

This is a pre print version of the following article:

Machine Learning Model for Efficient Nonthermal Tuning of the Charge Density Wave in Monolayer NbSe₂ / Benic, L., Grasselli, F., Ben Mahmoud, C., Novko, D., Loncaric, I.. - In: JOURNAL OF CHEMICAL THEORY AND COMPUTATION. - ISSN 1549-9618. - 21:16(2025), pp. 8130-8141. [10.1021/acs.jctc.5c00959]

Terms of use:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

01/07/2026 23:04

(Article begins on next page)

Machine learning model for efficient nonthermal tuning of the charge density wave in monolayer NbSe₂

Luka Benić,¹ Federico Grasselli,^{2,3} Chiheb Ben Mahmoud,⁴ Dino Novko,^{5,*} and Ivor Lončarić^{1,†}

¹*Ruđer Bošković Institute, Zagreb, Croatia*

²*Dipartimento di Scienze Fisiche, Informatiche e Matematiche,
Università degli Studi di Modena e Reggio Emilia, Modena, Italy*

³*CNR NANO S3, Modena, Italy*

⁴*Inorganic Chemistry Laboratory, Department of Chemistry,
University of Oxford, Oxford OX1 3QR, United Kingdom*

⁵*Centre for Advanced Laser Techniques, Institute of Physics, Zagreb, Croatia*

Understanding and controlling the charge density wave (CDW) phase diagram of transition metal dichalcogenides is a long-studied problem in condensed matter physics. However, due to complex involvement of electron and lattice degrees of freedom and pronounced anharmonicity, theoretical simulations of the CDW phase diagram at the density-functional-theory level are often numerically demanding. To reduce the computational cost of first principles modelling by orders of magnitude, we have developed an electronic free energy machine learning model for monolayer NbSe₂ that allows changing both electronic and ionic temperatures independently. Our approach relies on a machine learning model of the electronic density of states and zero-temperature interatomic potential. This allows us to explore the CDW phase diagram of monolayer NbSe₂ both under thermal and laser-induced nonthermal conditions. Our study provides an accurate estimate of the CDW transition temperature at low cost and can disentangle the role of hot electrons and phonons in nonthermal ultrafast melting process of the CDW phase in NbSe₂.

I. INTRODUCTION

In addition to having exceptional optical properties, transition-metal dichalcogenides (TMDs) have gained a lot of attention due to their complex phase diagram, containing various charge-ordered phases, such as charge density waves (CDW), superconductivity, and even signs of order fluctuations [1, 2]. One notable example is NbSe₂, which supports both CDW and superconductivity already at ambient pressure [3, 4], while both can be modified or quenched with an increase in applied pressure or temperature [5, 6]. An interesting interplay between CDW and superconductivity was observed when the number of layers in NbSe₂ is varied from bulk up to the monolayer limit [7, 8]. For instance, the value for superconductivity temperature T_{SC} drops from 7.2 K in bulk, to 3 K in the monolayer [7]. The corresponding modifications to the CDW phase when NbSe₂ is thinned down to a monolayer limit are not so clear, and are still a matter of active debate [8–11]. Namely, while certain experimental measurements suggest that the CDW transition temperature T_{CDW} in the NbSe₂ single layer is $T_{CDW} = 145$ K, namely significantly increased compared to the bulk [8, 10, 11], there are also reported experimental evidences that the CDW phase remains intact and T_{CDW} is close to the value of 33.5 K as in the bulk [9]. In order to resolve this controversy, a thorough quantitative and efficient computational study needs to be performed that is able to incorporate all the relevant many-body interactions ruling the CDW phase in NbSe₂.

Several theoretical studies have inspected the origin of the CDW in NbSe₂ and have clearly proven that there are no electronic instabilities in NbSe₂ that would lead to the standard Peierls transition, while the CDW order is induced mainly by the orbital-dependent and anisotropic electron-phonon coupling [12–14], as also confirmed by the experiments [15, 16]. On top of that, anharmonic effects were shown to be important for the proper and quantitative description of the CDW phase transition [14, 17, 18]. Furthermore, *ab initio* anharmonic study based on stochastic self-consistent harmonic approximation (SSCHA), which includes quantum and thermal fluctuations at a non-perturbative level, reported weak dimensionality effect in NbSe₂, with $T_{CDW} = 59$ K and 73 K for bulk and monolayer, respectively [19, 20]. On the other hand, the results based on *ab-initio* and path-integral molecular dynamics showed more stronger dimensionality dependence with $T_{CDW} = 25 - 50$ K for bulk, and $T_{CDW} = 75 - 100$ K for the monolayer [21]. With this, it is obvious that the dimensionality controversy in NbSe₂ is still not settled. Even though both of these theoretical approaches are suitable to resolve this issue, they are computationally demanding and it is not easy to reach the numerical convergence.

A powerful approach to control the CDW state, as well as to unravel the origin of CDW transition is laser-induced ultrafast dynamics, where it is possible to separate electron and phonon contributions to the CDW in time and with different excitation conditions [22–24]. Commonly, an ultrafast sub-picosecond rise or decay of the signal in time-resolved photoemission and transient optical spectroscopy is attributed to the purely electron-related origin of the CDW [25], such as excitonic insulator [26], while slower picosecond dynamics to the

* dino.novko@gmail.com

† ivor.loncaric@gmail.com

phonon-driven mechanisms, namely electron-phonon and phonon-phonon interactions [22]. Recent ultrafast time-resolved measurements in NbSe₂ reported a fast transient electronic component, the existence of the CDW amplitude phonon mode, melting excitation density, involvement of strong electron-phonon coupling, and slow dynamics of order parameter [27–30]. These experimental observations suggest an unconventional involvement of various many-body interactions in photo-excited NbSe₂, however, so far there are no *ab initio* theoretical studies that could provide a deeper insight and tackle electron- and lattice-related fluctuation in non-thermal conditions.

Modelling phase diagrams of materials efficiently and from first principles is a long-sought-after goal of computational material science. Machine learning (ML) interatomic potentials have recently reached the same accuracy as the baseline density functional theory (DFT) on which they are trained to predict total energies and forces, while offering orders of magnitude faster evaluation. Given that the training dataset samples the configurational space well, one could, in principle, use the resulting potential to efficiently model the phase diagram of any material across pressures and temperatures. This has been successfully demonstrated for a few materials or classes of materials such as elemental systems [31], alloys [32], high-entropy alloys [33], hybrid inorganic-organic systems [34], and molecular crystals [35]. However, important features of the phase diagram of some materials, such as CDW ordering, have not yet been modelled from the first principles with ML interatomic potentials.

One of the limitations in the current machine learning interatomic potentials is that they are trained on DFT data for fixed electronic temperature T_{el} . This T_{el} is conveniently used to improve the convergence of DFT calculation and different smearing methods exist, from Fermi-Dirac smearing representing the real, but in practice very large T_{el} up to methods that try to better represent $T_{el} \approx 0$ K case. However, T_{el} has an important effect on the potential energy surface and, thus, the phase diagram of NbSe₂ and similar materials. One solution would be to train a separate ML for each electronic temperature set by the Fermi-Dirac distribution. This is highly impractical for modelling phase diagram as a function of temperature. Additionally, setting a low Fermi-Dirac smearing temperature requires dense sampling of the Brillouin zone, making DFT calculations computationally more demanding. Moreover, if one would study the phase diagram under ultrafast laser conditions, ion and electron temperatures have to be separately controlled as light couples differently to electron and ion subsystems.

To solve these general issues, and to resolve some of the long-standing questions in the NbSe₂ phase diagram, we have constructed an ML model for electronic free energy that enables independent control of both electron and phonon temperatures. Our ML model is based on learning the $T_{el} = 0$ K ML interatomic potential from DFT energies and forces and learning the ML model of the

electronic density of states at $T_{el} = 0$ K [36]. From these two ML models it is then possible to calculate the electronic free energy potential for any T_{el} . We have used our model to compute the phase diagram of NbSe₂ as a function of electronic and ionic temperatures that compares with previous DFT-based calculations at orders of magnitude lower computational cost. In particular, for equilibrium conditions we get $T_{CDW} \approx 104$ K, which agrees with *ab initio* molecular dynamics results [21] that suggest a finite dimensionality effect in NbSe₂. We have then simulated response of NbSe₂ to a short laser pulse within the two-temperature model (TTM) based on DFT. Our model reveals a strong exchange of energy between lattice and electrons, which leads to sub-picosecond dynamics of electron component, and long-lived phonon part. In combination with the two-temperature ML interatomic potential, we have shown that hot electrons can efficiently melt the electron-phonon-driven CDW order in NbSe₂. These results introduce a fresh perspective on the understanding of time-resolved measurements, showing that it is not necessary to invoke additional ordering mechanisms of purely electronic origin to explain the sub-picosecond decay of the CDW-related signals.

II. MODELS AND METHODS

A. Theoretical model

To efficiently model finite electronic temperature effects, we use an approximation to the exact Helmholtz electronic free energy that depends only on quantities at $T_{el} = 0$ K. As shown in the Supplementary Material and in Ref. [36], the Helmholtz electronic free energy of the system with N_{ion} ions and N_{el} electrons can be approximated by

$$F(\underline{\mathbf{R}}; T_{el}) \approx E^0(\underline{\mathbf{R}}; 0) + \Delta F(\underline{\mathbf{R}}; 0, T_{el}) . \quad (1)$$

where $\underline{\mathbf{R}} = \{\mathbf{R}_1, \dots, \mathbf{R}_{N_{ion}}\}$ represents the set of ionic coordinates of a given system (we will call $\underline{\mathbf{R}}$ a structure from now on), $E^0(\underline{\mathbf{R}}; 0)$ is the $T_{el} = 0$ K DFT energy in its standard form [37] and $\Delta F(\underline{\mathbf{R}}; 0, T_{el})$ is the finite T_{el} correction to the Helmholtz electronic free energy, constructed out of the ground state electronic density of states (eDOS), which is defined as

$$g^0(\underline{\mathbf{R}}; \epsilon) = \sum_{s,n,\mathbf{k}} \delta(\epsilon - \epsilon_{s,n,\mathbf{k}}^0(\underline{\mathbf{R}})) , \quad (2)$$

where the sum runs over the spin projections s , bands n , and \mathbf{k} -points \mathbf{k} . The force $\mathbf{F}_n(\underline{\mathbf{R}}; T_{el})$ on the n -th ion, represented by the coordinates \mathbf{R}_n , is given by the negative gradient with respect to \mathbf{R}_n of (1)

$$\mathbf{F}_n(\underline{\mathbf{R}}; T_{el}) = -\nabla_{\mathbf{R}_n} F(\underline{\mathbf{R}}; T_{el}) \quad (3)$$

and is given in the Supplementary Material.

The Helmholtz electronic free energy $F(\underline{\mathbf{R}}; T_{el})$ in this approximation is fully determined by the $T_{el} = 0$ K

ground state energy $E^0(\mathbf{R}; 0)$ and the finite electronic temperature correction ($\Delta F(\mathbf{R}; 0, T_{el})$) constructed from the ground state eDOS $g^0(\mathbf{R}; \epsilon)$. Furthermore, the ionic force $\mathbf{F}_n(\mathbf{R}; T_{el})$ is determined by the gradients of $E^0(\mathbf{R}; 0)$ and the gradients of $g^0(\mathbf{R}; \epsilon)$.

The given approximation opens up the possibility to construct an electronic free energy ML model using the $T_{el} = 0$ K (ground state) DFT total energies (and forces) and DFT electronic density of states that would be able to predict the behaviour of a material at any electronic temperature.

B. Training dataset

To construct the dataset we performed DFT calculations using the Vienna Ab initio Simulation Package (VASP) software version 6.4.2 [38–42] with the Perdew-Burke-Ernzerhof (PBE) functional [43]. For the plane-wave basis set energy cutoff we used 270 eV which is 1.3 times larger than the default value for the VASP Se and Nb_pv pseudopotentials. To model $T_{el} = 0$ K, we have used the Methfessel-Paxton scheme of order 1 with a smearing of 0.0043 eV. To obtain accurate results, we used fine sampling of the Brillouin zone of $93 \times 93 \times 1$ for the unit cell, and for the larger supercells it was scaled accordingly. To collect data, we performed molecular dynamics with the Andersen thermostat [44] spanning temperatures from 10 to 300 K and the Bayesian-learning algorithm for on-the-fly machine learning as implemented in VASP.

The dataset consists of 2452 different structures in total, with supercell sizes of $2 \times 2 \times 1$, $3 \times 3 \times 1$, $4 \times 4 \times 1$, $6 \times 6 \times 1$, $9 \times 9 \times 1$, $6 \times 1 \times 1$, $1 \times 9 \times 1$, $12 \times 1 \times 1$ and $2 \times 9 \times 1$. The dataset was randomly split into 80% training and 20% test sets. Lattice constant for calculations was 3.4715 Å. A small fraction of $3 \times 3 \times 1$ structures in the dataset were strained, with uniaxial strains in the range of negative 1.8% to positive 1.2%.

C. Machine learning interatomic potential at $T_{el} = 0$ K

The model for $E^0(\mathbf{R}; 0)$ and its gradients was constructed with the message-passing atomic cluster expansion (MACE) architecture [45, 46]. Most of the training hyperparameters were kept to the default values (see Appendix A of [46]). To be able to accurately model acoustic phonons, the cutoff radius was set to 8 Å and we used 8 Bessel basis functions, 128 embedding channels and $l_{max} = 2$, namely, we included equivariant messages up to the second order. In the weighted loss function, energy (λ_E) and forces (λ_F) weights were set to 1 and 100, respectively. The model was trained for 2000 epochs, with batch size of 5. 80% of the original structures were taken as a training set, with 5% of them taken for the validation set, and the other 20% of the original struc-

tures comprised a test set. Stochastic weight averaging [47, 48] was performed for the last 20% of epochs, and in this part of training λ_E weight was set to 1000 and λ_F weight was kept at 100.

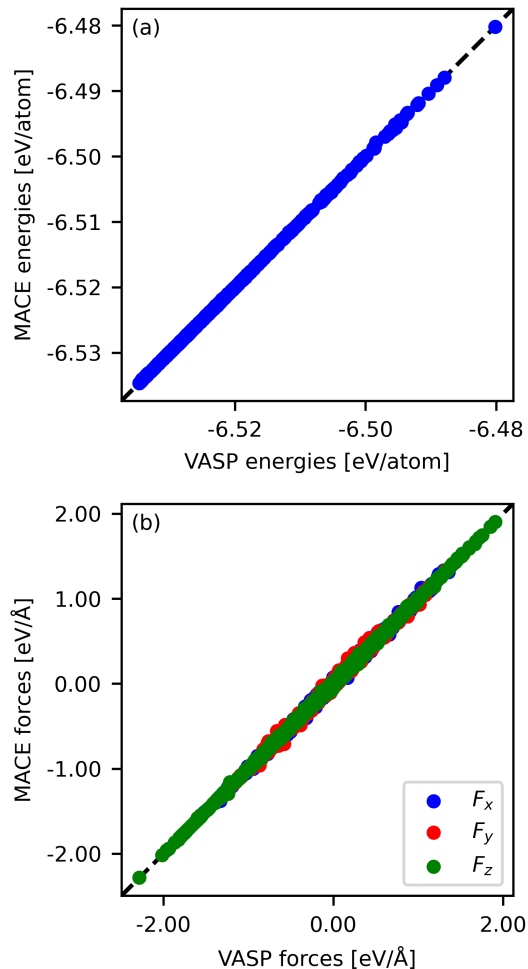


Figure 1. (a) A correlation plot between the VASP and MACE energies evaluated on the test set. (b) A correlation plot between the VASP and MACE forces per component evaluated on the test set.

In Fig. 1 we show the correspondence between the MACE and the VASP results for $E^0(\mathbf{R}; 0)$ evaluated on the test set. The performance metrics evaluated on the test set are shown in Table I.

Table I. The performance of the MACE model evaluated on the test set.

quantity	MAE	RMSE	R ²
energy	0.04 meV/atom	0.06 meV/atom	1.00
F_x	4.10 meV/Å	7.07 meV/Å	0.99
F_y	4.08 meV/Å	7.51 meV/Å	0.99
F_z	2.79 meV/Å	4.76 meV/Å	1.00

The accuracy of <0.1 meV/atom and <10 meV/Å is considered state-of-the-art [49]. However, CDW instability is related to acoustic phonons, which are typically much harder to converge. We calculated finite-difference

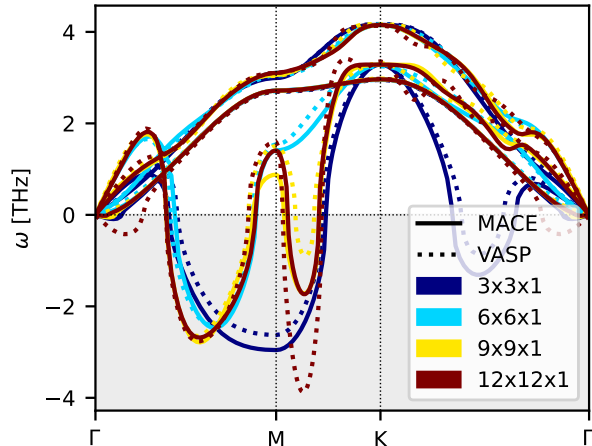


Figure 2. Comparison of the acoustic phonon dispersions for different supercell sizes at $T_{el} = 0$ K in the harmonic approximation between the MACE and VASP.

