Supplementary Fig. 10). The corresponding bandstructures for the

monoclinic geometry of TNSe can be seen in Fig. 3. Independently of the starting point, the bandgap converges towards a similar value, increasing in the PBE case, where it is commonly underestimated, and decreasing for the HSE03 functional. In all cases,  $G_0W_0$  predicts a bandgap between 100 and 163 meV, which is in good agreement with the experimental gap of 160 meV measured in optics<sup>26</sup>. The dispersion of the conduction bands shows similar features with a pronounced double well shape in  $X-\Gamma-X$  direction. The valence band dispersion, however, differs between the PBE and the HSE, with the PBE one being M-shaped and the HSE one being parabolic. This is inherited from the different initial DFT wavefunctions and electronic dispersion.

The comparison of the electronic dispersion with recent ARPES measurements by Baldini et al.<sup>28</sup> shows that the dispersion



**Fig. 3**  $G_0W_0$  bandstructure of monoclinic  $Ta_2NiSe_5$ . Bandstructure of the monoclinic geometry of TNSe after performing a  $G_0W_0$  calculation starting from the DFT groundstate of TNSe using the PBE (a) and the HSE03 hybrid functional (b). While the bandgaps are all within a small range of 100 to 163 meV, the valence band dispersions differ depending on the functional used for the initial wavefunctions. a has been adapted from our work in ref. 28.

obtained using the PBE functional as starting point agrees well with the experimentally measured one, reproducing both the flat valence band as well as the Mexican hat-shaped conduction band. Also Tang et al.<sup>25</sup> and Watson et al.<sup>24</sup> report an M-shaped valence band, which agrees well with the  $G_0W_0$ @PBE calculation. This reinforces the conclusion that the band structure in the monoclinic phase correctly accounts for the metal-to-insulator transition. A detailed quantitative comparison to experimental results, however, can only be achieved by including correlation beyond the PBE functional.

We also solved the Bethe–Salpether equation (BSE) starting from the  $G_0W_0$  results<sup>52,53</sup> to explicitly account for exciton formation. While full convergence is not feasible with our computational resources, we can extrapolate from our  $k$ -point-dependent results BSE results on top of  $G_0W_0$  and verify that the excitonic binding energy is smaller than the  $G_0W_0$  gap in the monoclinic phase (see Supplementary Notes 7 and 8 and Supplementary Figs. 13–19). This is an indication that the condition needed for an EI phase, i.e. an exciton-binding energy larger than the gap, is not fulfilled in TNSe for the monoclinic phase.

### The effect of electronic heating

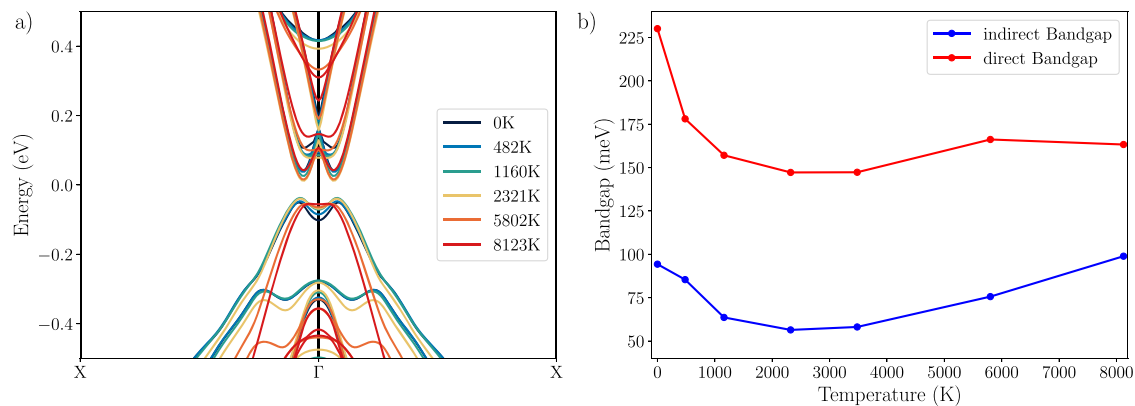
The interpretation of an excitonic insulator as a phase consisting of an excitonic condensate entails that an increase of temperature above a critical value leads to the melting of the condensate and hence to a bandgap closing. Experiments aimed at demonstrating such a behaviour to prove the possibility of excitonic condensation in TNSe have indeed reported bandgap opening and band flattening for increasing pump fluence<sup>25</sup>. Here we mimic the photoexcitation with a thermal distribution of electrons and holes and show that the variation of the bandgap and the band flattening can be explained via a temperature increase in the same manner as in standard semiconductors. Indeed, an increase of temperature of the thermal distribution generates free carriers and it has been shown that for low carrier densities the bandgap shrinks and eventually increases again for sufficiently high densities<sup>54,55</sup>. The mechanism behind such a bandgap variation can be rationalized in terms of effective screening of the Coulomb interaction which modulates the effect of bandgap opening caused by the exchange contribution.

To approximately investigate the enhanced screening effects due to free carriers, we have performed a set of  $G_0W_0$  calculations for different electronic temperatures. The electronic temperatures have been set by changing the temperature in the Fermi–Dirac function used for the occupation smearing of the Kohn–Sham states in the self-consistent DFT calculations. This effectively leads to a depletion of electrons in the valence bands and a finite electronic occupancy of the conduction bands. Specifically, we have investigated Fermi–Dirac distributions with temperatures ranging between 0 and 8000 K.  $G_0W_0$  calculations on such thermalized groundstates can then provide bandstructures, which include the effect of the screening due to the thermally excited carriers. The resulting bandstructures and bandgap renormalizations are presented in Fig. 4a, b. One observes that introducing charged carriers to the conduction bands rapidly decreases the bandgap. It reaches its minimum around 2500 K with a bandgap 50% smaller than its equilibrium value and reopens for higher temperatures. This behaviour is consistent with the typical bandgap renormalization in standard semiconductors mentioned above. The pronounced double well feature of the valence band slightly flattens out as the  $\Gamma$ -point energy is shifted upwards (see Fig. 4a) and only for electronic temperature >4000 K the bandstructure changes significantly, which explains the upwards trend in the bandgap for high temperatures. To reach such high electronic temperatures with a laser, however, is experimentally unrealistic as it would damage the material.

We remark that, even though the applied computational procedure does not properly describe the experimental photoexcitation of electrons into conduction bands, we could expect a behaviour similar to the one described above upon intensive laser pumping.

### Origin of the structural instability

Having established the importance of the structural instability in the determination of the insulating groundstate of TNSe, in this section we discuss possible origins of the structural instability. On general grounds, the structural instability is the result of the complex interplay between the electronic and the lattice degrees of freedom. The possibility of a purely electronic origin of the structural instability, namely a charge density that spontaneously breaks the orthorhombic symmetry at fixed ions position, has been recently discussed in the literature. Specifically, this possibility has been introduced at the model level in ref. 35 as the result of an electronic interaction-driven spontaneous hybridization between localized Wannier orbitals of different symmetry. Moreover, experimental data from symmetry-resolved Raman have been interpreted as signatures of quasi-critical charge



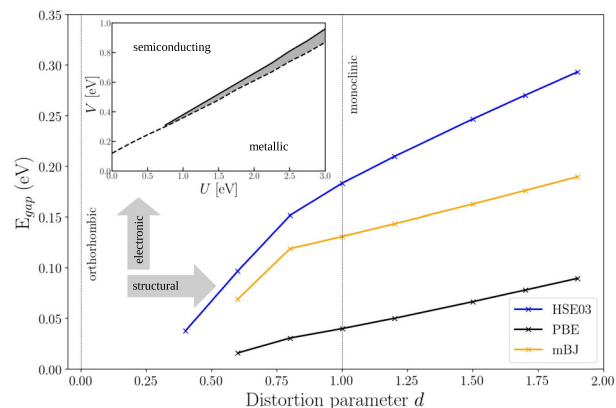
**Fig. 4 Bandstructure of  $\text{Ta}_2\text{NiSe}_5$  for different electronic temperatures.** DFT+ $G_0W_0$  calculation using different electronic smearing temperatures. **a** shows the electronic dispersion with increased temperature. The obtained bandstructures agree very well with the flat bands obtained in ARPES measurements. **b** shows the evolution of the bandgap with increasing electronic temperature. It decreases rapidly to 50% of its equilibrium value at 2000 K and increases afterwards due to an increased band flattening for very high temperatures. The direct bandgap has been measured at  $\Gamma$ . The figure has been adapted from our work in ref. <sup>28</sup>.

fluctuations that are active in the corresponding symmetry channel<sup>56</sup>.