DFT phonon dispersions in the harmonic approximation for different supercell sizes, as shown in Fig. 2. We used the Phonopy code [50] with a finite difference step of 0.015 Å. Clearly, to converge the complete phonon dispersion, one would have to use supercells larger than $12 \times 12 \times 1$ making DFT calculations expensive. This points to long-range interactions in NbSe₂, arising from its two-dimensional nature, namely weak screening. As discussed below in detail, the important instability that gives rise to the CDW order between the Γ and M points is already converged for the $9 \times 9 \times 1$ supercell. The instability between M and K points is quickly removed with anharmonic corrections and, therefore, is not of particular importance. Previous DFT-based works also used supercells up to $9 \times 9 \times 1$ [20] and therefore we aim to reproduce these results with our ML potential.

The fact that interactions up to 9 unit cells (> 30 Å) are important for converging acoustic phonons makes the construction of machine learning interatomic potential challenging, since typically a cutoff radius of 5-10 Å has to be imposed. Actually, we are not aware of any trained potential for this class of materials despite both the importance of these materials and large increase of trained potentials in recent years. Using the message-passing and the relatively large initial cutoff of 8 Å as well as including the data of large supercells, we have managed to converge our machine learning potential to reproduce DFT phonons, as shown in Fig. 2.

D. Machine learning electronic density of states at $T_{el} = 0$ K

The model for eDOS $g^0(\mathbf{R}; \epsilon)$ and its gradients was trained using a Kernel Ridge Regression (KRR) model [51] as implemented in *librascal* [52], and Smooth Overlap of Atomic Positions (SOAP) descriptors [53]. As in Ref. [55] the ML model for the eDOS leverages a decomposition into local contributions from each ionic environment specified by its ionic coordinates \mathbf{R}_i

$$g^0(\mathbf{R}; \epsilon) = \sum_{\mathbf{R}_i \in \mathbf{R}} g^0(\mathbf{R}_i; \epsilon) \stackrel{(2)}{=} \sum_{\substack{s,n,\mathbf{k} \\ \mathbf{R}_i \in \mathbf{R}}} \delta(\epsilon - \epsilon_{s,n,\mathbf{k}}^0(\mathbf{R}_i)). \quad (4)$$

The Dirac delta distributions in (4) are approximated by the Gaussian distributions, defined by the Gaussian broadening hyperparameter g_b .

$$\delta(\epsilon - \epsilon_{s,n,\mathbf{k}}^0(\mathbf{R}_i)) \rightarrow \frac{1}{\sqrt{2\pi}g_b} e^{-\frac{(\epsilon - \epsilon_{s,n,\mathbf{k}}^0(\mathbf{R}_i))^2}{2g_b^2}}. \quad (5)$$

In this work, the value of $g_b = 0.01$ eV was chosen. For the construction of eDOS, we use a hyperparameter $d\epsilon = 0.00025$ eV to control the discretization of the energy grid. These two hyperparameters were set to low enough values to ensure a detailed eDOS and a finely discretized energy grid, whenever precise numerical integration is needed, see the Supplementary Material. In Fig. S2 and Fig. S3 comparison between the approximation and the VASP results for the Helmholtz electronic free energy for different values of g_b and $d\epsilon$ with $T_{el} = 500$ K are reported. As can be seen, the values we have chosen for this work are the most suitable. Both of these hyperparameters should be part of the hyperparameter optimization process, but that would be computationally too expensive.

Since all the terms in (1) and (3) are invariant to the choice of zero energy in eDOS, there is freedom in the choice of zero energy in the eDOS construction. We have chosen to align the eDOS of every structure to its Fermi energy, namely the Fermi energy is at $\epsilon = 0$ eV for every structure in the dataset. In the literature, there are other options to align the eDOS, two examples are alignment to the core electron eDOS [36] and the adaptive energy reference [56]. We also performed our calculations using alignment of the eDOS to the core electron eDOS, but there were no noticeable differences in the results. Furthermore, the cutoff radius for the SOAP descriptors for this ML model was set to 7.6 Å after the Bayesian hyperparameter optimization process. Other relevant hyperparameters are given in the supplementary note S2.

Due to the high memory consumption of the full KRR models, sparse KRR was used during the training of an ML model for eDOS. The sparsification of ionic environments to construct the so-called active set (\mathbf{A}) was performed using the Farthest Point Sampling algorithm [57] and the number of sparse ionic environments was also

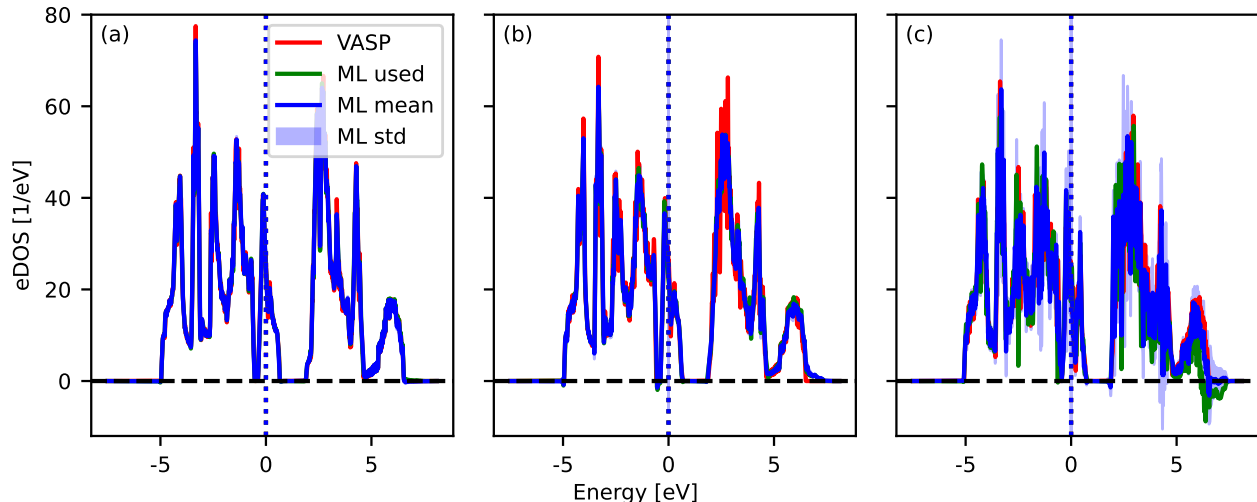


Figure 3. Results of the $T_{el} = 0$ K eDOS predictions for an ensemble of 10 models with the smallest errors (9) on the validation set during the Bayesian hyperparameter optimization process using Optuna [54]. From left to right: (a) the best, (b) median and (c) the worst predictions (with respect to the RMSE) on the test set. Red lines represent the baseline eDOS constructed from the VASP data, green lines represent the eDOS predicted by the model chosen (%RMSE on the validation set of this model is 14.49%) for the calculations done in III, blue lines represent the means of 10 models and light blue represents the standard deviation of the 10 model predictions. Vertical dotted lines represent the Fermi energies. One can observe that the largest deviations between the ML and DFT eDOS occur well beyond the Fermi level and are likely to impact calculations only at extremely high temperatures.

part of the Bayesian hyperparameter optimization process. There are 545 Se and 272 Nb ionic environments in the active set. The eDOS of the i -th local ionic environment \mathbf{R}_i of a structure $\underline{\mathbf{R}}$ can be expanded on a set of positive-definite functions (kernel functions), using the set $\underline{\mathbf{A}}$ [58],

$$g^0(\mathbf{R}_i; \epsilon) = \sum_{\mathbf{A}_j \in \underline{\mathbf{A}}} x_j(\epsilon) k(\mathbf{A}_j, \mathbf{R}_i), \quad (6)$$

where with $\mathbf{A}_j \in \underline{\mathbf{A}}$ we note the j -th ionic environment from the active set, $x_j(\epsilon)$ are the weights of the model and $k(\mathbf{A}_j, \mathbf{R}_i)$ are the kernel functions.

The kernel functions are defined using the SOAP framework and are given in [53, 55]. SOAP power spectrum was raised to a power of $\zeta > 1$, to effectively include higher-body order correlations, and ζ was part of the Bayesian hyperparameter optimization process. In this work, the value of $\zeta = 2$ was used.

From the first equality in (4) and (6) we have

$$g^0(\underline{\mathbf{R}}; \epsilon) = \sum_{\substack{\mathbf{R}_i \in \underline{\mathbf{R}} \\ \mathbf{A}_j \in \underline{\mathbf{A}}}} x_j(\epsilon) k(\mathbf{A}_j, \mathbf{R}_i) = \mathbf{k}_{\underline{\mathbf{R}}\underline{\mathbf{A}}}^T \cdot \mathbf{x}_{\underline{\mathbf{A}}}(\epsilon). \quad (7)$$

From (7) we can conclude [55]

$$\nabla_{\underline{\mathbf{R}}_i} g^0(\underline{\mathbf{R}}; \epsilon) = (\nabla_{\underline{\mathbf{R}}_i} \mathbf{k}_{\underline{\mathbf{R}}\underline{\mathbf{A}}}^T) \cdot \mathbf{x}_{\underline{\mathbf{A}}}(\epsilon). \quad (8)$$

From (8) we can see that in this model it is easy to obtain the gradients of the eDOS with respect to the ionic positions, and thus free energy forces, which is a great

advantage of this model. In our dataset we did not include any eDOS gradients, even though it could increase the precision of the model. This is because the gradients of the eDOS with DFT can be obtained only by finite differences, making it expensive and inconvenient.

To evaluate the performance of the model during training and optimize the hyperparameters, we use the metric proposed in [55]

$$\%RMSE = \frac{\sqrt{\frac{1}{N} \sum_i \int (g^0(\underline{\mathbf{R}}_i; \epsilon) - \bar{g}^0(\epsilon))^2 d\epsilon}}{\sqrt{\frac{1}{N} \sum_i \int (g^0(\underline{\mathbf{R}}_i; \epsilon) - \bar{g}^0(\epsilon))^2 d\epsilon}} \times 100, \quad (9)$$

where N is the number of structures in the target set, $\tilde{g}^0(\underline{\mathbf{R}}_i; \epsilon)$ is the predicted and $g^0(\underline{\mathbf{R}}_i; \epsilon)$ is the target 0 K eDOS of the i -th structure ($\underline{\mathbf{R}}_i$) and $\bar{g}^0(\epsilon)$ is the mean of the entire target set, given by

$$\bar{g}^0(\epsilon) = \frac{1}{N} \sum_i g^0(\underline{\mathbf{R}}_i; \epsilon). \quad (10)$$

As we can see, %RMSE is the root mean square error (RMSE) of the prediction set normalized by the standard deviation (STD) of the target set.

In Fig. 3, we show the results of an ensemble of 10 best-performing models in the Bayesian hyperparameter optimization process. Even though we trained the whole eDOS, for convenience, we only show the more challenging region around the Fermi energy. The performance of the "ML used" model from Fig. 3 on the test set is

given in Table II. In [55] it was shown that %RMSE in-

Table II. The performance of the "ML used" model from Fig. 3 on the test set.

model	MAE	RMSE	R ²	%RMSE
ML used	0.93 1/eV	4.48 1/eV	0.99	14.39%

creases as g_b decreases, and in that work %RMSE $\approx 10\%$ was achieved for $g_b = 0.3$ eV and %RMSE $\approx 20\%$ for $g_b = 0.1$ eV, while we achieve %RMSE = 14.39% for $g_b = 0.01$ eV. In Fig. 3, one can also see that the models sometimes predict negative values for the eDOS. Physically, this is impossible, but this is a consequence of the fact that the KRR model, by its construction, does not constrain eDOS predictions to be nonnegative.

E. Electronic free energy interatomic potential for finite T_{el}

By combining the models from IIC and IID we construct the full ML model for electronic free energy (1) and corresponding forces (3). To evaluate the performance of the full ML free energy potential we have randomly taken one $3 \times 3 \times 1$ structure from the test set and run VASP calculations for five different electronic temperatures. The correlation plots between the VASP and the full ML potential are shown in Fig. 4. Performance metrics are reported in Table III.

Table III. The performance of full ML potential evaluated on a one $3 \times 3 \times 1$ structure from the test set for five different electronic temperatures.

quantity	MAE	RMSE	R ²
energy	0.13 meV/atom	0.17 meV/atom	1.00
F_x	12.80 meV/Å	17.29 meV/Å	0.99
F_y	27.24 meV/Å	39.84 meV/Å	0.97
F_z	55.04 meV/Å	71.29 meV/Å	0.93

One can observe that the free electronic energy is well reproduced by our model for electronic temperatures in the range of 100 – 3000 K, with errors still considerably lower than 1 meV/atom. Forces show slightly larger errors, up to an order of magnitude larger than errors for forces of $T_{el} = 0$ ML model. Still, the errors are well below 100 meV/Å which is considered good for the kernel based methods [36]. It can also be noticed in Fig. 4 that as the electronic temperature increases, the forces become less accurate, which is also expected from the approximation in (3). It is also expected from the fact that as T_{el} increases, DFT forces also have larger fluctuations. The possible solutions to this are given in [36] and the issue is addressed in more detail in Sec. III.

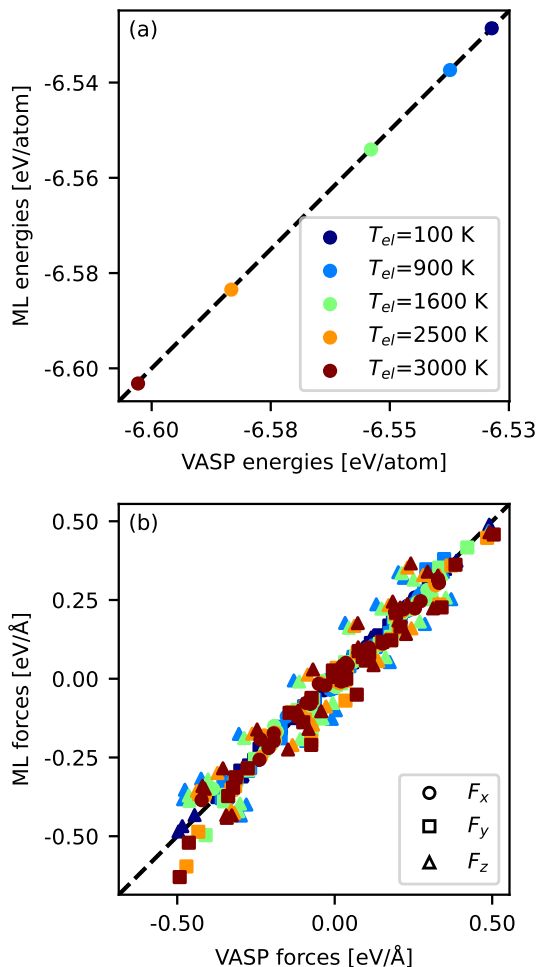


Figure 4. (a) Correlation plot between VASP and full free energy ML potential energies and (b) correlation plot between VASP and free energy forces per component for five different electronic temperatures calculated for one $3 \times 3 \times 1$ structure from the test set.

F. Modelling ionic temperature

In the previous subsections, we have presented the construction of the full ML model for the free electronic energy depending on electronic temperature (T_{el}) based on (1) and (3). To map a phase diagram, it is crucial to also include the effects of ionic temperature. While effects of ionic temperature can be modelled in several ways, we choose to use the stochastic self-consistent harmonic approximation (SSCHA) [59, 60] that is often used to model CDW materials [20, 61] as it conveniently accounts for anharmonicity at a non-perturbative level including quantum and thermal fluctuations. Previously, NbSe₂ was studied on the DFT level with SSCHA showing that significant anharmonic effects are present [20], and these results can serve as benchmark for our ML models. Since SSCHA relies only on the energy and forces from any

model, we can use it either with the $T_{el} = 0$ K MACE model to reproduce previous DFT based results, or with the full free electronic energy ML model.

Most of the SSCHA parameters were kept at their default values. Initial harmonic dynamical matrices for the SSCHA calculations were obtained with Atomic Simulation Environment [62, 63], with a finite difference step of 0.01 Å. Relaxation in SSCHA was performed with 500 configurations and, to generate ensembles, we have used 2000 configurations. To speed up calculations under laser conditions, the number of configurations was decreased to 250 and 1000, respectively. The meaningful factor was set to 10^{-5} . Both the dynamical matrix and the structure minimization steps were set to 0.05.

All SSCHA calculations were performed with the random seed set to 0, to ensure reproducibility.

III. RESULTS AND DISCUSSIONS

In the following, we discuss the phonons dispersions as a function of T_{el} and T_{ion} and the CDW phase diagram as a function of both temperatures. The signature of the CDW phase is the existence of the imaginary phonon mode in the acoustic phonon branch between Γ and M points of the Brillouin zone. As discussed earlier, and as in previous works [20], all calculations have been performed for $9 \times 9 \times 1$ unit cell that provides converged acoustic phonons in the Γ -M path.

A. Phonon dispersions as a function of T_{el}

We have calculated the phonon dispersions in the harmonic approximation as a function of electronic temperature using the full free electronic energy ML potential and VASP-DFT. The full phonon dispersions are given in Fig. S7, and the relevant acoustic phonons are given in Fig. 5. As discussed above, at $T_{el} = 0$ K the two phonon dispersions (from DFT and our model) are comparable, as our MACE model is accurate. As expected, both DFT and our model predict the stabilization of the imaginary mode with increasing electronic temperature. Although the behaviour with increasing T_{el} is similar, there are some noticeable differences. First, one can notice that DFT predicts stabilization of the normal phase already for $T_{el} = 900$ K, while in our model normal phase is obtained only for $T_{el} = 1200$ K. There are two main reasons for this difference in the temperature of transition: starting approximations in the theoretical model (1) and (3), and errors associated with ML models. To check the accuracy of the approximation (1) in the theoretical model, we have plotted in Fig. S1 the free electronic energy along the CDW mode using the exact DFT eDOS. One can see that for $T_{el} = 500$ K the approximation works well, but as T_{el} increases there is a deviation and DFT gives higher phonon frequencies. Still, the temperature of the CDW

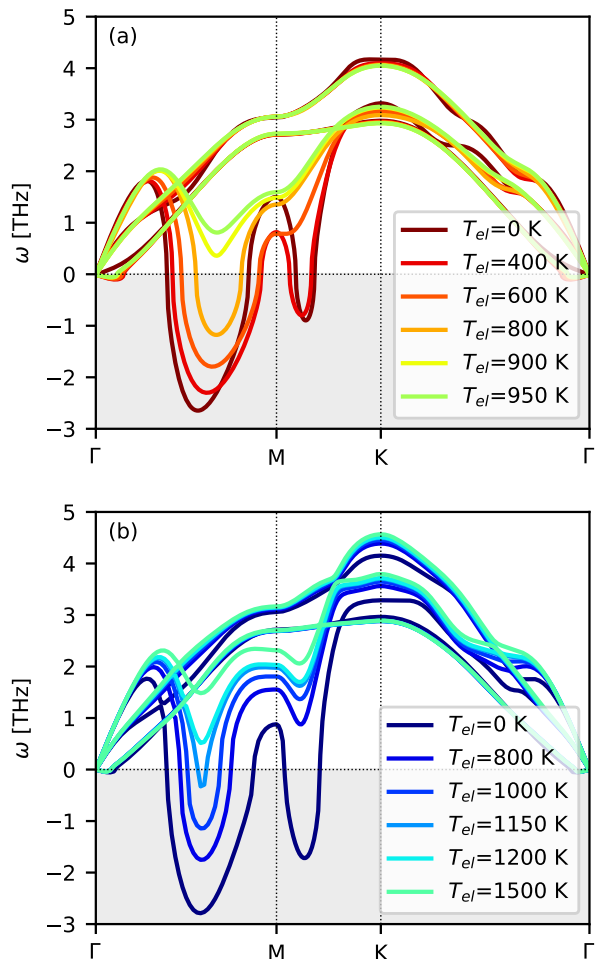


Figure 5. (a) Acoustic phonon dispersions calculated in the harmonic approximation with DFT in $9 \times 9 \times 1$ supercell. (b) Acoustic phonon dispersions calculated in the harmonic approximation with the full free electronic ML potential in $9 \times 9 \times 1$ supercell.

transition between DFT and the approximation with the correct eDOS is similar.

The second difference in phonon dispersions between the DFT and ML model is the location of the minimum of the imaginary mode (Kohn anomaly). In DFT, this minimum slightly shifts toward the M point as T_{el} increases, while in the results of ML model it does not shift. However, this is an expected behaviour of the approximation in the theoretical model. Namely, in DFT the shift in position is due to the small changes in the parts of the Fermi surface that give rise to the Kohn anomaly. In the model, on the other hand, eDOS at $T_{el} = 0$ K is used, and by construction it cannot capture the effects of changes in the electronic structure.

Since, near the CDW transition, the square of the phonon frequency ω^2 can be approximated to depend linearly on the temperature [20], in Fig. 6 we plot it

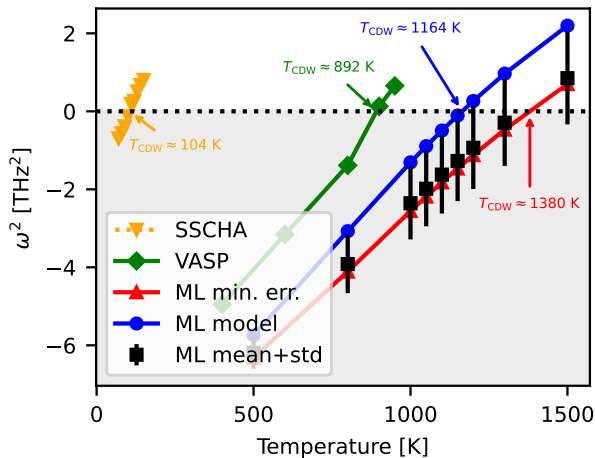


Figure 6. Plot of the ω^2 corresponding to the CDW instability along the Γ -M path versus temperature. It includes the VASP results, along with the results for the ensemble of 10 models explained earlier (in black). Also, in red we show the model with the minimal error (9) in the Bayesian hyperparameter optimization process and in blue we show the model chosen out of this ensemble that was used for the calculations in this section. In orange we show the results of the SSCHA calculations with $T_{el} = 0$ K.

for different models. We are plotting the ω^2 of the instability along the Γ -M path. DFT predicts $T_{CDW} \approx 892$ K and the ML model used for the calculations (blue in Fig. 6) predicts $T_{CDW} \approx 1164$ K. This model is chosen such that out of the 10 best models according to the %RMSE of the eDOS prediction on the validation set, it gives T_{CDW} closest to the one from DFT. The model with the lowest %RMSE on the validation set gives $T_{CDW} \approx 1380$ K. Since we have constructed the dataset using the molecular dynamics to be used generally and not only for phonon calculations, no emphasis was given to particularly sample the configurations corresponding to the CDW mode. Optimizing the ML model solely based on the %RMSE across all structures does not necessarily lead to optimal phonon predictions. While fine-tuning the dataset and model hyperparameters could improve the accuracy of phonons and T_{CDW} to match the limits of the underlying theoretical approximations, in this work we prioritized model generality.

Even with the said discrepancies, our model gives qualitatively correct results at orders of magnitude faster evaluation (for example, the calculations presented in Fig. 5(a) took around a week using VASP with 64 MPI processes on a CPU cluster, while the calculations with our ML potential presented in Fig. 5(b) took only a few minutes on a single CPU). It should also be noted that T_{CDW} and the corresponding CDW formation energy are sensitive to the calculation details and the DFT settings [21, 64, 65]. In particular, the use of hybrid HSE exchange correlation functional can increase the CDW

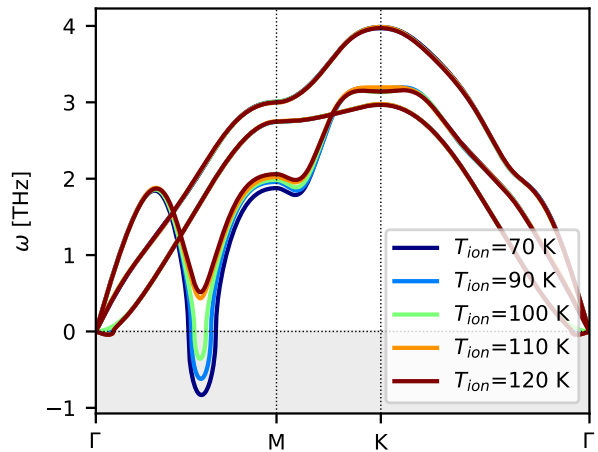


Figure 7. Acoustic phonon dispersions calculated in the anharmonic approximation using the SSCHA with $T_{el} = 0$ K and for different T_{ion} in $9 \times 9 \times 1$ supercell.

formation energy in NbSe₂ by an order of magnitude [21], while T_{CDW} in TiSe₂ is enhanced from around 1100 K to 1900 K when PBE is replaced with HSE functional [65]. This shows that our model has smaller differences in T_{CDW} compared to DFT, than the DFT approximations itself.

B. Phonon dispersions as a function of T_{ion}

In this subsection we report the acoustic phonon dispersions as a function of the ionic temperature using the SSCHA and keeping $T_{el} = 0$ K, namely for these calculations, only the MACE potential for $E^0(\mathbf{R}; 0)$ was used. Anharmonic acoustic phonon dispersions are shown in Fig. 7 and are comparable with the DFT-SSCHA calculations in Ref. [20]. Also, in Fig. 6 (orange line) it can be seen that our calculations give $T_{CDW} \approx 104$ K, which is somewhat higher than the previous SSCHA calculations that predicted $T_{CDW} \approx 70 - 80$ K. Path-integral and *ab initio* molecular dynamics simulations obtained $T_{CDW} = 75 - 100$ K [21], which is consistent with our results. The spread in values for DFT based results stems from the difficulty in converging the results due to the computational complexity of DFT for such calculations, which is avoided in our approach. Our efficient ML model for modelling the CDW phase diagram therefore suggests that CDW in monolayer NbSe₂ is enhanced compared to the bulk case. In other words, our simulations are supporting the conclusions of experiments in Refs. [8, 10, 11] which suggest $T_{CDW} > 100$ K, compared to the experiment from Ref. [9] which suggested $T_{CDW} \approx 33$ K, close to the bulk value.

C. CDW under nonequilibrium conditions

An application of the full ML interatomic potential as obtained with Eqs. (1) and (3) enables us to study the phase diagram as a function of both T_{ion} and T_{el} as shown in Fig. 8. These calculations have been performed using the SSCHA with fixed T_{el} , for which the CDW phase still exists in the harmonic limit. For each T_{el} , we vary T_{ion} until we get T_{CDW} in the same way as was done in Fig. 6. In Fig. 8(a) we report the (T_{ion}, T_{el}) phase diagram, which illustrates how the effects of electron and lattice entropy can melt the CDW phase, where obviously the latter is more dominant [20]. Nevertheless, two important conclusions regarding the role of electronic entropy emerge from this: (i) When trying to reach the quantitative values of thermal T_{CDW} it is important to choose appropriately small T_{el} in the DFT calculations. (ii) Hot electrons (high T_{el}) can efficiently melt the CDW phase that originates from electron-phonon and phonon-phonon interactions. Namely, it is commonly believed that the ultrafast creation of hot electrons melts only the electron-induced order such as the excitonic insulator and the standard Peierls transition [22, 24]. For instance, sub-picosecond decay of the CDW signals in time-resolved measurements, which follows the increase of T_{el} , is usually considered a fingerprint of the purely electronic origin of the CDW [22, 24, 25, 28]. Here we show that this can also be a fingerprint for phonon-related mechanism. Namely, soft Kohn anomaly in CDW-bearing TMDs is mostly induced by anisotropic electron-phonon scatterings, which means that anharmonicity is electron induced [14, 18, 66–68]. In NbSe₂, instability of the acoustic phonon around the 2/3 Γ M point is driven by the electron scatterings between Nb- d states around Fermi level [14], which are then diminished with elevated broadening or high T_{el} [66].

To further corroborate this result, we perform TTM simulations using the DFT input parameters [69, 70] (see supplementary note S4) and combine them with our ML model and SSCHA calculations [see Figs. 8(b)-(d)]. Considering that NbSe₂ is a metal in normal state, we assume that its electron thermalization time due to electron-electron scattering is very fast (~ 10 fs) and that the hot Fermi-Dirac distribution with T_{el} is formed in a very early stage of dynamics [70]. With this combination of techniques, we mimic the laser-induced time-resolved dynamics, as in time-resolved photoemission [29] and transient pump-probe optical spectroscopy [28, 30]. In Fig. 8(b) we show the results of the TTM simulations (see also Fig. S9), where we can see a very fast dynamics of electron component of about 0.2 ps, which agrees well with time-resolved photoemission results [29], and it comes from the strong electron-phonon coupling of $\lambda = 1.3$. For laser fluence, we chose $F = 0.1$ J/m², which was found to be a critical value above which the CDW is melted in NbSe₂ [30]. Further, in Fig. 8(c) we show phonon dispersions of acoustic phonons with anharmonic corrections that correspond to three different

snapshots, namely combinations of (T_{el}, T_{ion}) , marked in panel (b). From this it is obvious that the CDW is melted by the time when T_{el} reaches a peak, meaning slightly away from the zero time delay. In panel (d) we additionally show hardening of the CDW-related mode when T_{ion} is kept fixed at the temperature of 46 K, while T_{el} is raised according to TTM as in panels (b) and (c). By comparing these results and the one when T_{ion} is varied in time according to TTM, it is obvious that the CDW is initially melted by hot electrons and that the timescale of CDW decay will follow sub-picosecond electron excitation timescale, i.e., T_{el} . Later on, in dynamics, the CDW is further quenched by the elevated T_{ion} beyond $T_{CDW} \approx 100$ K, which supports the observation that the melting dynamics in NbSe₂ is long-lived [30].

IV. CONCLUSIONS

In this work, we have constructed the first ML interatomic potential for the monolayer NbSe₂ with fairly low errors on energies and forces. The relatively long range of our potential made it possible to accurately capture acoustic phonons that are connected with CDW order. To model the effect of electronic temperature we have also constructed an ML model of electronic density of states that together with ML interatomic potential allows for accurate model of electronic free energies and forces, thus providing a full model capable of simulating NbSe₂ dynamics and phase diagram.

We have shown that our ML models can successfully predict the CDW phase diagram of NbSe₂ in agreement with DFT at orders of magnitude lower computational cost. With this model we obtain that the CDW transition temperature is 104 K, which supports recent experimental claims that the CDW is enhanced for the monolayer NbSe₂. We then used our model to investigate the CDW phase diagram as a function of both electronic and ion temperature that arise in the non-thermal conditions such as under the short laser pulse. We show that elevated electronic temperatures efficiently melt the electron-phonon-driven CDW order, which is then long-lived as the ionic temperature equilibrates due to the electron-phonon coupling. Our results show that melting of CDW in time-resolved measurements cannot be exclusively assigned to the purely electronic origin of the CDW as often assumed. This can be especially useful for CDW-bearing systems, where time-resolved measurements are often utilized for the disentanglement of purely electronic and phonons origins to the formation of the CDW, such as for understanding the timescales of excitonic insulator and Jahn-Teller mechanisms in TiSe₂ [23] and similar TMDs.

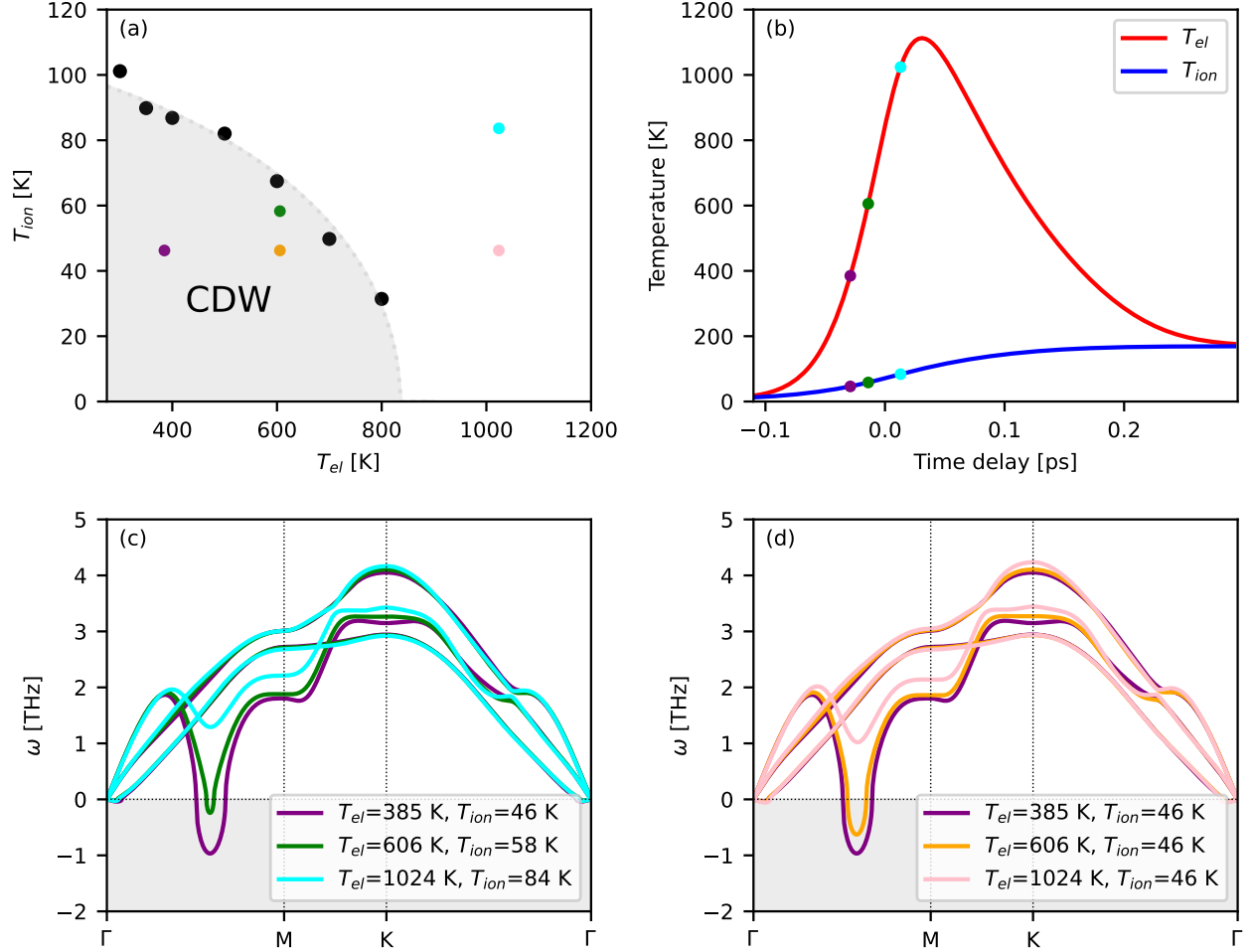


Figure 8. (a) CDW order as a function of T_{ion} and T_{el} , where the black circles indicate the pairs of T_{ion} and T_{el} after which the CDW instability disappears; the shaded area is bounded by fitting the black data points to the equation $T_{CDW}(T_{el}) = A \tanh \sqrt{1 - T_{el}/B}$ [64], where $A = 143$ K and $B = 836$ K. (b) Results of a TTM for T_{el} and T_{ion} ; three pairs of T_{el} and T_{ion} are indicated, that are used in (c). (c) Acoustic phonon dispersions calculated using full interatomic ML potential with SSCHA for three pairs of T_{el} and T_{ion} indicated in (a) and (b). (d) Acoustic phonon dispersions calculated using the full interatomic ML potential with SSCHA for the same T_{el} as in (c), but with $T_{ion} = 46$ K kept fixed, as indicated in (a).