We start by recalling the mechanism behind such an electronic instability. The mechanism can be readily understood in terms of the energetic competition between on-site (Ta-Ta and Ni-Ni) and the nearest neighbouring (Ta-Ni) density–density interaction that can render the groundstate of the system to be either metallic or insulating. In a simplified description, this energetic competition can be well understood as a problem of two electrons in two sites with on-site interaction  $U$  and inter-site interaction  $V$ . If the energy of the lowest lying orbitals in the two sites are comparable, the groundstate configuration of the problem is entirely determined by the ratio between  $U$  and  $V$ . In the limit  $U \gg V$ , the lowest energy configuration corresponds to the single occupation of the lowest lying orbitals on the two sites, whereas in the opposite limit  $U \ll V$  the most favourable configuration is the doubly occupation of the site with the lowest energy orbital. The two configurations become quasi-degenerate for  $U \sim V$  such that a coherent superposition between the two configuration due to an exchange interaction mediated by  $V$  is possible.

Assuming that in the extended system the lowest lying orbitals in the two sites contribute to the formation of two bands close to the Fermi level, the two configurations correspond respectively to a metal in which both bands are partially occupied or to an insulator in which the lowest lying band is fully occupied. The coherent superposition, instead, indicates the possibility of a spontaneous hybridization between the bands that leads to the opening of an hybridization gap. As discussed in ref. <sup>35</sup>, this can happen only if the bands have different symmetry character and therefore an hybridization of this type would correspond to a charge density that breaks the orthorhombic lattice symmetry. Calculations based on a mean-field ansatz for a realistic model of TNSe show that the spontaneous hybridization transition can occur in a very narrow region around  $U \sim 4V$  (see inset of Fig. 5), where the factor four comes from the fact that there are four neighbouring Ta atoms around each Nickel. We emphasize that the phase diagram in the inset of Fig. 5 from ref. <sup>35</sup> is intended to be qualitative as it is specifically calculated for a set of realistic parameters corresponding to the mBJ functional and a selection of Wannier orbitals, hence not necessarily representative for other functionals.

The smallness of this region suggests that the electronic instability relies on a delicate energetic balance between the competing metallic and insulating states, which is not guaranteed to be exactly satisfied in the real material. We therefore exploit the present full ab initio results to test this energetic balance in the



**Fig. 5 Bandgap of  $\text{Ta}_2\text{NiSe}_5$  along the phase transition.** Bandgap as a function of the distortion parameter  $d$  defined in the text for different exchange correlation functionals for TNSe: for all considered exchange functionals, we obtain a positive bandgap only for finite distortions  $d$ , which then increases as a function of the distortion parameter and the functional chosen. A minimal distortion breaking all relevant lattice symmetries ( $d = 0.05$ ) does not open a bandgap, which would be expected for an electronically driven instability (see Supplementary Note 5 and Supplementary Fig. 8). We displayed only data points that have a finite bandgap. Grey arrows indicates the bandgap opening due to structural distortion and electronic correlation using different functionals. The inset shows the phase diagram of TNSe adapted from ref. <sup>35</sup>. The region between dashed and solid (grey shaded) line shows the region with an electronic instability breaking the orthorhombic lattice symmetry towards the monoclinic. Above is the semiconducting region without electronic symmetry breaking and below the metallic region.

case of TNSe and, in particular, how it is affected by the lattice distortions that are neglected in the electronic instability discussion above. We do so by performing bandgap calculations, reported in Fig. 5, for different functionals on increasingly distorted structures using the distortion parameter  $d$ . To define the distortion parameter we linearly interpolate between the high-symmetric orthorhombic and low-symmetric monoclinic geometry: we parameterized this transition with a transition vector  $\mathbf{v}_t$  defined as the difference between the atomic configuration of the high-symmetry and low-symmetry structure  $\mathbf{v}_t = \mathbf{v}_m - \mathbf{v}_o$ , with  $\mathbf{v}_m$  describing the atomic configuration of the monoclinic cell and  $\mathbf{v}_o$  describing the atomic configuration of the orthorhombic cell.

Using this vector, we can linearly interpolate between the orthorhombic and monoclinic structure via  $\mathbf{v}_i(d) = \mathbf{v}_o + d \cdot \mathbf{v}_v$ , where  $d$  is the distortion parameter. The lattice vectors are parameterized in a similar way. This means that also each of the three lattice vectors of both geometries are linearly interpolated using the difference between the monoclinic and orthorhombic lattice vectors  $\mathbf{v}_{\text{lat},t} = \mathbf{v}_{\text{lat},m} - \mathbf{v}_{\text{lat},o}$ . Hence, the interpolated lattice vectors are  $\mathbf{v}_{\text{lat},i}(d) = \mathbf{v}_{\text{lat},o} + d \cdot \mathbf{v}_{\text{lat},t}$  for each of the three lattice vectors. If a gap opens as a result of the electronic instability only, we could expect that an infinitesimal distortion  $d \rightarrow 0^+$  would be sufficient to stabilize an insulating groundstate with a distorted charge density. However, we observe that, independently of the functional used, a finite critical value of the distortion parameter  $d_c$  is needed in order to open a gap (see Fig. 5). This means that TNSe is indeed likely outside the region where a purely electronic instability can be expected and that a finite distortion of the lattice is critical to account for the metal–insulator transition. Nonetheless, we observe a cooperation between electronic correlations and lattice distortion in the opening of the gap. Indeed, by increasing the accuracy of the functional (from PBE to HSE), which corresponds to increasing amount of electronic exchange and correlation, the bandgap at fixed distortion is increased and the critical value of the distortion parameter  $d_c$  needed for opening a gap is significantly reduced.

We therefore conclude that, in relation to the simple mean-field phase diagram with tunable  $U$  and  $V$  interaction derived for the orthorhombic phase of TNSe, the material must be placed on the semimetallic region of the phase diagram and that the lattice distortions act in the same direction of the nearest neighbour interaction  $V$ . On the other hand, based on the previous electronic band structure results, the Sulfur compound must be placed on the semiconducting side of the phase diagram, corresponding to a larger effect of the interaction  $V$ .

As a further confirmation of the fact that the non-local interaction  $V$  is responsible for the opening of the gap in these materials, we also performed fully ab initio Hartree–Fock calculations for TNSe, which are supposed to include a larger contribution from non-local correlations: both the orthorhombic and the monoclinic phase have semiconducting groundstates with bandgaps of 436 and 836 meV, respectively (see Supplementary Fig. 11).

We finally mention that despite the fact that a purely electronic contribution is absent for TNSe we cannot exclude the system to be close enough to an electronic instability so that quasi-critical charge fluctuations can be present in the symmetry channel corresponding to the breaking of the mirror symmetries that characterizes the transition from the orthorhombic to the monoclinic phase<sup>56</sup>. This analysis goes beyond the scope of the present work and will be the subject of future investigations.