DATA AVAILABILITY

All DFT training data and ML models, as well as calculation scripts, are available in Zenodo with the identifier 10.5281/zenodo.15125087.

CODE AVAILABILITY

All code versions and modifications in the codes are available in Zenodo with the identifier 10.5281/zenodo.15125087.

ACKNOWLEDGMENTS

LB would like to express his gratitude to Juraj Ovčar for many useful discussions. This work has been supported by Croatian Science Foundation under the project UIP-2020-02-5675, the European Regional Development Fund within the “Center of Excellence for Advanced Materials and Sensing Devices” (Grant No. KK.01.1.1.01.0001), and COST Action CA22154 - Data-driven Applications towards the Engineering of functional Materials: an Open Network (DAEMON) supported by COST (European Cooperation in Science and Technology). FG acknowledges financial support from the EMPEROR Project, CUP E93C24001040001, sponsored by the National Quantum Science and Technology Institute (Spoke 5), Grant No. PE00000023,

funded by the European Union – NextGeneration EU. CBM acknowledges funding from the Swiss National Science Foundation (SNSF) under grant number 217837. DN acknowledges financial support from the project “Podizanje znanstvene izvrsnosti Centra za napredne laserske tehnike (CALTboost)” financed by the European Union through the National Recovery and Resilience Plan 2021-2026 (NRPP).

AUTHOR CONTRIBUTIONS

LB has trained all ML models and performed all calculations with them under the supervision of IL. IL and LB

performed DFT calculations for the training dataset. DN performed TTM calculations. IL designed the study with input from FG and DN. FG and CBM helped in the construction of electronic free energy model. LB, DN and IL wrote the manuscript draft. All authors provided comments on the manuscript. All authors read and approved the final manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

-
- [1] K. Rossnagel, On the origin of charge-density waves in select layered transition-metal dichalcogenides, *Journal of Physics: Condensed Matter* **23**, 213001 (2011).
- [2] S. Manzeli, D. Ovchinnikov, D. Pasquier, O. V. Yazyev, and A. Kis, 2D transition metal dichalcogenides, *Nature Reviews Materials* **2**, 17033 (2017).
- [3] S. Foner and E. McNiff, Upper critical fields of layered superconducting NbSe₂ at low temperature, *Physics Letters A* **45**, 429 (1973).
- [4] J. Harper, T. Geballe, and F. Di Salvo, Heat capacity of 2H-NbSe₂ at the charge density wave transition, *Physics Letters A* **54**, 27 (1975).
- [5] K. Cho, M. Kończykowski, S. Teknowijoyo, M. A. Tanatar, J. Guss, P. B. Gartin, J. M. Wilde, A. Kreyssig, R. J. McQueeney, A. I. Goldman, V. Mishra, P. J. Hirschfeld, and R. Prozorov, Using controlled disorder to probe the interplay between charge order and superconductivity in NbSe₂, *Nature Communications* **9**, 2796 (2018).
- [6] O. Moulding, I. Osmond, F. Flicker, T. Muramatsu, and S. Friedemann, Absence of superconducting dome at the charge-density-wave quantum phase transition in 2H-NbSe₂, *Phys. Rev. Res.* **2**, 043392 (2020).
- [7] N. E. Staley, J. Wu, P. Eklund, Y. Liu, L. Li, and Z. Xu, Electric field effect on superconductivity in atomically thin flakes of NbSe₂, *Phys. Rev. B* **80**, 184505 (2009).
- [8] X. Xi, L. Zhao, Z. Wang, H. Berger, L. Forró, J. Shan, and K. F. Mak, Strongly enhanced charge-density-wave order in monolayer NbSe₂, *Nature Nanotechnology* **10**, 765–769 (2015).
- [9] M. M. Ugeda, A. J. Bradley, Y. Zhang, S. Onishi, Y. Chen, W. Ruan, C. Ojeda-Aristizabal, H. Ryu, M. T. Edmonds, H.-Z. Tsai, A. Riss, S.-K. Mo, D. Lee, A. Zettl, Z. Hussain, Z.-X. Shen, and M. F. Crommie, Characterization of collective ground states in single-layer NbSe₂, *Nature Physics* **12**, 92–97 (2015).
- [10] D. Lin, S. Li, J. Wen, H. Berger, L. Forró, H. Zhou, S. Jia, T. Taniguchi, K. Watanabe, X. Xi, and M. S. Bahramy, Patterns and driving forces of dimensionality-dependent charge density waves in 2H-type transition metal dichalcogenides, *Nature Communications* **11**, 2406 (2020).
- [11] D. Azoury, E. Baldini, A. Devarakonda, J. Li, S. Fang, P. Williams, R. Comin, J. Checkelsky, and N. Gedik, performed DFT calculations for the training dataset. DN performed TTM calculations. IL designed the study with input from FG and DN. FG and CBM helped in the construction of electronic free energy model. LB, DN and IL wrote the manuscript draft. All authors provided comments on the manuscript. All authors read and approved the final manuscript.
- [12] M. D. Johannes, I. I. Mazin, and C. A. Howells, Fermi-surface nesting and the origin of the charge-density wave in NbSe₂, *Phys. Rev. B* **73**, 205102 (2006).
- [13] X. Zhu, Y. Cao, J. Zhang, E. W. Plummer, and J. Guo, Classification of charge density waves based on their nature, *Proceedings of the National Academy of Sciences* **112**, 2367 (2015).
- [14] F. Flicker and J. van Wezel, Charge order from orbital-dependent coupling evidenced by NbSe₂, *Nature Commun.* **6**, 7034 (2015).
- [15] T. Valla, A. V. Fedorov, P. D. Johnson, P.-A. Glans, C. McGuinness, K. E. Smith, E. Y. Andrei, and H. Berger, Quasiparticle Spectra, Charge-Density Waves, Superconductivity, and Electron-Phonon Coupling in 2H-NbSe₂, *Phys. Rev. Lett.* **92**, 086401 (2004).
- [16] F. Weber, S. Rosenkranz, J.-P. Castellan, R. Osborn, R. Hott, R. Heid, K.-P. Bohnen, T. Egami, A. H. Said, and D. Reznik, Extended Phonon Collapse and the Origin of the Charge-Density Wave in 2H-NbSe₂, *Phys. Rev. Lett.* **107**, 107403 (2011).
- [17] W. L. McMillan, Microscopic model of charge-density waves in 2H-TaSe₂, *Phys. Rev. B* **16**, 643 (1977).
- [18] C. M. Varma and A. L. Simons, Strong-Coupling Theory of Charge-Density-Wave Transitions, *Phys. Rev. Lett.* **51**, 138 (1983).
- [19] M. Leroux, I. Errea, M. Le Tacon, S.-M. Souliou, G. Garbarino, L. Cario, A. Bosak, F. Mauri, M. Calandra, and P. Rodière, Strong anharmonicity induces quantum melting of charge density wave in 2H-NbSe₂ under pressure, *Phys. Rev. B* **92**, 140303 (2015).
- [20] R. Bianco, L. Monacelli, M. Calandra, F. Mauri, and I. Errea, Weak Dimensionality Dependence and Dominant Role of Ionic Fluctuations in the Charge-Density-Wave Transition of NbSe₂, *Phys. Rev. Lett.* **125**, 106101 (2020).
- [21] Y. Zheng, X. Jiang, X.-X. Xue, X. Yao, J. Zeng, K.-Q. Chen, E. Wang, and Y. Feng, Nuclear Quantum Effects on the Charge-Density Wave Transition in NbX₂ (X = S, Se), *Nano Letters* **22**, 1858 (2022).
- [22] S. Hellmann, T. Rohwer, M. Källäne, K. Hanff, C. Sohrt, A. Stange, A. Carr, M. Murnane, H. Kapteyn, L. Kipp, *et al.*, Time-domain classification of charge-density-wave

- insulators, *Nature communications* **3**, 1069 (2012).
- [23] M. Porer, U. Leierseder, J.-M. Ménard, H. Dachraoui, L. Mouchliadis, I. E. Perakis, U. Heinzmann, J. Demsar, K. Rossnagel, and R. Huber, Non-thermal separation of electronic and structural orders in a persisting charge density wave, *Nature Materials* **13**, 857 (2014).
- [24] M. R. Otto, J.-H. Pöhls, L. P. R. de Cotret, M. J. Stern, M. Sutton, and B. J. Siwick, Mechanisms of electron-phonon coupling unraveled in momentum and time: The case of soft phonons in TiSe_2 , *Science Advances* **7**, eabf2810 (2021).
- [25] C. Monney, M. Puppini, C. W. Nicholson, M. Hoesch, R. T. Chapman, E. Springate, H. Berger, A. Magrez, C. Cacho, R. Ernstorfer, and M. Wolf, Revealing the role of electrons and phonons in the ultrafast recovery of charge density wave correlations in $1T\text{-TiSe}_2$, *Phys. Rev. B* **94**, 165165 (2016).
- [26] D. Jérôme, T. M. Rice, and W. Kohn, Excitonic insulator, *Phys. Rev.* **158**, 462 (1967).
- [27] A. Anikin, R. D. Schaller, G. P. Wiederrecht, E. R. Margine, I. I. Mazin, and G. Karapetrov, Ultrafast dynamics in the high-symmetry and in the charge density wave phase of $2H\text{-NbSe}_2$, *Phys. Rev. B* **102**, 205139 (2020).
- [28] D. T. Payne, P. Barone, L. Benfatto, F. Parmigiani, and F. Cilento, Lattice contribution to the unconventional charge density wave transition in $2H\text{-NbSe}_2$: a non-equilibrium optical approach (2020), arXiv:2010.09826 [cond-mat.str-el].
- [29] M. Watanabe, T. Suzuki, T. Someya, Y. Ogawa, S. Michimae, M. Fujisawa, T. Kanai, J. Itatani, T. Saitoh, S. Shin, and K. Okazaki, Photo-Excitation Band-Structure Engineering of $2H\text{-NbSe}_2$ Probed by Time- and Angle-Resolved Photoemission Spectroscopy, *Journal of the Physical Society of Japan* **91**, 064703 (2022).
- [30] R. Venturini, A. Sarkar, P. Sutar, Z. Jagličić, Y. Vaskivskyi, E. Goreschnik, D. Mihailovic, and T. Mertelj, Unconventional photoinduced charge density wave dynamics in $2H\text{-NbSe}_2$, *Phys. Rev. B* **108**, 235160 (2023).
- [31] H. Zong, G. Pilania, X. Ding, G. Ackland, and T. Lookman, Developing an interatomic potential for martensitic phase transformations in zirconium by machine learning, *npj Computational Materials* **4** (2018).
- [32] C. Rosenbrock, K. Gubaev, A. Shapeev, L. Bartok-Partay, N. Bernstein, G. Csányi, and G. Hart, Machine-learned interatomic potentials for alloys and alloy phase diagrams, *npj Computational Materials* **7** (2021).
- [33] T. Kostiuchenko, F. Körmann, J. Neugebauer, and A. Shapeev, Impact of lattice relaxations on phase transitions in a high-entropy alloy studied by machine-learning potentials, *npj Computational Materials* **5** (2019).
- [34] T. Braeckvelt, R. Goeminne, S. Vandenhaute, S. Borgmans, T. Verstraelen, J. A. Steele, M. B. J. Roefiaers, J. Hofkens, S. M. J. Rogge, and V. Van Speybroeck, Accurately Determining the Phase Transition Temperature of CsPbI_3 via Random-Phase Approximation Calculations and Phase-Transferable Machine Learning Potentials, *Chemistry of Materials* **34**, 8561 (2022), <https://doi.org/10.1021/acs.chemmater.2c01508>.
- [35] B. Mladineo and I. Lončarić, Thermosalient Phase Transitions from Machine Learning Interatomic Potential, *Crystal Growth & Design* **24**, 8167 (2024), <https://doi.org/10.1021/acs.cgd.4c00905>.
- [36] C. Ben Mahmoud, F. Grasselli, and M. Ceriotti, Predicting hot-electron free energies from ground-state data, *Phys. Rev. B* **106**, L121116 (2022).
- [37] F. Giustino, *Materials Modelling Using Density Functional Theory: Properties and Predictions* (Oxford University Press, 2014).
- [38] G. Kresse and J. Hafner, Ab initio molecular dynamics for liquid metals, *Phys. Rev. B* **47**, 558 (1993).
- [39] G. Kresse and J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, *Computational Materials Science* **6**, 15 (1996).
- [40] G. Kresse and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Phys. Rev. B* **54**, 11169 (1996).
- [41] G. Kresse and J. Hafner, Norm-conserving and ultrasoft pseudopotentials for first-row and transition elements, *Journal of Physics: Condensed Matter* **6**, 8245 (1994).
- [42] G. Kresse and D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, *Phys. Rev. B* **59**, 1758 (1999).
- [43] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [44] H. C. Andersen, Molecular dynamics simulations at constant pressure and/or temperature, *The Journal of Chemical Physics* **72**, 2384 (1980), <https://pubs.aip.org/aip/jcp/article-pdf/72/4/2384/18920903/2384.1.online.pdf>.
- [45] I. Batatia, S. Batzner, D. P. Kovács, A. Musaelian, G. N. C. Simm, R. Drautz, C. Ortner, B. Kozinsky, and G. Csányi, The Design Space of E(3)-Equivariant Atom-Centered Interatomic Potentials (2022), arXiv:2205.06643.
- [46] I. Batatia, D. P. Kovacs, G. Simm, C. Ortner, and G. Csanyi, Mace: Higher order equivariant message passing neural networks for fast and accurate force fields, in *Advances in Neural Information Processing Systems*, Vol. 35, edited by S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (Curran Associates, Inc., 2022) pp. 11423–11436.
- [47] P. Izmailov, D. Podoprikin, T. Garipov, D. Vetrov, and A. G. Wilson, Averaging Weights Leads to Wider Optima and Better Generalization (2019), arXiv:1803.05407 [cs.LG].
- [48] B. Athiwaratkun, M. Finzi, P. Izmailov, and A. G. Wilson, There Are Many Consistent Explanations of Unlabeled Data: Why You Should Average (2019), arXiv:1806.05594 [cs.LG].
- [49] D. P. Kovács, I. Batatia, E. S. Arany, and G. Csányi, Evaluation of the MACE force field architecture: From medicinal chemistry to materials science, *The Journal of Chemical Physics* **159**, 044118 (2023).
- [50] A. Togo, L. Chaput, T. Tadano, and I. Tanaka, Implementation strategies in phonopy and phono3py, *Journal of Physics: Condensed Matter* **35**, 353001 (2023).
- [51] V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti, and G. Csányi, Gaussian Process Regression for Materials and Molecules, *Chemical Reviews* **121**, 10073 (2021), pMID: 34398616, <https://doi.org/10.1021/acs.chemrev.1c00022>.
- [52] F. Musil, M. Veit, A. Goscinski, G. Fraux, M. J. Willatt, M. Stricker, T. Junge, and M. Ceriotti, Efficient implementation of atom-density representations, *The Journal*

- of chemical physics **154** **11**, 114109 (2021).
- [53] A. P. Bartók, R. Kondor, and G. Csányi, On representing chemical environments, *Phys. Rev. B* **87**, 184115 (2013).
- [54] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, *Optuna: A Next-generation Hyperparameter Optimization Framework* (2019), arXiv:1907.10902 [cs.LG].
- [55] C. Ben Mahmoud, A. Anelli, G. Csányi, and M. Ceriotti, Learning the electronic density of states in condensed matter, *Phys. Rev. B* **102**, 235130 (2020).
- [56] W. B. How, S. Chong, F. Grasselli, K. K. Huguenin-Dumittan, and M. Ceriotti, Adaptive energy reference for machine-learning models of the electronic density of states, *Phys. Rev. Mater.* **9**, 013802 (2025).
- [57] Y. Eldar, M. Lindenbaum, M. Porat, and Y. Zeevi, The farthest point strategy for progressive image sampling, *IEEE Transactions on Image Processing* **6**, 1305 (1997).
- [58] C. Ben Mahmoud, *Machine-learning the electronic density of states: electronic properties without quantum mechanics*, Ph.D. thesis, EPFL, Lausanne (2023).
- [59] R. Bianco, I. Errea, L. Paulatto, M. Calandra, and F. Mauri, Second-order structural phase transitions, free energy curvature, and temperature-dependent anharmonic phonons in the self-consistent harmonic approximation: Theory and stochastic implementation, *Phys. Rev. B* **96**, 014111 (2017).
- [60] L. Monacelli, R. Bianco, M. Cherubini, M. Calandra, I. Errea, and F. Mauri, The stochastic self-consistent harmonic approximation: calculating vibrational properties of materials with full quantum and anharmonic effects, *Journal of Physics: Condensed Matter* **33**, 363001 (2021).
- [61] J. S. Zhou, L. Monacelli, R. Bianco, I. Errea, F. Mauri, and M. Calandra, Anharmonicity and Doping Melt the Charge Density Wave in Single-Layer TiSe₂, *Nano Letters* **20**, 4809 (2020), pMID: 32496779, <https://doi.org/10.1021/acs.nanolett.0c00597>.
- [62] S. R. Bahn and K. W. Jacobsen, An object-oriented scripting interface to a legacy electronic structure code, *Comput. Sci. Eng.* **4**, 56 (2002).
- [63] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dulák, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, and K. W. Jacobsen, The atomic simulation environment—a Python library for working with atoms, *Journal of Physics: Condensed Matter* **29**, 273002 (2017).
- [64] D. Novko, Z. Torbatian, and I. Lončarić, Electron correlations rule the phonon-driven instability in single-layer TiSe₂, *Phys. Rev. B* **106**, 245108 (2022).
- [65] M. Hellgren, J. Baima, R. Bianco, M. Calandra, F. Mauri, and L. Wirtz, Critical role of the exchange interaction for the electronic structure and charge-density-wave formation in tise₂, *Phys. Rev. Lett.* **119**, 176401 (2017).
- [66] A. Schobert, J. Berges, E. G. C. P. van Loon, M. A. Sentef, S. Brener, M. Rossi, and T. O. Wehling, Ab initio electron-lattice downfolding: Potential energy landscapes, anharmonicity, and molecular dynamics in charge density wave materials, *SciPost Phys.* **16**, 046 (2024).
- [67] Y. Yoshida and K. Motizuki, Electron Lattice Interaction and Lattice Instability of 1T-TiSe₂, *Journal of the Physical Society of Japan* **49**, 898 (1980).
- [68] H. Yoshiyama, Y. Takaoka, N. Suzuki, and K. Motizuki, Effects on lattice fluctuations on the charge-density-wave transition in transition-metal dichalcogenides, *Journal of Physics C: Solid State Physics* **19**, 5591 (1986).
- [69] D. Novko, F. Caruso, C. Draxl, and E. Cappelluti, Ultrafast Hot Phonon Dynamics in MgB₂ Driven by Anisotropic Electron-Phonon Coupling, *Phys. Rev. Lett.* **124**, 077001 (2020).
- [70] F. Caruso and D. Novko, Ultrafast dynamics of electrons and phonons: from the two-temperature model to the time-dependent Boltzmann equation, *Advances in Physics: X* **7**, 2095925 (2022).

Supplemental Material:
**Machine learning model for efficient nonthermal tuning of the
charge density wave in monolayer NbSe₂**

Luka Beni¹, Federico Grasselli^{2,3}, Chiheb Ben
Mahmoud⁴, Dino Novko^{5,*} and Ivor Lončarić^{1,†}

¹*Ruđer Bošković Institute, Zagreb, Croatia*

²*Dipartimento di Scienze Fisiche, Informatiche e Matematiche,
Università degli Studi di Modena e Reggio Emilia, Modena, Italy*

³*CNR NANO S3, Modena, Italy*

⁴*Inorganic Chemistry Laboratory, Department of Chemistry,
University of Oxford, Oxford OX1 3QR, United Kingdom*

⁵*Centre for Advanced Laser Techniques, Institute of Physics, Zagreb, Croatia*

*dino.novko@gmail.com, †ivor.loncaric@gmail.com

S1. APPROXIMATION FOR THE HELMHOLTZ ELECTRONIC FREE ENERGY AND THE IONIC FORCES

We report the derivation of the approximation used for simulating the finite electronic temperature effects. The idea is to use the 0 K density functional theory data to construct an approximation for the Helmholtz electronic free energy to be able to efficiently simulate finite electronic temperature effects. It is a modified derivation of the approximation developed in Ref. [S1]. Firstly, we define the notation we will use. We start by considering an atomic system with N_{ion} ions and N_{el} electrons, where $\underline{\mathbf{R}} = \{\mathbf{R}_1, \dots, \mathbf{R}_{N_{ion}}\}$ represents the set of ionic coordinates and we let T_{el} to be the electronic temperature of the system. Fermi-Dirac distribution is given by

$$f(x) = \frac{1}{e^x + 1} . \quad (\text{S1})$$

In the following derivations, the temperatures in the superscripts indicate to which temperatures do the eigenvalues that are used to construct desired quantities correspond, and the temperatures in parentheses indicate at which temperatures the desired quantities are evaluated.

We define occupation numbers as

$$f_i^{T_1}(T_2) \stackrel{(\text{S1})}{=} f\left(\frac{\epsilon_i^{T_1}(\underline{\mathbf{R}}) - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) , \quad (\text{S2})$$

where k_B is the Boltzmann constant. In (S2) i -th occupation number $f_i^{T_1}(T_2)$ evaluated at electronic temperature T_2 is determined by the i -th Kohn-Sham eigenenergy $\epsilon_i^{T_1}(\underline{\mathbf{R}})$ corresponding to T_1 and the chemical potential (Fermi level) $\mu^{T_1}(\underline{\mathbf{R}}; T_2)$ evaluated at T_2 and constructed out of the T_1 eigenenergies. The chemical potential is determined by the constraint of conservation of the number of electrons in the system

$$N_{el} = \sum_i f_i^{T_1}(T_2) \quad \forall T_1, T_2 , \quad (\text{S3})$$

where the summation over i represents the summation over the spin projections s , bands n and k-points \mathbf{k} . Furthermore, in the limit $T_2 \rightarrow 0^+$ we have

$$\mu^{T_1}(\underline{\mathbf{R}}; 0) = \epsilon_F^{T_1}(\underline{\mathbf{R}}) , \quad (\text{S4})$$

where we call $\epsilon_F^{T_1}(\underline{\mathbf{R}})$ the Fermi energy of the system at the electronic temperature T_1 , which in the limit $T_1 \rightarrow 0^+$ becomes the real Fermi energy of the system $\epsilon_F^0(\underline{\mathbf{R}})$.

Next, we define the electronic density of states (eDOS) at a electronic temperature T as

$$g^T(\underline{\mathbf{R}}; \epsilon) = \sum_i \delta(\epsilon - \epsilon_i^T(\underline{\mathbf{R}})) , \quad (\text{S5})$$

where $\delta(x)$ is the Dirac delta distribution. Using (S5) we have the standard correspondence

$$\sum_i \rightarrow \int g^T(\underline{\mathbf{R}}; \epsilon) d\epsilon . \quad (\text{S6})$$

Using (S6) the equation (S3) can be written as

$$N_{el} = \int g^{T_1}(\underline{\mathbf{R}}; \epsilon) f\left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) d\epsilon . \quad (\text{S7})$$

The fermionic entropy in this notation is defined as

$$S^{T_1}(\underline{\mathbf{R}}; T_2) = -k_B \sum_i [f_i^{T_1}(T_2) \ln(f_i^{T_1}(T_2)) + (1 - f_i^{T_1}(T_2)) \ln(1 - f_i^{T_1}(T_2))] . \quad (\text{S8})$$

For the electronic density, we have

$$\rho^{T_1}(\mathbf{r}; T_2) = \sum_i \rho_i^{T_1}(\mathbf{r}; T_2) = \sum_i f_i^{T_1}(T_2) |\phi_i^{T_1}(\mathbf{r})|^2 , \quad (\text{S9})$$

where $\phi_i^{T_1}(\mathbf{r})$ is the i -th Kohn-Sham orbital at the electronic temperature T_1 and $\mathbf{r} = r_x \hat{\mathbf{x}} + r_y \hat{\mathbf{y}} + r_z \hat{\mathbf{z}}$.

The Helmholtz electronic free energy of a system at the electronic temperature T_{el} , with fixed number of electrons N_{el} and fixed volume V of the system is given by [S2]

$$F(\underline{\mathbf{R}}; N_{el}, V, T_{el}) \equiv F(\underline{\mathbf{R}}; T_{el}) = E^{T_{el}}(\underline{\mathbf{R}}; T_{el}) - T_{el} S^{T_{el}}(\underline{\mathbf{R}}; T_{el}) , \quad (\text{S10})$$

where $E^{T_{el}}(\underline{\mathbf{R}}; T_{el})$ is the standard DFT energy defined as [S3]

$$E^{T_{el}}(\underline{\mathbf{R}}; T_{el}) = E_b^{T_{el}}(\underline{\mathbf{R}}; T_{el}) - E_{dc}^{T_{el}}(\underline{\mathbf{R}}; T_{el}) + E_{nn}(\underline{\mathbf{R}}) . \quad (\text{S11})$$

In (S11) the first, so-called *band-energy*, term, is given by

$$E_b^{T_{el}}(\underline{\mathbf{R}}; T_{el}) = \sum_i f_i^{T_{el}}(T_{el}) \epsilon_i^{T_{el}}(\underline{\mathbf{R}}) . \quad (\text{S12})$$

The second term, the so-called *double-counting* term, is given by

$$\begin{aligned} E_{dc}^{T_{el}}(\underline{\mathbf{R}}; T_{el}) &= \frac{1}{2} \iint \frac{\rho^{T_{el}}(\mathbf{r}; T_{el}) \rho^{T_{el}}(\mathbf{r}'; T_{el})}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' - E_{XC}[\rho^{T_{el}}(\mathbf{r}; T_{el})](T_{el}) \\ &+ \int V_{XC}[\rho^{T_{el}}(\mathbf{r}; T_{el})](\mathbf{r}; T_{el}) \rho^{T_{el}}(\mathbf{r}; T_{el}) d\mathbf{r} , \end{aligned} \quad (\text{S13})$$

where $E_{XC}[\rho^{T_{el}}(\mathbf{r}; T_{el})](T_{el})$ is the so-called *exchange-correlation* (XC) functional and

$$V_{XC}[\rho^{T_{el}}(\mathbf{r}; T_{el})](\mathbf{r}; T_{el}) = \frac{\delta E_{XC}[\rho^{T_{el}}(\mathbf{r}; T_{el})](T_{el})}{\delta \rho^{T_{el}}(\mathbf{r}; T_{el})} \quad (\text{S14})$$

is the XC potential. Finally, the last term in (S11) is the electrostatic energy between the ions and is independent of T_{el} .

We start by taking the difference between the Helmholtz energies (S10) at the two different temperatures, where some finite reference temperature T_1 is taken to make the derivations more consistent. All of the results will be based at $T_1 \rightarrow 0^+$ and at the end we will let $T_1 \rightarrow 0^+$ everywhere. So, we start with

$$\begin{aligned} F(\underline{\mathbf{R}}; T_2) - F(\underline{\mathbf{R}}; T_1) &= E_b^{T_2}(\underline{\mathbf{R}}; T_2) - E_b^{T_1}(\underline{\mathbf{R}}; T_1) - E_{dc}^{T_2}(\underline{\mathbf{R}}; T_2) + E_{dc}^{T_1}(\underline{\mathbf{R}}; T_1) \\ &+ \underline{E}_{nn}(\underline{\mathbf{R}}) - \underline{E}_{nn}(\underline{\mathbf{R}}) - T_2 S^{T_2}(\underline{\mathbf{R}}; T_2) + T_1 S^{T_1}(\underline{\mathbf{R}}; T_1) . \end{aligned} \quad (\text{S15})$$

Now, the approximation consists of using a similar idea as in the Ref. [S4]. However, keeping in mind that the results from [S4] are obtained for $T_{el} = 0$ K. The idea there was that even at $T_{el} = 0$ K one can introduce occupational numbers like in (S9) as a mathematical tool that enables to calculate variations with respect to occupational numbers, and at the end of the calculations one lets those occupational numbers follow the Fermi-Dirac distribution at $T_{el} = 0$ K, or in other words, occupational numbers can take only the values of 0 and 1. The main result of that paper is the so-called Janak's theorem, which states that from (S11) at $T_{el} = 0$ K one has

$$\left. \frac{\delta E^{T_{el}}(\underline{\mathbf{R}}; T_{el})}{\delta f_i^{T_{el}}(T_{el})} \right|_{T_{el}=0} = \epsilon_i^0(\underline{\mathbf{R}}) . \quad (\text{S16})$$

It is worth mentioning that Janak's theorem has its thermal version [S5], but this is not what this approximation aims to use. The variation with respect to occupational numbers can be traced to the following idea

$$\psi_{\underline{\mathbf{R}}}(\underline{\mathbf{r}}; 0) \iff \rho^0(\mathbf{r}; 0) \stackrel{(\text{S9})}{\iff} \{f_i^0(0)\}, \{\phi_i^0(\mathbf{r})\} , \quad (\text{S17})$$

where $\psi_{\underline{\mathbf{R}}}(\underline{\mathbf{r}}; 0)$ is the exact electronic wave function at 0 K. The first equivalence in (S17) is due to the Hohenberg-Kohn theorem [S6]. In (S17) we want to emphasize that once we have density functional description of the quantum many-body problem at 0 K given by the Hohenberg-Kohn theorem, we can formally describe it using the occupational numbers and Kohn-Sham orbitals, as implied by (S9), where we impose that those occupational numbers

obey the Fermi-Dirac distribution at 0 K. This is the reasoning behind taking the variations with respect to occupational numbers in (S16). It also allows one to treat the variations with respect to occupational numbers and Kohn-Sham orbitals independently. So, the idea is to treat the occupational numbers introduced in [S4] as objects with physical meaning to capture the effects of changing the electronic temperature and to construct an approximation for the Helmholtz electronic free energy. Using only the information we have at $T_1 \rightarrow 0^+$, we perform the following set of approximations

$$\rho^{T_2}(\mathbf{r}; T_2) \approx \rho^{T_1}(\mathbf{r}; T_1) + \sum_i \left. \frac{\delta \rho^{T_1}(\mathbf{r}; T)}{\delta f_i^{T_1}(T)} \right|_{T=T_1} (f_i^{T_1}(T_2) - f_i^{T_1}(T_1)) , \quad (\text{S18})$$

$$E_b^{T_2}(\underline{\mathbf{R}}; T_2) \approx E_b^{T_1}(\underline{\mathbf{R}}; T_1) + \sum_i \left. \frac{\delta E_b^{T_1}(\underline{\mathbf{R}}; T)}{\delta f_i^{T_1}(T)} \right|_{T=T_1} (f_i^{T_1}(T_2) - f_i^{T_1}(T_1)) , \quad (\text{S19})$$

$$E_{dc}^{T_2}(\underline{\mathbf{R}}; T_2) \approx E_{dc}^{T_1}(\underline{\mathbf{R}}; T_1) + \sum_i \left. \frac{\delta E_{dc}^{T_1}(\underline{\mathbf{R}}; T)}{\delta f_i^{T_1}(T)} \right|_{T=T_1} (f_i^{T_1}(T_2) - f_i^{T_1}(T_1)) , \quad (\text{S20})$$

$$T_2 S^{T_2}(\underline{\mathbf{R}}; T_2) \approx T_2 S^{T_1}(\underline{\mathbf{R}}; T_2) . \quad (\text{S21})$$

The approximations given by (S18)-(S21) tell us that we approximate the quantities in (S18)-(S20) using the perturbations in occupational numbers caused by change in temperature.