In the following, we focus on the possibility of having an energetically favourable electronic instability that, however, does not open a bandgap but instead stays in the metallic regime close to the orthorhombic geometry. We introduce the total energy  $E \equiv E[d, \rho_-]$ , as a function of two coordinates. The first one, which accounts for the lattice degrees of freedom, is the lattice distortion  $d$ . The second coordinate of the energy functional is the component  $\rho_-$  of the electronic charge density, which corresponds to a symmetry lowering from orthorhombic to monoclinic (of ref. <sup>35</sup>). It is defined as

$$\rho_- = \int_{x \in \text{unitcell}} |\rho(x) - \rho(R(x))| dx, \quad (1)$$

with  $R$  being the orthorhombic symmetry operators and  $\rho$  the charge density. For the orthorhombic cmcm geometry of TNSe, four different mirror symmetries  $R = \{\sigma_{A,o}, \sigma_{A,p}, \sigma_{B,o}, \sigma_{B,p}\}$  have been identified as relevant to the orthorhombic to monoclinic transition<sup>35</sup>, where  $p$  and  $o$  denote mirror symmetries parallel and orthogonal to the chain direction and  $A$  and  $B$  denote the

corresponding one-dimensional chains in the unit cell. We will focus on these mirror symmetries in the following. In this picture, the groundstate is defined such that at  $d^*$  and  $\rho_-^*$  the total energy is minimal:

$$\left. \frac{\partial E}{\partial d} \right|_{d=d^*} = \left. \frac{\partial E}{\partial \rho_-} \right|_{\rho_-=\rho_-^*} = 0. \quad (2)$$

If the system undergoes a spontaneous electronic instability, the charge density breaks the orthorhombic lattice symmetry for infinitesimal distortions and hence:

$$\lim_{d \rightarrow 0} \rho_-(d) \neq 0. \quad (3)$$

The two possible scenarios for a lattice and an electronically driven transition are shown schematically in Fig. 6. Figure 6a shows the case of a lattice-driven phase transition. The charge density component  $\rho_-$  does not break the orthorhombic symmetries at  $d=0$  and evolves smoothly towards the groundstate of the system with increasing distortion. Figure 6b shows the case of an electronically driven phase transition where the charge density breaks the orthorhombic lattice symmetries. In this case, the charge density component  $\rho_-$  does not vanish at  $d=0$  but it is finite instead and its value increases with increasing distortion. In both cases of Fig. 6a, b, a finite  $d$ -distortion may still be needed to obtain a positive bandgap.

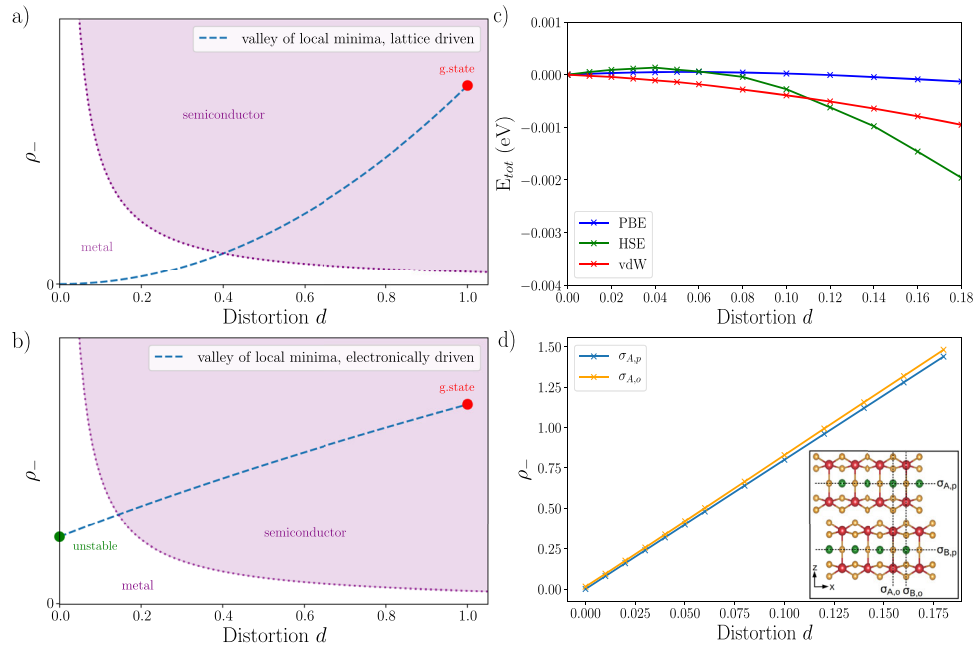
When we perform ab initio calculations imposing the orthorhombic symmetries, the charge density is automatically symmetrized so that  $\rho_-(d=0) = 0$  is true by construction. However, if we compute  $\rho_-$  for small lattice distortions that break the orthorhombic symmetry, we can evaluate the limit  $\lim_{d \rightarrow 0} \rho_-(d)$  and

verify whether an electronic instability is present. Furthermore, an electronic symmetry breaking should also result in a discontinuity in the total energy for infinitesimal lattice distortion, as a phase with a finite  $\rho_-^*$  has to lower the energy with respect to the symmetric orthorhombic phase in the electronic instability scenario. Both the computed total energy as well as  $\rho_-$  for fixed lattice distortions  $d$  are depicted in Fig. 6c, d.

Figure 6c shows that there is no discontinuity in the total energy. For the vdw-optB88 functional, the total energy decreases monotonously to its minimum for increasing  $d$  (see Supplementary Note 2 and Supplementary Fig. 1 for  $q > 0.2$ ). For the PBE and HSE functionals, it increases first and subsequently decreases towards its minimum, which for the HSE03 functional is at around  $d = 1.6$ . We attribute such a non-monotonous behaviour to the fact that the distortion parameter  $d$  is defined with respect to the vdw-optB88 functional structures. This means that in the phase-space of the normal phononic mode we are not moving along the soft phonon modes but along a direction where the orthorhombic phase is a saddle point. Solving this issue would require a structure relaxation using hybrid functionals to obtain the exact soft mode coordinate, which is computationally not feasible for TNSe. Figure 6d shows the evolution of the charge density component  $\rho_-$  for different symmetry operations defined in ref. <sup>35</sup>.  $\sigma_{A,p}$  describes the reflection symmetry parallel the Nickel chain in the  $xz$ -plane and  $\sigma_{A,o}$  describes the reflection in the  $yz$ -plane orthogonal to the chain direction and along the Tantalum atoms (see inset in Fig. 6). For both reflection symmetries,  $\rho_-$  approaches 0 for  $d \rightarrow 0$ . This reinforces the conclusion that the transition in TNSe requires a finite lattice distortion and we do not observe a spontaneous electronic instability.

### Phonon dispersion

So far, we have investigated the electronic degrees of freedom focusing on the effect of the lattice geometries. While we have introduced the concept of a structural instability in the prior



**Fig. 6 Origin of the phase transition.** Sketch of the energy landscape of  $E[d, \rho_-]$  along the local minima valleys for **a** a lattice-driven and **b** an electronically driven phase transition. The green dot marks the groundstate defined via  $d^*$  and  $\rho_-^*$ . In the case of a lattice-driven phase transition (**a**), the charge density component  $\rho_-$  evolves smoothly from 0 to  $\rho_-^*$  as the lattice is distorted. In the case of an electronically driven phase transition, the charge density breaks the orthorhombic symmetries for infinitesimal distortions. Thus,  $\rho_-$  jumps discontinuously to a finite value (green dot marked as unstable) and evolves afterwards smoothly to the groundstate. The purple line marks that even with an electronic instability a finite  $d$  distortion may in general be needed to obtain a positive bandgap. **c** Total energy as a function of the lattice distortion for the PBE, vdW-optB88 and HSE03 functional. The total energy of the respective orthorhombic geometry was taken as reference value and set to 0. **d** Increase of the symmetry breaking parameter  $\rho_-$  with increasing lattice distortion using the HSE03 functional. If there was an electronic instability, both curves would have exhibited a discontinuity for infinitesimal lattice distortions  $d$ , with  $\rho_-$  extrapolating to a positive non-zero value. A linear fit of both curves is given by  $f_{\sigma_{A,o}}(d) = 7.9964(\pm 0.0041) \cdot d + 2.2e - 4(\pm 4.1e - 4)$  and  $f_{\sigma_{A,p}}(d) = 8.1658(\pm 0.0120) \cdot d + 2.64e - 3(\pm 0.26e - 3)$ . The inset shows the orthorhombic symmetry planes defined by Mazza et al.<sup>35</sup>.

sections, we show how it is related to the lattice dynamics of the crystal.

The phonon dispersion of the orthorhombic as well as the monoclinic cell of TNSe are shown in Fig. 7a: the orthorhombic cell shows two unstable optical phonon modes around the  $\Gamma$ -point, characterized by imaginary frequencies (indicated as negative values here), while the monoclinic cell shows no phonon instabilities.