From (S17) and (S18) we have

$$\begin{aligned} \rho^{T_2}(\mathbf{r}; T_2) &\approx \rho^{T_1}(\mathbf{r}; T_1) + \sum_i \left. \frac{\delta \rho^{T_1}(\mathbf{r}; T)}{\delta f_i^{T_1}(T)} \right|_{T=T_1} (f_i^{T_1}(T_2) - f_i^{T_1}(T_1)) \\ &= \rho^{T_1}(\mathbf{r}; T_1) + \sum_i (f_i^{T_1}(T_2) - f_i^{T_1}(T_1)) |\phi_i^{T_1}(\mathbf{r})|^2 \\ &= \underbrace{\rho^{T_1}(\mathbf{r}; T_1)}_{\substack{\text{(S9)} \\ = \rho^{T_1}(\mathbf{r}; T_2)}} + \sum_i f_i^{T_1}(T_2) |\phi_i^{T_1}(\mathbf{r})|^2 - \underbrace{\sum_i f_i^{T_1}(T_1) |\phi_i^{T_1}(\mathbf{r})|^2}_{\substack{\text{(S9)} \\ = \rho^{T_1}(\mathbf{r}; T_1)}} \\ &\implies \rho^{T_2}(\mathbf{r}; T_2) \approx \rho^{T_1}(\mathbf{r}; T_2) . \end{aligned} \quad (\text{S22})$$

Now, using (S18)-(S21) in (S15) we have

$$\begin{aligned} F(\underline{\mathbf{R}}; T_2) - F(\underline{\mathbf{R}}; T_1) &\approx \sum_i \left. \frac{\delta E_b^{T_1}(\underline{\mathbf{R}}; T)}{\delta f_i^{T_1}(T)} \right|_{T=T_1} (f_i^{T_1}(T_2) - f_i^{T_1}(T_1)) \\ &- \sum_i \left. \frac{\delta E_{dc}^{T_1}(\underline{\mathbf{R}}; T)}{\delta f_i^{T_1}(T)} \right|_{T=T_1} (f_i^{T_1}(T_2) - f_i^{T_1}(T_1)) - T_2 S^{T_1}(\underline{\mathbf{R}}; T_2) + T_1 S^{T_1}(\underline{\mathbf{R}}; T_1) . \end{aligned} \quad (\text{S23})$$

Also, we know that

$$\frac{\delta E_{nn}(\underline{\mathbf{R}})}{\delta f_i^{T_1}(T)} = 0 . \quad (\text{S24})$$

Using (S24) in (S23) we have

$$\begin{aligned} F(\underline{\mathbf{R}}; T_2) - F(\underline{\mathbf{R}}; T_1) &\approx \sum_i \frac{\delta (E_b^{T_1}(\underline{\mathbf{R}}; T) - E_{dc}^{T_1}(\underline{\mathbf{R}}; T) + E_{nn}(\underline{\mathbf{R}}))}{\delta f_i^{T_1}(T)} \Bigg|_{T=T_1} (f_i^{T_1}(T_2) - f_i^{T_1}(T_1)) \\ &\quad - T_2 S^{T_1}(\underline{\mathbf{R}}; T_2) + T_1 S^{T_1}(\underline{\mathbf{R}}; T_1) \\ &\stackrel{(\text{S11})}{=} \sum_i \frac{\delta E^{T_1}(\underline{\mathbf{R}}; T)}{\delta f_i^{T_1}(T)} \Bigg|_{T=T_1} (f_i^{T_1}(T_2) - f_i^{T_1}(T_1)) - T_2 S^{T_1}(\underline{\mathbf{R}}; T_2) + T_1 S^{T_1}(\underline{\mathbf{R}}; T_1) \\ &\stackrel{(\text{S16})}{=} \sum_i \epsilon_i^{T_1}(\underline{\mathbf{R}}) (f_i^{T_1}(T_2) - f_i^{T_1}(T_1)) - T_2 S^{T_1}(\underline{\mathbf{R}}; T_2) + T_1 S^{T_1}(\underline{\mathbf{R}}; T_1) . \end{aligned} \quad (\text{S25})$$

In conclusion, using (S10) and (S25) we approximate the Helmholtz electronic free energy evaluated at a finite electronic temperature T_2 as

$$F(\underline{\mathbf{R}}; T_2) \approx E^{T_1}(\underline{\mathbf{R}}; T_1) + \sum_i \epsilon_i^{T_1}(\underline{\mathbf{R}}) (f_i^{T_1}(T_2) - f_i^{T_1}(T_1)) - T_2 S^{T_1}(\underline{\mathbf{R}}; T_2) . \quad (\text{S26})$$

Furthermore, as we have already mentioned, T_1 was used just to have a more mathematically consistent derivation, so we can finally write everything using $T_1 \rightarrow 0^+$, where we also let $T_2 = T_{el}$

$$F(\underline{\mathbf{R}}; T_{el}) \approx E^0(\underline{\mathbf{R}}; 0) + \sum_i \epsilon_i^0(\underline{\mathbf{R}}) (f_i^0(T_{el}) - f_i^0(0)) - T_{el} S^0(\underline{\mathbf{R}}; T_{el}) . \quad (\text{S27})$$

We can write (S26) as

$$F(\underline{\mathbf{R}}; T_2) \approx E^{T_1}(\underline{\mathbf{R}}; T_1) + \Delta F(\underline{\mathbf{R}}; T_1, T_2) , \quad (\text{S28})$$

where we have defined

$$\Delta F(\underline{\mathbf{R}}; T_1, T_2) = \sum_i \epsilon_i^{T_1}(\underline{\mathbf{R}}) (f_i^{T_1}(T_2) - f_i^{T_1}(T_1)) - T_2 S^{T_1}(\underline{\mathbf{R}}; T_2) . \quad (\text{S29})$$

Or, using $T_1 \rightarrow 0^+$ and the definitions from (S27) we have

$$F(\underline{\mathbf{R}}; T_{el}) \approx E^0(\underline{\mathbf{R}}; 0) + \Delta F(\underline{\mathbf{R}}; 0, T_{el}) . \quad (\text{S30})$$

From (S6), (S8), and (S26) we have

$$\begin{aligned} F(\underline{\mathbf{R}}; T_2) &\approx E^{T_1}(\underline{\mathbf{R}}; T_1) + \int \epsilon g^{T_1}(\underline{\mathbf{R}}; \epsilon) \left[f \left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2} \right) - f \left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_1)}{k_B T_1} \right) \right] d\epsilon \\ &\quad + k_B T_2 \int g^{T_1}(\underline{\mathbf{R}}; \epsilon) s \left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2} \right) d\epsilon , \end{aligned} \quad (\text{S31})$$

where $s(x)$ is defined as

$$s(x) = f(x) \ln(f(x)) + (1 - f(x)) \ln(1 - f(x)) . \quad (\text{S32})$$

Or, using $T_1 \rightarrow 0^+$ and the definitions from (S27), along with (S4), we have

$$\begin{aligned} F(\underline{\mathbf{R}}; T_{el}) \approx E^0(\underline{\mathbf{R}}; 0) + \int \epsilon g^0(\underline{\mathbf{R}}; \epsilon) \left[f\left(\frac{\epsilon - \mu^0(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) - \theta(\epsilon_F^0(\underline{\mathbf{R}}) - \epsilon) \right] d\epsilon \\ + k_B T_2 \int g^0(\underline{\mathbf{R}}; \epsilon) s\left(\frac{\epsilon - \mu^0(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) d\epsilon , \end{aligned} \quad (\text{S33})$$

where $f^0(0) = \theta(\epsilon_F^0(\underline{\mathbf{R}}) - \epsilon)$ was used.

In essence, this approximation exploits the idea from [S4] where the occupation numbers are let to be fractional and have been used as a mathematical tool to derive Janak's theorem. But in this approximation they are treated as valid physical objects and are used to simulate changes in temperature and to construct an approximation for the Helmholtz electronic free energy. Inspecting the presented derivations, one can say that the approximation developed is based on the following approximation $f_i^{T_2}(T_2) \approx f_i^{T_1}(T_2)$, which, as already said, means that occupational numbers introduced in the Janak's theorem were used to simulate the effects of change in temperature. In conclusion, thermal effects are included by extrapolating the 0 K density functional theory data to finite temperatures using the constructed approximation for the Helmholtz electronic free energy. The main difference between this derivation of (S33) and the one from [S1] is in the fact that here we did not use $f_i^{T_2}(T_2)$ quantities explicitly, for which we do not have any information in this model.

Now, we can derive the expression for the force on the n -th ion $\mathbf{F}_n(\underline{\mathbf{R}}; T_{el})$ by taking the negative gradient with respect to \mathbf{R}_n of (S31)

$$\begin{aligned} \mathbf{F}_n(\underline{\mathbf{R}}; T_2) = -\nabla_{\mathbf{R}_n} F(\underline{\mathbf{R}}; T_2) \approx -\nabla_{\mathbf{R}_n} E^{T_1}(\underline{\mathbf{R}}; T_1) \\ - \nabla_{\mathbf{R}_n} \int \epsilon g^{T_1}(\underline{\mathbf{R}}; \epsilon) \left[f\left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) - f\left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_1)}{k_B T_1}\right) \right] d\epsilon \\ - k_B T_2 \nabla_{\mathbf{R}_n} \int g^{T_1}(\underline{\mathbf{R}}; \epsilon) s\left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) d\epsilon . \end{aligned} \quad (\text{S34})$$

We start by calculating

$$\begin{aligned} \nabla_{\mathbf{R}_n} \int \epsilon g^{T_1}(\underline{\mathbf{R}}; \epsilon) f\left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) d\epsilon = \int d\epsilon \epsilon f\left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) \nabla_{\mathbf{R}_n} g^{T_1}(\underline{\mathbf{R}}; \epsilon) \\ + \int d\epsilon \epsilon g^{T_1}(\underline{\mathbf{R}}; \epsilon) \frac{\partial f}{\partial \mu^{T_1}}\left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) \nabla_{\mathbf{R}_n} \mu^{T_1}(\underline{\mathbf{R}}; T_2) , \end{aligned} \quad (\text{S35})$$

now we need an expression for $\nabla_{\mathbf{R}_n} \mu^{T_1}(\underline{\mathbf{R}}; T_2)$, which we get by taking the gradient $\nabla_{\mathbf{R}_n}$ of the expression (S7)

$$\begin{aligned}
0 &= \nabla_{\mathbf{R}_n} \int g^{T_1}(\underline{\mathbf{R}}; \epsilon) f\left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) d\epsilon = \int d\epsilon f\left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) \nabla_{\mathbf{R}_n} g^0(\underline{\mathbf{R}}; \epsilon) \\
&\quad + \nabla_{\mathbf{R}_n} \mu^{T_1}(\underline{\mathbf{R}}; T_2) \int d\epsilon g^{T_1}(\underline{\mathbf{R}}; \epsilon) \frac{\partial f}{\partial \mu^{T_1}}\left(\frac{\epsilon - \mu^0(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) \\
\implies \nabla_{\mathbf{R}_n} \mu^{T_1}(\underline{\mathbf{R}}; T_2) &= -\frac{\int d\epsilon f\left(\frac{\epsilon - \mu^0(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) \nabla_{\mathbf{R}_n} g^{T_1}(\underline{\mathbf{R}}; \epsilon)}{\int d\epsilon g^{T_1}(\underline{\mathbf{R}}; \epsilon) \frac{\partial f}{\partial \mu^{T_1}}\left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right)}. \tag{S36}
\end{aligned}$$

Now, by using (S36) in (S35) we have

$$\begin{aligned}
\nabla_{\mathbf{R}_n} \int \epsilon g^{T_1}(\underline{\mathbf{R}}; \epsilon) f\left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) d\epsilon &= \int d\epsilon \epsilon f\left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) \nabla_{\mathbf{R}_n} g^{T_1}(\underline{\mathbf{R}}; \epsilon) \\
&\quad + \underbrace{\frac{\int d\epsilon' \epsilon' g^{T_1}(\underline{\mathbf{R}}; \epsilon') \frac{\partial f}{\partial \mu^{T_1}}\left(\frac{\epsilon' - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right)}{\int d\epsilon' g^{T_1}(\underline{\mathbf{R}}; \epsilon') \frac{\partial f}{\partial \mu^{T_1}}\left(\frac{\epsilon' - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right)}}_{\equiv \Sigma^{T_1}(\underline{\mathbf{R}}; T_2)} \int d\epsilon f\left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) \nabla_{\mathbf{R}_n} g^{T_1}(\underline{\mathbf{R}}; \epsilon) \\
\implies \nabla_{\mathbf{R}_n} \int \epsilon g^{T_1}(\underline{\mathbf{R}}; \epsilon) f\left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) d\epsilon &= \int d\epsilon (\epsilon - \Sigma^{T_1}(\underline{\mathbf{R}}; T_2)) \\
&\quad \times f\left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right) \nabla_{\mathbf{R}_n} g^{T_1}(\underline{\mathbf{R}}; \epsilon). \tag{S37}
\end{aligned}$$

In (S37) we have defined

$$\Sigma^{T_1}(\underline{\mathbf{R}}; T_2) = \frac{\int d\epsilon \epsilon g^{T_1}(\underline{\mathbf{R}}; \epsilon) \frac{\partial f}{\partial \mu^{T_1}}\left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right)}{\int d\epsilon g^{T_1}(\underline{\mathbf{R}}; \epsilon) \frac{\partial f}{\partial \mu^{T_1}}\left(\frac{\epsilon - \mu^{T_1}(\underline{\mathbf{R}}; T_2)}{k_B T_2}\right)} \tag{S38}$$

and we call it the average shift term. In the limit $T_2 \rightarrow 0^+$, using (S4) and the fact $f^{T_1}(0) = \theta(\epsilon_F^{T_1}(\underline{\mathbf{R}}) - \epsilon)$, we have

$$\begin{aligned}
\Sigma^{T_1}(\underline{\mathbf{R}}; 0) &= \frac{\int d\epsilon \epsilon g^{T_1}(\underline{\mathbf{R}}; \epsilon) \frac{\partial \theta(\epsilon_F^{T_1}(\underline{\mathbf{R}}) - \epsilon)}{\partial \epsilon_F^{T_1}(\underline{\mathbf{R}})}}{\int d\epsilon g^{T_1}(\underline{\mathbf{R}}; \epsilon) \frac{\partial \theta(\epsilon_F^{T_1}(\underline{\mathbf{R}}) - \epsilon)}{\partial \epsilon_F^{T_1}(\underline{\mathbf{R}})}} = \frac{\int d\epsilon \epsilon g^{T_1}(\underline{\mathbf{R}}; \epsilon) \delta(\epsilon_F^{T_1}(\underline{\mathbf{R}}) - \epsilon)}{\int d\epsilon g^{T_1}(\underline{\mathbf{R}}; \epsilon) \delta(\epsilon_F^{T_1}(\underline{\mathbf{R}}) - \epsilon)} = \epsilon_F^{T_1}(\underline{\mathbf{R}}) \\
\implies \Sigma^{T_1}(\underline{\mathbf{R}}; 0) &= \epsilon_F^{T_1}(\underline{\mathbf{R}}), \tag{S39}
\end{aligned}$$

which in the limit $T_1 \rightarrow 0^+$ becomes

$$\Sigma^0(\underline{\mathbf{R}}; 0) = \epsilon_F^0(\underline{\mathbf{R}}). \tag{S40}$$

Now, we can calculate the following

$$\begin{aligned} \nabla_{\mathbf{R}_n} \int g^{T_1}(\mathbf{R}; \epsilon) s \left(\frac{\epsilon - \mu^{T_1}(\mathbf{R}; T_2)}{k_B T_2} \right) d\epsilon &= \int d\epsilon s \left(\frac{\epsilon - \mu^{T_1}(\mathbf{R}; T_2)}{k_B T_2} \right) \nabla_{\mathbf{R}_n} g^{T_1}(\mathbf{R}; \epsilon) \\ &+ \int d\epsilon g^{T_1}(\mathbf{R}; \epsilon) \frac{\partial s}{\partial \mu^{T_1}} \left(\frac{\epsilon - \mu^{T_1}(\mathbf{R}; T_2)}{k_B T_2} \right) \nabla_{\mathbf{R}_n} \mu^{T_1}(\mathbf{R}; T_{el}) . \end{aligned} \quad (\text{S41})$$

Furthermore. we have

$$\begin{aligned} \frac{\partial s}{\partial \mu^{T_1}} \left(\frac{\epsilon - \mu^{T_1}(\mathbf{R}; T_2)}{k_B T_2} \right) &= \frac{\partial s}{\partial f} \frac{\partial f}{\partial \mu^{T_1}} \left(\frac{\epsilon - \mu^{T_1}(\mathbf{R}; T_2)}{k_B T_2} \right) \\ \stackrel{(\text{S32})}{=} \left(\ln \left(\frac{f}{1-f} \right) \right) \frac{\partial f}{\partial \mu^{T_1}} \left(\frac{\epsilon - \mu^{T_1}(\mathbf{R}; T_2)}{k_B T_2} \right) &\stackrel{(\text{S2})}{=} - \frac{\epsilon - \mu^{T_1}(\mathbf{R}; T_2)}{k_B T_2} \frac{\partial f}{\partial \mu^{T_1}} \left(\frac{\epsilon - \mu^{T_1}(\mathbf{R}; T_2)}{k_B T_2} \right) . \end{aligned} \quad (\text{S42})$$

Now, by inserting (S42) into (S41), and using (S36) and (S38) we have

$$\begin{aligned} \nabla_{\mathbf{R}_n} \int g^{T_1}(\mathbf{R}; \epsilon) s \left(\frac{\epsilon - \mu^{T_1}(\mathbf{R}; T_2)}{k_B T_2} \right) d\epsilon &= \int \nabla_{\mathbf{R}_n} g^{T_1}(\mathbf{R}; \epsilon) \left[s \left(\frac{\epsilon - \mu^{T_1}(\mathbf{R}; T_2)}{k_B T_2} \right) \right. \\ &\left. + \frac{\Sigma^{T_1}(\mathbf{R}; T_2) - \mu^{T_1}(\mathbf{R}; T_2)}{k_B T_2} f \left(\frac{\epsilon - \mu^{T_1}(\mathbf{R}; T_2)}{k_B T_2} \right) \right] d\epsilon . \end{aligned} \quad (\text{S43})$$

Finally, by using (S34), (S37) and (S43) we arrive at

$$\begin{aligned} \mathbf{F}_n(\mathbf{R}; T_2) &\approx -\nabla_{\mathbf{R}_n} E^{T_1}(\mathbf{R}; T_1) - \int d\epsilon \nabla_{\mathbf{R}_n} g^{T_1}(\mathbf{R}; \epsilon) \left[(\epsilon - \Sigma^{T_1}(\mathbf{R}; T_2)) f \left(\frac{\epsilon - \mu^{T_1}(\mathbf{R}; T_2)}{k_B T_2} \right) \right. \\ &- (\epsilon - \Sigma^{T_1}(\mathbf{R}; T_1)) f \left(\frac{\epsilon - \mu^{T_1}(\mathbf{R}; T_1)}{k_B T_1} \right) \left. \right] - k_B T_2 \int d\epsilon \nabla_{\mathbf{R}_n} g^{T_1}(\mathbf{R}; \epsilon) \left[s \left(\frac{\epsilon - \mu^{T_1}(\mathbf{R}; T_2)}{k_B T_2} \right) \right. \\ &\left. + \frac{\Sigma^{T_1}(\mathbf{R}; T_2) - \mu^{T_1}(\mathbf{R}; T_2)}{k_B T_2} f \left(\frac{\epsilon - \mu^{T_1}(\mathbf{R}; T_2)}{k_B T_2} \right) \right] . \end{aligned} \quad (\text{S44})$$

Or, using $T_1 \rightarrow 0^+$ and the definitions from (S27), along with (S4) and (S40), we have

$$\begin{aligned} \mathbf{F}_n(\mathbf{R}; T_{el}) &\approx -\nabla_{\mathbf{R}_n} E^0(\mathbf{R}; 0) - \int d\epsilon \nabla_{\mathbf{R}_n} g^0(\mathbf{R}; \epsilon) \left[(\epsilon - \Sigma^0(\mathbf{R}; T_{el})) f \left(\frac{\epsilon - \mu^0(\mathbf{R}; T_{el})}{k_B T_{el}} \right) \right. \\ &- (\epsilon - \epsilon_F^0(\mathbf{R})) \theta(\epsilon_F^0(\mathbf{R}) - \epsilon) \left. \right] - k_B T_{el} \int d\epsilon \nabla_{\mathbf{R}_n} g^0(\mathbf{R}; \epsilon) \left[s \left(\frac{\epsilon - \mu^0(\mathbf{R}; T_{el})}{k_B T_{el}} \right) \right. \\ &\left. + \frac{\Sigma^0(\mathbf{R}; T_{el}) - \mu^0(\mathbf{R}; T_{el})}{k_B T_{el}} f \left(\frac{\epsilon - \mu^0(\mathbf{R}; T_{el})}{k_B T_{el}} \right) \right] , \end{aligned} \quad (\text{S45})$$

where $f^0(0) = \theta(\epsilon_F^0(\mathbf{R}) - \epsilon)$ was used.

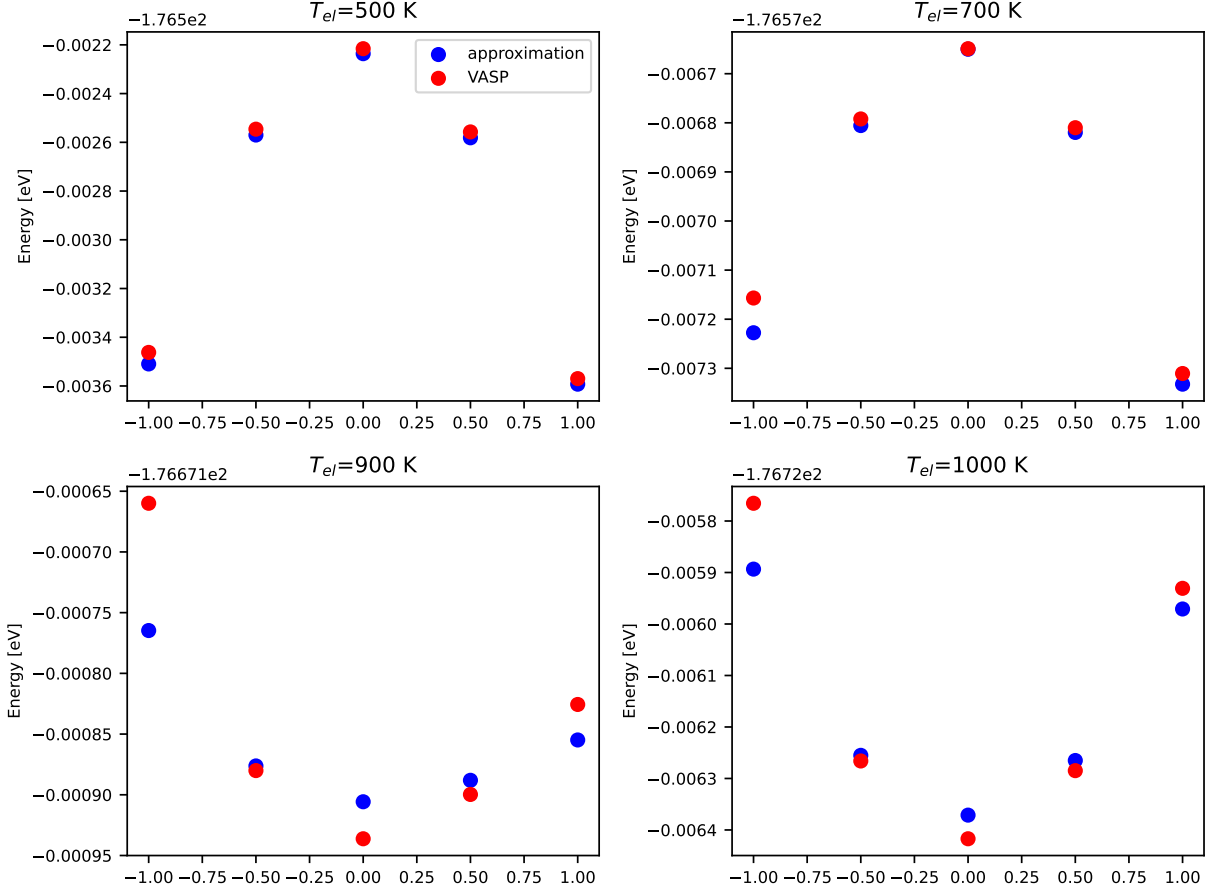


Figure S1. Comparison of the approximation (S33) and the VASP results for the Helmholtz electronic free energy for four different electronic temperatures T_{el} . Calculations were performed along the $\mathbf{q} = 2/3$ FM CDW phonon mode in $3 \times 3 \times 1$ supercell. For these calculations every eDOS was constructed using the same hyperparameters as in the main text. Namely, $g_b = 0.01$ eV and $d\epsilon = 0.00025$ eV.

In Fig. S1, we show the comparison between the approximation (S33) and the VASP results calculated along the $\mathbf{q} = 2/3$ FM CDW phonon mode in $3 \times 3 \times 1$ supercell.

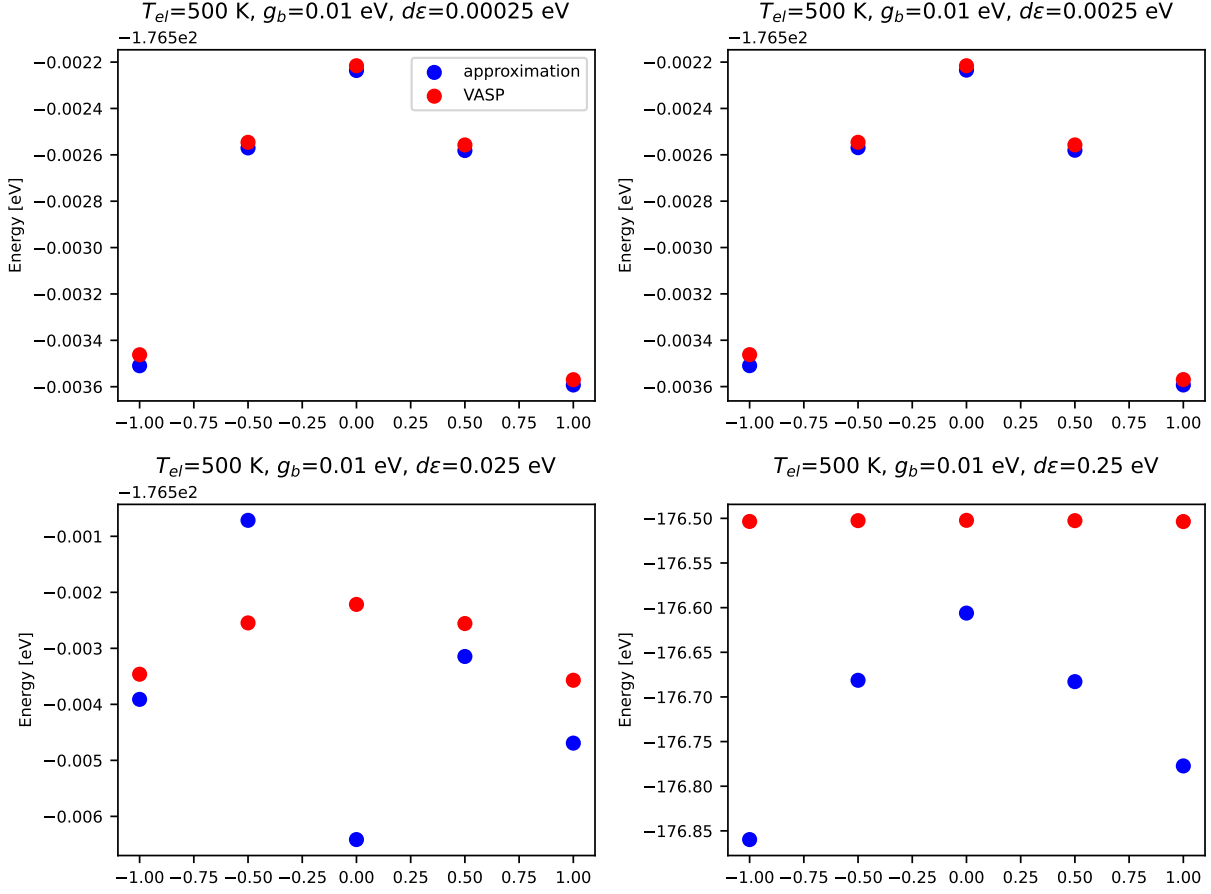


Figure S2. Comparison between the approximation (S33) and the VASP results for the Helmholtz electronic free energy calculated for one $3 \times 3 \times 1$ structure along the $\mathbf{q} = 2/3$ Γ M CDW phonon mode for $T_{el} = 500$ K, $g_b = 0.01$ eV and four different $d\epsilon$ hyperparameters.

In Fig. S2 we report the comparison between the approximation (S33) and the VASP results for the Helmholtz electronic free energy calculated for one $3 \times 3 \times 1$ structure along the $\mathbf{q} = 2/3$ Γ M CDW phonon mode for $T_{el} = 500$ K, $g_b = 0.01$ eV and four different $d\epsilon$ hyperparameters. As can be seen, smaller values of $d\epsilon$ result in better agreement between the approximation and the VASP results.

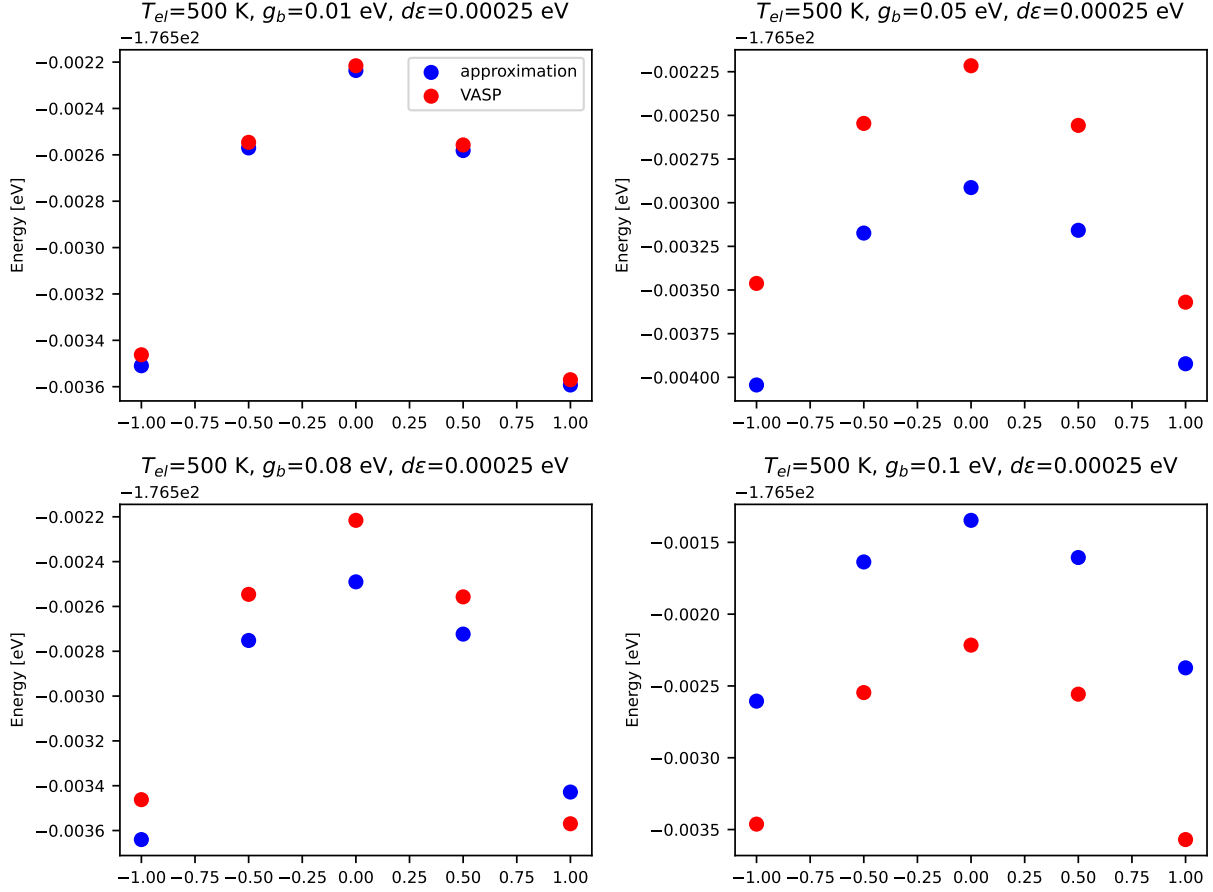


Figure S3. Comparison between the approximation (S33) and the VASP results for the Helmholtz electronic free energy calculated for one $3 \times 3 \times 1$ structure along the $\mathbf{q} = 2/3$ Γ M CDW phonon mode for $T_{el} = 500$ K, $d\epsilon = 0.00025$ eV and four different g_b hyperparameters.

In Fig. S3 we report the comparison between the approximation (S33) and the VASP results for the Helmholtz electronic free energy calculated for one $3 \times 3 \times 1$ structure along the $\mathbf{q} = 2/3$ Γ M CDW phonon mode for $T_{el} = 500$ K, $d\epsilon = 0.00025$ eV and four different g_b hyperparameters. As can be seen, smaller values of g_b result in better agreement between the approximation and the VASP results.

S2. SOAP AND KRR HYPERPARAMETER OPTIMIZATION

The ML model for the finite T_{el} part of the Helmholtz electronic free energy given by (S33) depends on both SOAP and KRR hyperparameters. We have performed the hyperparameter optimization of both of them at once using Optuna software [S7]. We have used TPESampler [S8], which is a Bayesian algorithm for hyperparameter optimization. The optimization was done in 100 steps, where in every step the training set was randomly split into temporary training and validation sets in the ratio 80% to 20%. This splitting was done using a different random seed in each step. We have kept test set comprised of 20% of the original data untouched until the last inference on it to get the unbiased metrics about the models' performance. The objective function during the optimization was given by Eq. (9) in the main text and the goal was to minimize this objective function evaluated on the validation set.

Hyperparameters we wanted to optimize were both SOAP and KRR ones, because this system to our knowledge has not been treated within this workflow, so we could not find SOAP hyperparameters that we could use as a starting point.

For constructing descriptors we have chosen the power spectrum SOAP representation, within it we had to optimize the cutoff radius (7.6 Å), maximum radial (10) and maximum angular (10) numbers [S9]. Furthermore, for representing ionic environments we have used Gaussian type orbitals of constant width, which also was a part of optimization process (0.7 Å). To optimize the hyperparameters that control the smoothness of decay of the interactions we have used radial scaling option [S10] and we optimized hyperparameters that come with it (scale= 6.84 Å, rate=1.2, exponent= 1).