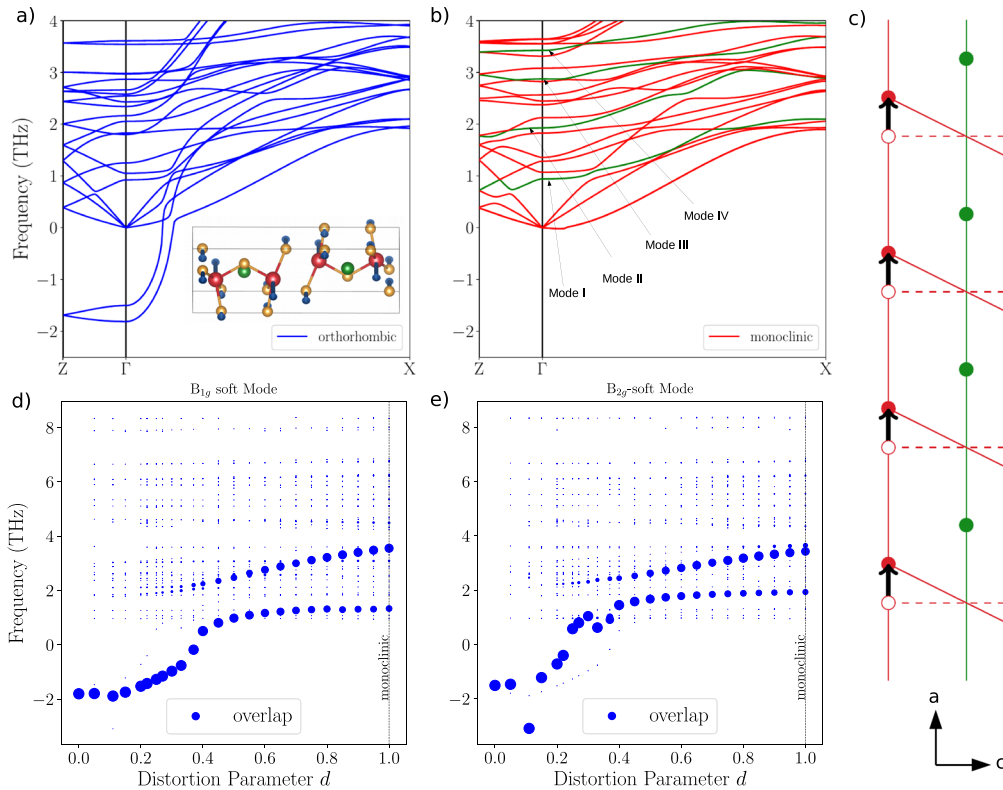
The two unstable phonon modes belong to the  $B_{1g}$  and  $B_{2g}$  symmetry groups. Both unstable phonons lead to monoclinic distortions and become stable in the monoclinic phase. The  $B_{1g}$  leads to a  $P2_1/m$  geometry and the  $B_{2g}$  mode leads to the  $C2/c$  symmetry. Just displacing the atoms within the unit cell and computing the total energy of the structure shows that the  $P2_1/m$  would be favoured<sup>36</sup>, however, also allowing for lattice distortions, the  $C2/c$  symmetry is favoured with an energy difference of 14.4 meV per unit cell. This is in agreement with the full relaxation as well as experimental findings observing a monoclinic  $C2/c$  symmetric geometry in the low-temperature phase of TNSe<sup>32,57</sup>. Therefore, we can identify the  $B_{2g}$  mode as the relevant structural instability, which drives the phase transition. The eigenvector of this mode is displayed in the inset of Fig. 7a. It shows the characteristic monoclinic displacement along the chain direction, which results in a shearing of the Ta atoms around the Nickel atom (see Fig. 7c), which has been identified as the order parameter for the orthorhombic to monoclinic phase transition by Kaneko et al.<sup>21</sup>. Note that a similar phonon analysis using ab initio methods has been performed in ref. <sup>36</sup> and our calculations agree with this study.

For the sake of comparison, we have performed a calculation of the phononic spectrum at the  $\Gamma$ -point for both geometries of TNS. We find one phononic instability in the orthorhombic structure,

which becomes stable in the monoclinic phase (see Supplementary Table 7). The instability has  $B_{2g}$  symmetry and drives, similarly to the TNSe case, the orthorhombic-to-monoclinic phase transition, which is consistent with the monoclinic groundstate being the energetically favoured one.

To investigate the evolution of the phonon instabilities when transforming the crystal from orthorhombic to monoclinic, we have performed calculations on increasingly distorted structures using the  $d$  distortion parameter introduced in the previous section. Performing phonon calculations, we obtained the phonon spectra and eigenvectors at the  $\Gamma$ -point for each of the distorted geometries. To follow the evolution of the phonon energies during the structural transition, we have computed the overlap between the orthorhombic eigenvectors of the two phonon instabilities with the phononic eigenvectors of all modes at a given value of the distortion parameter  $d$ . The result is displayed in Fig. 7c, d where the marker size is proportional to such overlaps. Note that the phonon eigenvectors for a given distortion  $d$  form an orthonormal system. Therefore, no overlap is visible at  $d = 0$  for the two unstable modes. As the structure is transformed to the monoclinic cell, the different phonons start hybridizing that leads to non-zero overlap of the unstable orthorhombic phonons and the phonons at  $d \neq 0$ . We observe that the two unstable phonons are being stabilized along the phase transition and hybridize predominantly into a pair of monoclinic phonons. The unstable  $B_{2g}$  has the strongest overlap with the monoclinic 3.4 and 1.9 THz phonon modes and the unstable  $B_{1g}$  has the strongest overlap with the monoclinic 3.6 and 1.4 THz modes. In section “Electron–phonon coupling,” we will show that the two monoclinic  $B_{2g}$  phonons couple strongly to the electronic degrees of freedom.

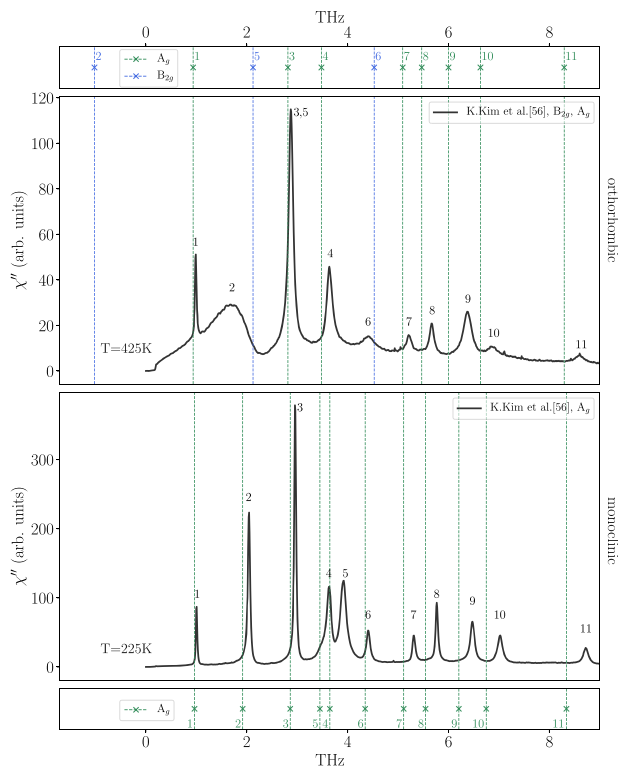




**Fig. 7 Phonon properties of  $\text{Ta}_2\text{NiSe}_5$ .** **a** TNSe phonon dispersion using the vdW-optB88 functional for the orthorhombic phase. It shows two optical instabilities, one of which is the  $B_{2g}$  soft mode. The inset shows the eigenvector of this mode. **b** Phonon dispersion of the monoclinic phase. The two unstable modes become stable. We highlighted the phonon branches of the four optical phonons I–IV for which the coupling to the electron bandstructure is discussed in the section “Electron–phonon coupling.” The lowest acoustic phonon in the monoclinic phase shows an unnatural behaviour close to  $\Gamma$ , which is an artefact of the numerical procedure applied. **c** Schematic illustration of the shearing of the Tantalum atoms around the Nickel atoms, which has also been discussed in refs. <sup>21,36</sup> and is a key signature of the phase transition. **d, e** show the overlap of the unstable orthorhombic phonon eigenvectors with the phonon eigenvectors at different distortions  $d$ . The position of the blue dots show the phonon eigenenergy of the corresponding mode and the size the overlap between the mode eigenvector and the unstable orthorhombic mode  $B_{1g}$  mode (**d**) and  $B_{2g}$  mode (**e**) eigenvector. At  $d = 0$ , we have the orthorhombic geometry and at  $d = 1$  the monoclinic geometry. One sees that the orthorhombic eigenvectors have maximal overlaps with the modes at 3.4 and 3.6 THz in the monoclinic phase.

Further analysis of the phonon symmetries allows us to identify all three  $B_{2g}$  and eight  $A_g$  phonon modes of the orthorhombic and all  $A_g$  phonon modes of the monoclinic geometry. Their comparison with the phononic peak positions from recent experimental Raman measurements that are sensitive to  $B_{2g}$  and  $A_g$  symmetry channel<sup>33,56,58,59</sup> shows a good agreement of all monoclinic  $A_g$  modes, except for one, and all orthorhombic  $A_g$  modes, even if our results are at  $T = 0$ . Only the first two  $B_{2g}$  modes, which are labelled 2 and 5 in the orthorhombic phase, and mode 5 in the monoclinic phase show a substantial energy difference between the calculated and the experimental results (see Figs. 8 and 9). In the most recent Raman experiments, no soft mode is reported<sup>33,56,58–60</sup> and the second  $B_{2g}$  mode (labelled 5) differs by roughly 0.7 THz from the theoretically calculated value at  $T = 0$  K. Note that Kim et al.<sup>56</sup> identify mode 5 as being almost degenerate with mode 3 in the orthorhombic phase and it evolves into mode 5 of the monoclinic phase (see curve at 225 K), where we also see a small discrepancy with our calculation. The two  $B_{2g}$  modes exhibit a significant Fano-like broadening, while all other Raman active modes have a sharp lineshape. It is conceivable that the electronic structure methods used here do not properly capture all coupling channels of the phonon modes to the electronic degrees of freedom, which could be responsible for these effects. Another possible explanation are inelastic phonon–phonon contributions due to anharmonic effects that are present for the two  $B_{2g}$  modes. At the experimental temperature, phonon anharmonicity effects are expected to play a

major role in renormalizing the phonon energies and stabilizing the unstable modes, as has been demonstrated in other materials exhibiting a soft mode-mediated phase transition<sup>61</sup>. This is especially relevant for the soft modes, which would explain why a very good agreement is observed for the other Raman modes even without including temperature effects. The anharmonic coupling of the soft mode to the continuum of other phononic modes could also be the source of the broadening of the Raman peak in the region just below 2 THz, which has been instead attributed to the coupling of the phonon to an excitonic instability. The idea of the fundamental role of phonon anharmonicity is supported by recent experimental Raman measurements on TNS by Ye et al.<sup>33</sup>. Indeed it is shown that, even though TNS undergoes a structural phase transition, no phonon softening is observed in the Raman data around the transition temperature at  $T = 120$  K. As in the case of TNSe, the authors claim that the phase transition is the result of an acoustic mode softening, which in this case is coupled to a ferroelastic instability rather than an excitonic instability. In contrast, our calculations do not show softening of the acoustic modes for both TNSe and TNS but rather soft Raman active  $B_{2g}$  modes (see Supplementary Fig. 7). Therefore, the lack of phonon softening in the Raman measurements in TNS is likely the result of the temperature stabilization of that phonon, as explained above, and provides a further support to the structural similarities between TNS and TNSe. A detailed investigation of the temperature effects in the phonon spectrum to fully corroborate this argumentation will be

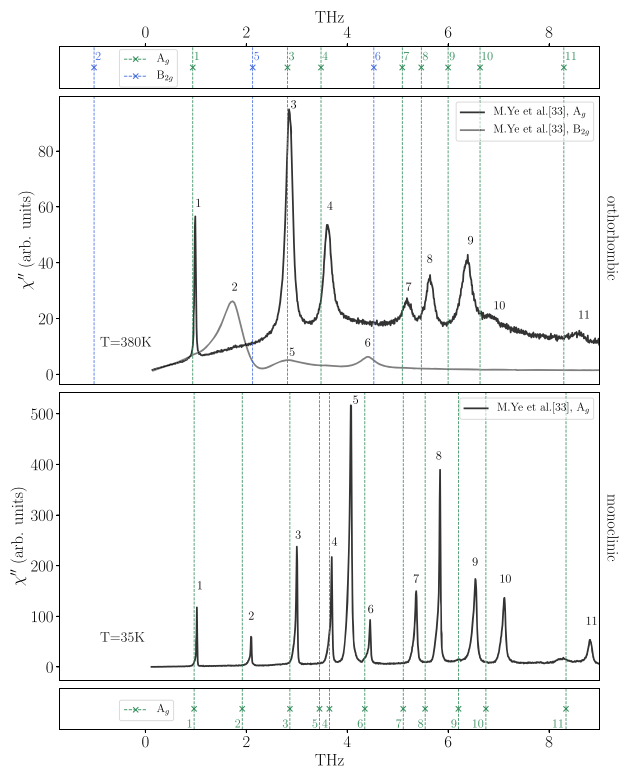


**Fig. 8 Comparison of the Raman data from Kim et al.<sup>56</sup> with our theory results.** Comparison of the theoretically calculated phonon eigenenergies at  $T=0$  K with the Raman spectra provided by Kim et al.<sup>56</sup>. The top panel shows the orthorhombic phonon spectrum and the bottom panel the monoclinic phonon spectrum. Note that peaks 3 and 5 are almost degenerate in energy and peak 5 only appears in the  $B_{2g}$ -channel with a much smaller amplitude than peak 3, which appears in the  $A_g$  channel. The theory spectra are obtained at  $T=0$  K and are in good agreement with Raman spectra. Only the first two  $B_{2g}$  modes of the orthorhombic geometry (peaks 2 and 5) differ substantially from their experimental values. This is consistent with these peaks having a significant broadening. The numerals in the experimental plots label the experimentally measured peaks by Kim et al.<sup>56</sup>. We assigned those numbers to the corresponding energies from our theory calculations by minimizing the energy differences and following the modes across the transition.

topic of future work. In the following section, we investigate the coupling of such  $B_{2g}$  modes to the electronic bandstructure using a frozen phonon approach. We found that especially phononic modes, which displace the Tantalum and Nickel atoms along the chain direction of the crystal, such as the orthorhombic  $B_{2g}$  modes, strongly modify the electronic bandstructure and hence could in turn be renormalized.

### Electron–phonon coupling

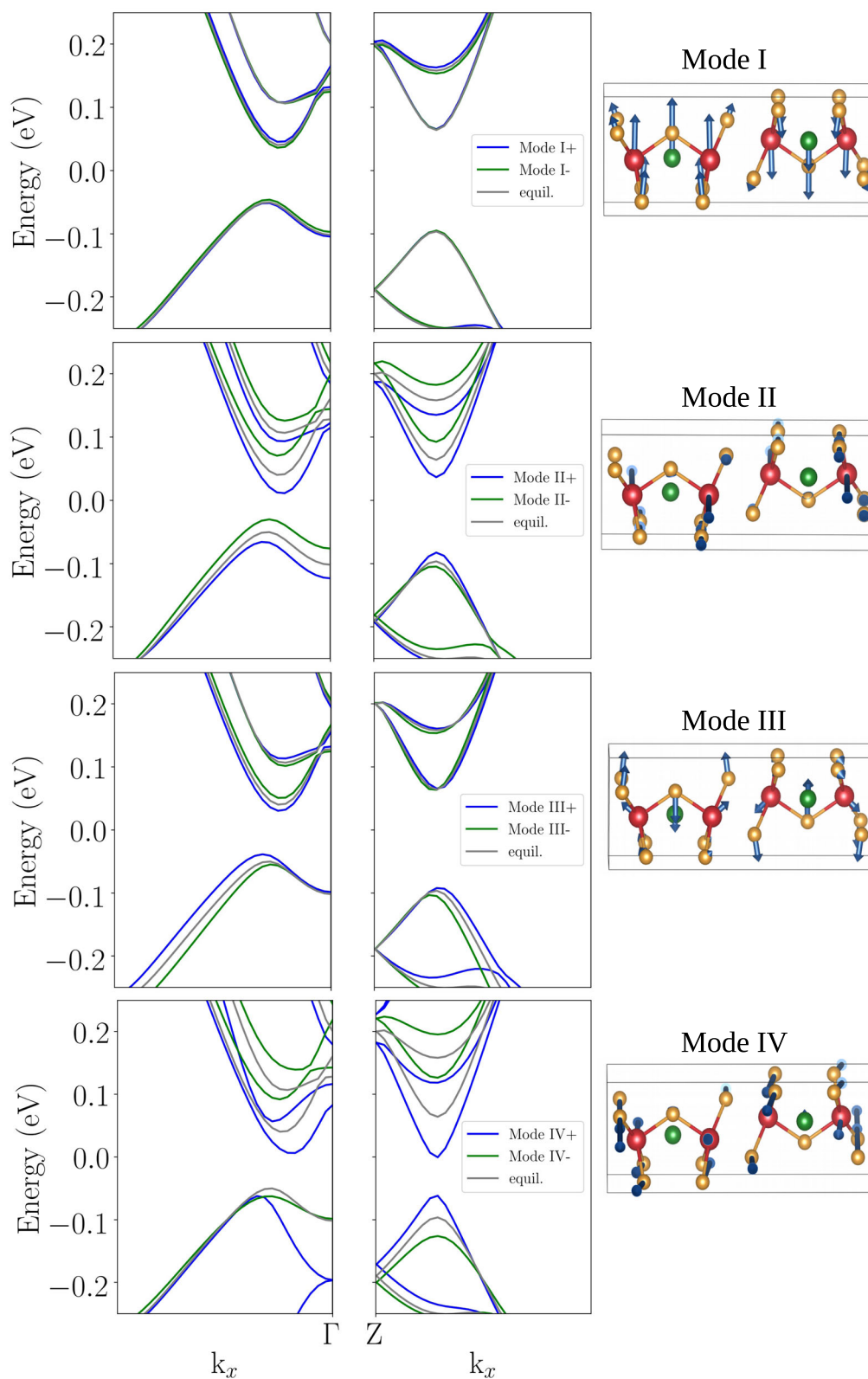
We conclude our work by demonstrating the intimate coupling between the electronic and nuclear degrees of freedom by highlighting the influence of specific phononic modes on the electronic bandstructure. Time-resolved ARPES measurements by Baldini et al. find that the gap presents oscillatory behaviour under photoexcitation<sup>28</sup>. The authors show that the momentum integrated photoelectron intensity is dominated by 4 frequency components: 0.98, 2.11, 3.0, and 3.67 THz. The four frequencies can be identified as the frequencies of Raman active phonons, and using our phonon dispersion calculation (Fig. 7a), we can identify the corresponding modes, with mode I at 1.1 THz, mode II at



**Fig. 9 Comparison of the Raman data from Ye et al.<sup>33</sup> with the theory results.** Comparison of the theoretically calculated phonon eigenenergies at  $T=0$  K with the Raman spectra provided by Ye et al.<sup>33</sup>. The top panel shows the orthorhombic phonon spectrum and the bottom panel the monoclinic phonon spectrum on a linear scale. The theory spectra are obtained at  $T=0$  K and are in good agreement with the Raman spectrum. Only the first two  $B_{2g}$  modes of the orthorhombic geometry differ substantially from their experimental values. This is consistent with these peaks having a significant broadening. The numerals in the experimental plots label the experimentally measured peaks. We assigned those numbers to the corresponding energies from our theory calculations by minimizing the energy differences and following the modes across the transition.

1.9 THz, mode III at 2.9 THz and mode IV at 3.4 THz. We observe that the modes II and IV have  $B_{2g}$  symmetry and show the characteristic shearing with respect to the Tantalum atoms around the Nickel atoms, while the modes I and III that have  $A_g$  symmetry do not exhibit this shearing. They can be identified with the peaks 1, 2, 3 and 5 in the Raman measurements of Figs. 8 and 9. Mode IV mode can also be identified as the monoclinic counterpart of the  $B_{2g}$  soft mode, which is unstable in the orthorhombic geometry (see Fig. 7e).

To quantify how strongly the above discussed phonons couple to the electrons, we have calculated their effect on the electronic bandstructure by displacing the atoms along the phonon eigenmodes by an amount equal to the square root of their mean-squared displacement at zero temperature for both the positive and negative direction. The resulting bandstructures using the  $G_0W_0$  method are shown in Fig. 10 (see Supplementary Fig. 12 for the PBE results). We see that the modes I and III show little influence, whereas both  $B_{2g}$  modes have a strong impact on the electronic structure. We stress that both of these modes exhibit the characteristic shearing of the Ta atoms around the Nickel atom in chain direction, which we also observed during the structural phase transition (see Fig. 7c). For mode IV, the average gap between positive and negative displacement is larger than the equilibrium gap. Therefore, we expect a small increase of around 21 meV of the bandgaps predicted in the previous



**Fig. 10 Frozen phonon bandstructure.** PBE+ $G_0W_0$  calculation using the mean-squared displacement at  $T=0$  K along different phonon eigenvectors in both positive and negative direction. The bandstructure couples strongly to the  $B_{2g}$  phonon modes II and IV at 2 and 3.4 THz, which exhibit the characteristic shearing of the Ta atoms. The figure has been adapted from our work in ref. <sup>28</sup>.

section due to the coupling of this phonon to the electronic degrees of freedom.

These results, beside confirming the strong influence of phonons on the electronic structure, hint towards the possibility of modulating the electronic bandgap through an external pumping of the above phononic modes.

## DISCUSSION

We have shown that the orthorhombic phase of both TNSe and TNS is unstable at low temperatures and exhibits a soft phonon mode, which signals the structural transition to the monoclinic phase.

Performing electronic bandstructure calculations using different exchange correlation functionals, as well as calculations using the  $G_0W_0$  approximation, we have proven that in TNSe the structural transition drives the metal-to-semiconducting phase transition, with an indirect bandgap that is comparable to the experimental results<sup>26,27</sup>, while in TNS the semiconductive electronic nature is preserved. This demonstrates that without the structural transition the experimentally observed gap opening at the critical temperature cannot be explained. At the same time, our ab initio results show that no excitonic effects are necessary to explain the experimental evidences. Furthermore, during the structural transition a significant flattening of the valence bands is observed which is enhanced by electronic heating leading to almost flat bands, which are in agreement with recently measured (pump-probe) ARPES signatures<sup>28</sup>.

The electronic heating shows that the bandgap never closes for increasing electronic temperatures but shows the expected renormalization effects of standard semiconductors due to the increased carrier density.

We discussed the possible origins of the structural transition in relation to the scenario of a purely electronic instability occurring at fixed ion positions and concluded that a structural distortion is required in order to observe a metal-insulator transition in TNSe. This is supported by scans of the total energy landscape and charge density as a function of the relevant distortion parameters, which show no sign of an electronic instability.

Investigating the change of the electronic bandstructure for experimentally relevant phonon modes, we found that the electrons tend to couple stronger to the phonon modes, which displace along the chain direction and induce the characteristic shearing of the Ta atoms around the Ni atom. Especially the 3.4 THz phonon mode is expected to lead to a small additional bandgap opening through the thermal occupation of this mode. We conclude that, in the low-temperature phase, TNSe and TNS behave as standard semiconductors.

## METHODS

### Computational details

All DFT and  $G_0W_0$  calculations have been performed with the Vienna Ab initio Simulation Package (VASP)<sup>62–66</sup>, and in the case of phonon calculations, we additionally used phonopy to compute phonon dispersions<sup>67</sup>. For the interpolation of the  $G_0W_0$  bandstructure, we have used the wannier90 package<sup>68</sup>. All calculations have made use of the PBE pseudopotential set generated with VASP<sup>69</sup>.

The structural relaxations have been performed on a  $24 \times 4 \times 6$   $k$ -mesh with an energy cutoff of 460 meV. A total free energy change of  $1e-6$  eV has been used as cutoff criterion, which led to forces of 2.8 meV/angstrom per ion or less. The resulting input structures are shown in section 1 of the Supplementary Information. The phonon dispersion calculation has been performed using the vdw-optB88 functional and a  $4 \times 2 \times 1$  supercell with a  $4 \times 4 \times 3$   $k$ -mesh. The post-processing has been performed with the phonopy package<sup>67</sup>. The calculation of the phonon spectra at  $\Gamma$  that are compared to the Raman data has been performed on a  $48 \times 4 \times 6$   $k$ -mesh using a single unit cell. The calculation of the phonon overlaps as a

function of the distortion parameter  $d$  has been performed on a  $16 \times 8 \times 3$   $k$ -mesh.

The electronic bandstructure calculations have been performed on a  $24 \times 8 \times 6$   $k$ -mesh for the PBE functional and a  $16 \times 4 \times 4$   $k$ -mesh for the mBJ and HSE hybrid functionals and the Hartree-Fock calculation. The mBJ functional calculated self-consistently a  $c$  value of 1.2578 for the orthorhombic and a  $c$ -value of 1.2636 for the monoclinic phase of TNSe. For TNS,  $c$  is 1.2620 for both structural phases. The bandstructures have been computed along the path  $M = (\frac{\pi}{a}, 0, \frac{\pi}{c})$ ,  $Z = (0, 0, \frac{\pi}{c})$ ,  $\Gamma = (0, 0, 0)$  and  $X = (\frac{\pi}{a}, 0, 0)$ .

The  $G_0W_0$  calculations have been performed on a  $12 \times 4 \times 2$   $k$ -mesh with 160 frequencies  $\omega$  and 1160 states with a 100 meV energy cutoff. The Wannierization has been performed with the Wannier90 package<sup>68</sup>. A convergence study of the relevant parameters for the presented  $G_0W_0$  calculation can be found in the Supplementary Note 7 and Supplementary Figs. 13–15.

## DATA AVAILABILITY

The authors declare that the main data supporting the findings of this study are contained within the paper and its associated Supplementary Information. All other relevant data are available from the corresponding author upon reasonable request.

## CODE AVAILABILITY

The VASP software package is available at <https://www.vasp.at/>. The open-source software packages Wannier90 is available at <http://www.wannier.org/> under the GNU General Public License (v2). The open-source phonopy package is available at <https://phonopy.github.io/phonopy/> under the BSD license.

Received: 16 June 2021; Accepted: 19 November 2021;

Published online: 20 December 2021

## REFERENCES

- Kohn, W. Excitonic phases. *Phys. Rev. Lett.* **19**, 439–442 (1967).
- Jérome, D., Rice, T. M. & Kohn, W. Excitonic insulator. *Phys. Rev.* **158**, 462–475 (1967).
- Mott, N. F. The transition to the metallic state. *Philos. Mag.* **6**, 287–309 (1961).
- Keldysh, L. V. & Kozlov, A. N. Collective properties of excitons in semiconductors. *Sov. Phys. JETP* **27**, 978 (1968).
- Halperin, B. I. & Rice, T. M. Possible anomalies at a semimetal-semiconductor transition. *Rev. Mod. Phys.* **40**, 755–766 (1968).
- Kozlov, A. The metal-dielectric divalent crystal phase transition. *Sov. Phys. JETP* **21**, 790 (1965).
- Cercellier, H. et al. Evidence for an excitonic insulator phase in  $1T - \text{TiSe}_2$ . *Phys. Rev. Lett.* **99**, 146403 (2007).
- Kogar, A. et al. Signatures of exciton condensation in a transition metal dichalcogenide. *Science* **358**, 1314–1317 (2017).
- Bucher, B., Steiner, P. & Wächter, P. Excitonic insulator phase in  $\text{TmSe}_{0.45}\text{Te}_{0.55}$ . *Phys. Rev. Lett.* **67**, 2717–2720 (1991).
- Ma, L. et al. Strongly correlated excitonic insulator in atomic double layers. *Nature* **598**, 585–589 (2021).
- Jia, Y. et al. Evidence for a monolayer excitonic insulator Preprint at <https://arxiv.org/abs/2010.05390> (2020).
- Varsano, D., Palumbo, M., Molinari, E. & Rontani, M. A monolayer transition-metal dichalcogenide as a topological excitonic insulator. *Nat. Nanotechnol.* **15**, 367–372 (2020).
- Gupta, S., Kutana, A. & Yakobson, B. I. Heterobilayers of 2d materials as a platform for excitonic superfluidity. *Nat. Commun.* **11**, 2989 (2020).
- Varsano, D. et al. Carbon nanotubes as excitonic insulators. *Nat. Commun.* **8**, 1461 (2017).
- Li, Z. et al. Possible excitonic insulating phase in quantum-confined sb nano-flakes. *Nano Lett.* **19**, 4960–4964 (2019).
- Li, J. I. A., Taniguchi, T., Watanabe, K., Hone, J. & Dean, C. R. Excitonic superfluid phase in double bilayer graphene. *Nat. Phys.* **13**, 751–755 (2017).
- Perali, A., Neilson, D. & Hamilton, A. R. High-temperature superfluidity in double-bilayer graphene. *Phys. Rev. Lett.* **110**, 146803 (2013).
- Hu, Y., Venderbos, J. W. F. & Kane, C. L. Fractional excitonic insulator. *Phys. Rev. Lett.* **121**, 126601 (2018).
- Du, L. et al. Evidence for a topological excitonic insulator in  $\text{InAs}/\text{GaSb}$  bilayers. *Nat. Commun.* **8**, 1971 (2017).

20. Ataei, S. S., Varsano, D., Molinari, E. & Rontani, M. Evidence of ideal excitonic insulator in bulk  $\text{mos}(2)$  under pressure. *Proc. Natl Acad. Sci. USA* **118**, e2010110118 (2021).
21. Kaneko, T., Toriyama, T., Konishi, T. & Ohta, Y. Orthorhombic-to-monoclinic phase transition of  $\text{Ta}_2\text{NiSe}_5$  induced by the Bose-Einstein condensation of excitons. *Phys. Rev. B* **87**, 035121 (2013).
22. Wakisaka, Y. et al. Excitonic insulator state in  $\text{Ta}_2\text{NiSe}_5$  probed by photoemission spectroscopy. *Phys. Rev. Lett.* **103**, 026402 (2009).
23. Seki, K. et al. Excitonic Bose-Einstein condensation in  $\text{Ta}_2\text{NiSe}_5$  above room temperature. *Phys. Rev. B* **90**, 155116 (2014).
24. Watson, M. D. et al. Band hybridization at the semimetal-semiconductor transition of  $\text{Ta}_2\text{NiSe}_5$  enabled by mirror-symmetry breaking. *Phys. Rev. Res.* **2**, 013236 (2020).
25. Tang, T. et al. Non-Coulomb strong electron-hole binding in  $\text{Ta}_2\text{NiSe}_5$  revealed by time- and angle-resolved photoemission spectroscopy. *Phys. Rev. B* **101**, 235148 (2020).
26. Lu, Y. F. et al. Zero-gap semiconductor to excitonic insulator transition in  $\text{Ta}_2\text{NiSe}_5$ . *Nat. Commun.* **8**, 14408 (2017).
27. Lee, J. et al. Strong interband interaction in the excitonic insulator phase of  $\text{Ta}_2\text{NiSe}_5$ . *Phys. Rev. B* **99**, 075408 (2019).
28. Baldini, E. et al. The spontaneous symmetry breaking in  $\text{Ta}_2\text{NiSe}_5$  is structural in nature. Preprint at <https://arxiv.org/abs/2007.02909> (2020).
29. Kaneko, T., Toriyama, T., Konishi, T. & Ohta, Y. Electronic structure of  $\text{Ta}_2\text{NiSe}_5$  as a candidate for excitonic insulators. *J. Phys. Conf. Ser.* **400**, 032035 (2012).
30. Nakano, A. et al. Antiferroelectric distortion with anomalous phonon softening in the excitonic insulator  $\text{Ta}_2\text{NiSe}_5$ . *Phys. Rev. B* **98**, 045139 (2018).
31. Di Salvo, F. et al. Physical and structural properties of the new layered compounds  $\text{Ta}_2\text{NiS}_5$  and  $\text{Ta}_2\text{NiSe}_5$ . *J. Less Common Met.* **116**, 51–61 (1986).
32. Sunshine, S. A. & Ibers, J. A. Structure and physical properties of the new layered ternary chalcogenides tantalum nickel sulfide ( $\text{Ta}_2\text{NiS}_5$ ) and tantalum nickel selenide ( $\text{Ta}_2\text{NiSe}_5$ ). *Inorg. Chem.* **24**, 3611–3614 (1985).
33. Ye, M. et al. Lattice dynamics of the excitonic insulator  $\text{Ta}_2\text{Ni}(\text{Se}_{1-x}\text{S}_x)_5$ . *Phys. Rev. B* **104**, 045102 (2021).
34. Mu, K. et al. Electronic structures of layered  $\text{Ta}_2\text{NiS}_5$  single crystals revealed by high-resolution angle-resolved photoemission spectroscopy. *J. Mater. Chem. C* **6**, 3976–3981 (2018).
35. Mazza, G. et al. Nature of symmetry breaking at the excitonic insulator transition:  $\text{Ta}_2\text{NiSe}_5$ . *Phys. Rev. Lett.* **124**, 197601 (2020).
36. Subedi, A. Orthorhombic-to-monoclinic transition in  $\text{Ta}_2\text{NiSe}_5$  due to a zone-center optical phonon instability. *Phys. Rev. Mater.* **4**, 083601 (2020).
37. Klimeš, J., Jev, Bowler, D. R. & Michaelides, A. Van der Waals density functionals applied to solids. *Phys. Rev. B* **83**, 195131 (2011).
38. Klimeš, J., Bowler, D. R. & Michaelides, A. Chemical accuracy for the van der Waals density functional. *J. Phys. Condens. Matter* **22**, 022201 (2009).
39. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
40. Tran, F. & Blaha, P. Accurate band gaps of semiconductors and insulators with a semilocal exchange-correlation potential. *Phys. Rev. Lett.* **102**, 226401 (2009).
41. Becke, A. D. & Johnson, E. R. A simple effective potential for exchange. *J. Chem. Phys.* **124**, 221101 (2006).
42. Becke, A. D. & Roussel, M. R. Exchange holes in inhomogeneous systems: a coordinate-space model. *Phys. Rev. A* **39**, 3761–3767 (1989).
43. Heyd, J., Scuseria, G. E. & Ernzerhof, M. Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.* **118**, 8207–8215 (2003).
44. Sugimoto, K., Nishimoto, S., Kaneko, T. & Ohta, Y. Strong coupling nature of the excitonic insulator state in  $\text{Ta}_2\text{NiSe}_5$ . *Phys. Rev. Lett.* **120**, 247602 (2018).
45. Li, L. et al. Strong in-plane anisotropies of optical and electrical response in layered dimetal chalcogenide. *ACS Nano* **11**, 10264–10272 (2017).
46. Larkin, T. I. et al. Giant exciton Fano resonance in quasi-one-dimensional  $\text{Ta}_2\text{NiSe}_5$ . *Phys. Rev. B* **95**, 195144 (2017).
47. van Schilfgaarde, M., Kotani, T. & Faleev, S. Quasiparticle self-consistent GW theory. *Phys. Rev. Lett.* **96**, 226402 (2006).
48. Onida, G., Reining, L. & Rubio, A. Electronic excitations: density-functional versus many-body Green's-function approaches. *Rev. Mod. Phys.* **74**, 601–659 (2002).
49. Hybertsen, M. S. & Louie, S. G. First-principles theory of quasiparticles: calculation of band gaps in semiconductors and insulators. *Phys. Rev. Lett.* **55**, 1418–1421 (1985).
50. Hybertsen, M. S. & Louie, S. G. Electron correlation in semiconductors and insulators: band gaps and quasiparticle energies. *Phys. Rev. B* **34**, 5390–5413 (1986).
51. Hedin, L. New method for calculating the one-particle Green's function with application to the electron-gas problem. *Phys. Rev.* **139**, A796–A823 (1965).
52. Salpeter, E. E. & Bethe, H. A. A relativistic equation for bound-state problems. *Phys. Rev.* **84**, 1232–1242 (1951).
53. Sander, T., Maggio, E. & Kresse, G. Beyond the Tamm-Dancoff approximation for extended systems using exact diagonalization. *Phys. Rev. B* **92**, 045209 (2015).
54. Spataru, C. D., Benedict, L. X. & Louie, S. G. Ab initio calculation of band-gap renormalization in highly excited GaAs. *Phys. Rev. B* **69**, 205204 (2004).
55. Beni, G. & Rice, T. M. Theory of electron-hole liquid in semiconductors. *Phys. Rev. B* **18**, 768–785 (1978).
56. Kim, K. et al. Direct observation of excitonic instability in  $\text{Ta}_2\text{NiSe}_5$ . *Nat. Commun.* **12**, 1969 (2021).
57. Nakano, A. et al. Pressure-induced coherent sliding-layer transition in the excitonic insulator  $\text{Ta}_2\text{NiSe}_5$ . *IUCr* **5**, 158–165 (2018).
58. Volkov, P. A. et al. Critical charge fluctuations and emergent coherence in a strongly correlated excitonic insulator. *NPJ Quantum Mater.* **6**, 52 (2021).
59. Volkov, P. A. et al. Failed excitonic quantum phase transition in  $\text{Ta}_2\text{Ni}(\text{Se}_{1-x}\text{S}_x)_5$ . Preprint at <https://arxiv.org/abs/2104.07032> (2021).
60. Kim, M.-J. et al. Phononic soft mode and strong electronic background behavior across the structural phase transition in the excitonic insulator  $\text{Ta}_2\text{NiSe}_5$  (with erratum). Preprint at <https://arxiv.org/abs/2007.01723> (2021).
61. Zhou, J.-J., Hellman, O. & Bernardi, M. Electron-phonon scattering in the presence of soft modes and electron mobility in  $\text{SrTiO}_3$  perovskite from first principles. *Phys. Rev. Lett.* **121**, 226603 (2018).
62. Kresse, G. & Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558–561 (1993).
63. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
64. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comp. Mater. Sci.* **6**, 15–50 (1996).
65. Kresse, G. & Hafner, J. Norm-conserving and ultrasoft pseudopotentials for first-row and transition elements. *J. Phys. Condens. Matter* **6**, 8245–8257 (1994).
66. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999).
67. Togo, A. & Tanaka, I. First principles phonon calculations in materials science. *Scr. Mater.* **108**, 1–5 (2015).
68. Mostofi, A. A. et al. An updated version of wannier90: a tool for obtaining maximally-localised Wannier functions. *Comput. Phys. Commun.* **185**, 2309–2310 (2014).
69. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999).

## ACKNOWLEDGEMENTS

We are grateful to E. Baldini, I. Mazin and Yann Gallais for enlightening discussions throughout the course of this work. We would like to thank M. Ye, G. Blumberg, K. Kim, B.J. Kim, M.J. Kim and S. Kaiser for sharing the experimental data of their Raman measurements and valuable discussions. This work is supported by the European Research Council (ERC-2015-AdG-694097), Grupos Consolidados (IT1249-19) and the Flatiron Institute, a division of the Simons Foundation. We acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG) under Germany's Excellence Strategy - Cluster of Excellence Advanced Imaging of Matter (AIM) EXC 2056 - 390715994 and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - SFB-925 - project 170620586. Support by the Max Planck Institute - New York City Center for Non-Equilibrium Quantum Phenomena is acknowledged. S.L. acknowledges support from the Alexander von Humboldt foundation. G.M. acknowledges support of the Swiss National Science Foundation FNS/SNF through an Ambizione grant.

## AUTHOR CONTRIBUTIONS

A.R. and S.L. designed the project. L.W. performed the computations under the supervision of S.L. All authors discussed the computational results and contributed to the writing and revision of the manuscript.

## FUNDING

Open Access funding enabled and organized by Projekt DEAL.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41524-021-00675-6>.

**Correspondence** and requests for materials should be addressed to Lukas Windgätter, Simone Latini or Angel Rubio.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021