Regarding the KRR hyperparameters, we had to optimize the number of sparse ionic environments out of total number of ions for each ionic species in our system. Since it would take very long time to train models if we had taken these numbers of sparse environments too big, we have chosen to optimize them in range from 0.5% to 1.25% of total number of ions for each ionic species (Se: 545 ionic environments, Nb: 272 ionic environments). Furthermore, there is a parameter that effectively includes higher than two-body interactions even when we use power spectrum, namely to construct the kernels we take the dot product between SOAP vectors and raise it to some power, called zeta ($\zeta = 2$). The equation for weights in KRR model is given by Eq. (11) in [S11]. A way to solve this equation in numerically

stable way, as implemented in *librascal*, is to choose the most relevant eigenvalues of the $K_{\underline{A}\underline{A}}$ matrix from the main text. This is controlled by threshold parameter (5.62×10^{-12}). This method is similar to principal component analysis.

Above we have written in parentheses the values of all hyperparameters of the "ML used" model from the main text. After the minimization of the %RMSE on the validation set, the

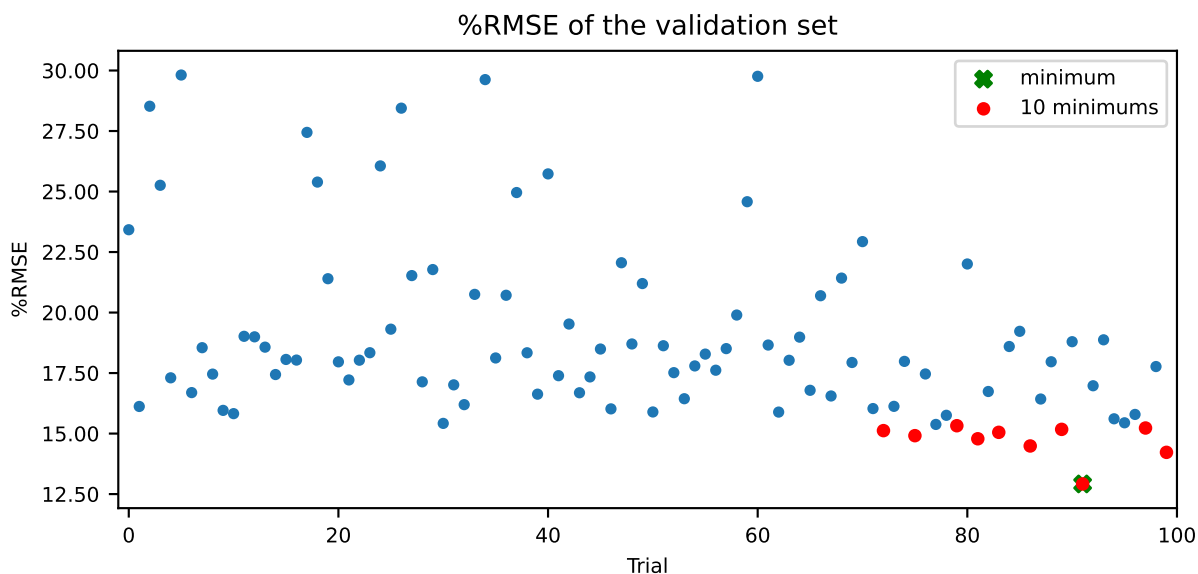


Figure S4. Distribution of %RMSE evaluated on the validation set during the hyperparameter optimization process.

models were trained again using the total training set and at the end evaluated on the test set.

In Fig. S4 we report the distribution of %RMSE on the validation set during the hyperparameter optimization process. Also, the 10 minimal errors are shown in red (in green we show the model with minimum %RMSE), those models were used to make an ensemble explained in the main text.

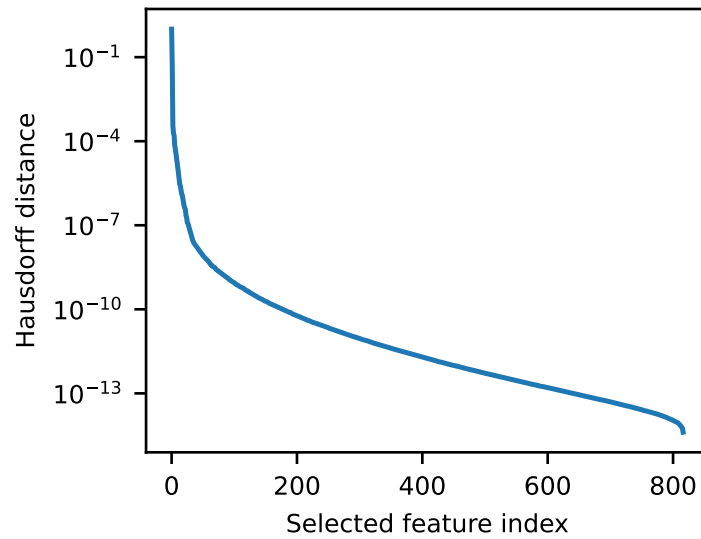


Figure S5. Plot of the Hausdorff distance used for selecting the number of relevant eigenvalues of the $K_{\underline{A} \underline{A}}$ matrix for the "ML used" model.

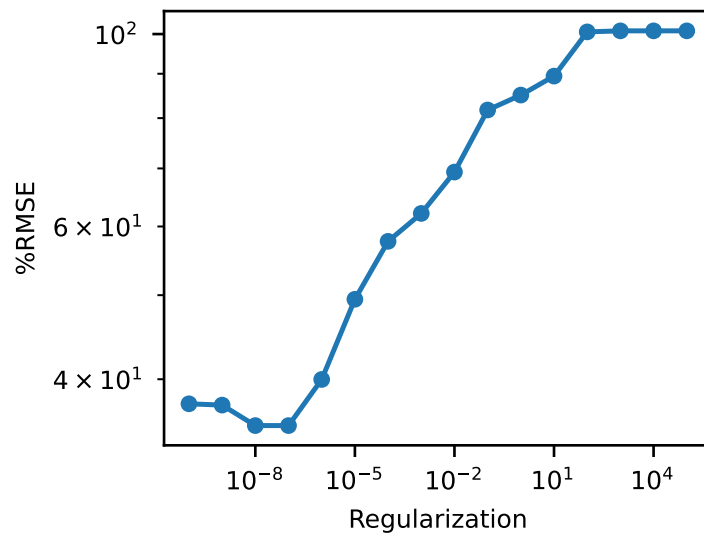


Figure S6. Plot of the different regularization parameters and their %RMSE for the "ML used" model.

In Figs. S5 we report the plot of the Hausdorff distance for "ML used" model used for selecting the number of relevant eigenvalues of the $K_{\underline{A} \underline{A}}$ matrix. Earlier explained threshold parameter is the Hausdorff distance chosen from this plot.

In Fig. S6 we show %RMSE vs. regularization plot, that helps us to determine the L2 type regularization parameter from the KRR model. This plot was obtained by 5-fold cross-validation and the minimum of this curve is taken as the best regularization parameter for a given model. The plot given here also corresponds to "ML used" model and the minimum of the curve is at 10^{-7} .

In literature one can also find a method [S12] where SOAP and KRR hyperparameters are optimized in separate phases. We have not used this method in this work.

S3. PHONON DISPERSIONS IN THE HARMONIC APPROXIMATION

In Fig. S7 we report all of the nine phonon dispersions (three acoustic and six optical branches) calculated in the harmonic approximation with both the VASP and the full interatomic ML potential from the main text. One can see that optical phonon dispersions are even in a better agreement between VASP and ML calculations.

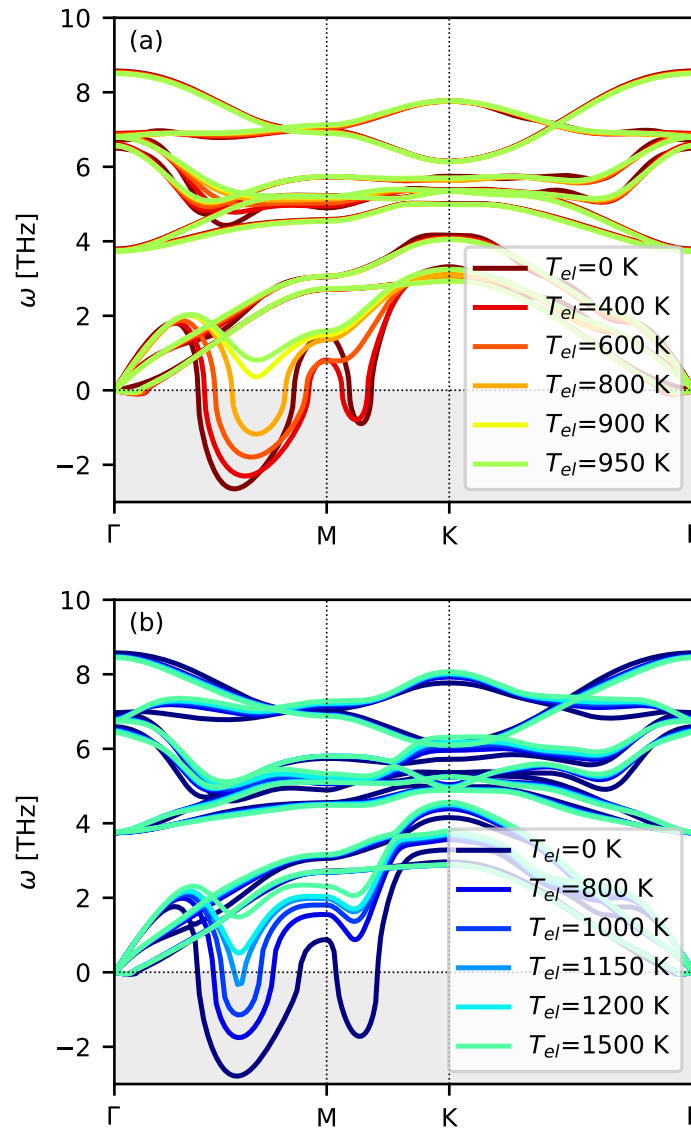


Figure S7. (a) All 9 phonon dispersions calculated using the VASP. (b) All 9 phonon dispersions calculated using ML model from the main text.

S4. TWO TEMPERATURE MODEL WITH DFT INPUT PARAMETERS

The energy flow between electron and phonon subsystems, represented with the corresponding temperatures $T_{el} = T_{el}(t)$ and $T_{ion} = T_{ion}(t)$, can be simulated using the following coupled equations [S13, S14]

$$\frac{\partial T_{el}}{\partial t} = \frac{I(t)}{C_{el}(T_{el})} - \frac{G}{C_{el}(T_{el})}(T_{el} - T_{ion}) , \quad (\text{S46})$$

$$\frac{\partial T_{ion}}{\partial t} = \frac{G}{C_{ion}(T_{ion})}(T_{el} - T_{ion}) . \quad (\text{S47})$$

The corresponding specific heats $C_{el}(T_{el})$ and $C_{ion}(T_{ion})$ are defined as:

$$C_{el}(T_{el}) = \int_{-\infty}^{\infty} d\varepsilon N(\varepsilon) \varepsilon \frac{\partial f(\varepsilon; T_{el})}{\partial T_{el}} , \quad (\text{S48})$$

$$C_{ion}(T_{ion}) = \int_0^{\infty} d\omega F(\omega) \omega \frac{\partial n(\omega; T_{ion})}{\partial T_{ion}} , \quad (\text{S49})$$

where $N(\varepsilon)$ is electron and $F(\omega)$ is phonon density of states. The electron-phonon relaxation rate G is calculated from density functional perturbation theory (DFPT) and electron-phonon calculations as

$$G = \frac{\pi k_B}{\hbar N(\varepsilon_F)} \lambda \langle \omega^2 \rangle \int_{-\infty}^{\infty} d\varepsilon N^2(\varepsilon) \left(-\frac{\partial f(\varepsilon; T_{el})}{\partial \varepsilon} \right) , \quad (\text{S50})$$

where λ is the total electron-phonon coupling strength,

$$\lambda = 2 \int d\Omega \frac{\alpha^2 F(\Omega)}{\Omega} , \quad (\text{S51})$$

and $\langle \omega^2 \rangle$ is the corresponding second moments of the phonon spectrum,

$$\langle \omega^2 \rangle = \frac{2}{\lambda} \int d\Omega \Omega \alpha^2 F(\Omega) . \quad (\text{S52})$$

Here $\alpha^2 F(\Omega)$ is the Eliashberg function, which represents the strength of electron-phonon coupling for each phonon energy Ω .

In addition, $I(t)$ describes a femtosecond Gaussian pump pulse with fluence F and duration t_p ,

$$I(t) = \frac{2F}{t_p} \sqrt{\frac{\log 2}{\pi}} \exp \left[-4 \log 2 \left(\frac{t}{t_p} \right)^2 \right] . \quad (\text{S53})$$

All the input parameters needed for the above TTM equations (except F and t_p) are obtained from DFT and DFPT calculations as implemented in QUANTUM ESPRESSO (QE)

package [S15]. The electron-phonon coupling constants were obtained with Wannier interpolation of EPC matrix elements as implemented in the EPW code [S16–S18]. In order to reduce the differences between the two type of calculations, here we chose the same structural parameters for NbSe₂ as the one used in VASP calculations of ML interatomic potential. For the same reason, we used the PBE functional [S19]. The DFPT calculations of the phonons were done on coarse grids of $\mathbf{k} = 24 \times 24 \times 1$ and $\mathbf{q} = 6 \times 6 \times 1$. We used Fermi-Dirac smearing with a broadening of 1600 K. For the electron-phonon calculations within EPW, we use 11 maximally localized Wannier functions [S16] with the initial projections of d orbitals on Nb and p orbitals on the Se atoms. To calculate Eliashberg function $\alpha^2F(\omega)$ we use a fine grids of $\mathbf{k} = 240 \times 240 \times 1$ and $\mathbf{q} = 80 \times 80 \times 1$. In Fig. S8 we show

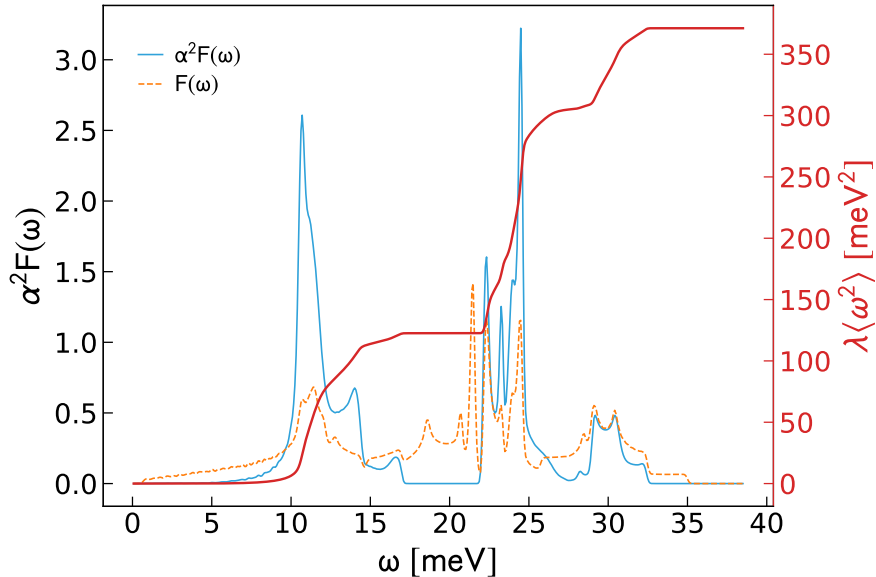


Figure S8. Theoretical results for Eliashberg function $\alpha^2F(\omega)$ (blue full line) and phonon density of states $F(\omega)$ (orange dashed line) in NbSe₂, with the corresponding value for the second moments of the phonon spectrum $\langle\omega^2\rangle$ (red).

the results for $\alpha^2F(\omega)$ and phonon density of states $F(\omega)$. From this we get $\lambda = 1.3$ and $\lambda\langle\omega^2\rangle = 371$ meV. The obtained value of λ is in a good agreement with estimation from the ARPES measurements, where the momentum-dependent λ is between 0.9 and 1.9 [S20], as well as with the previous DFT-based calculation [S21]. Further, in Fig. S9 we show the time evolution of electron and phonon temperatures as obtained with TTM and the DFT-derived input parameters for two different values of fluence F . For pulse width we took $t_p = 30$ fs,

which is very close to the experimental width [S21, S22]. For $F = 0.05 \text{ J/m}^2$ we get the maximal values of $T_{el} \approx 600 \text{ K}$ and $T_{ion} \approx 100 \text{ K}$, while for $F = 0.1 \text{ J/m}^2$ we get the maximal values of $T_{el} \approx 1000 \text{ K}$ and $T_{ion} \approx 170 \text{ K}$. Since we obtain that the CDW phase is melted above $T_{ion} \approx 100 \text{ K}$ and $T_{el} = 800 \text{ K}$, we consider that the critical value of fluence for melting the CDW phase in NbSe₂ is $F = 0.05 - 0.1 \text{ J/m}^2$, which agrees well with the pump-probe optical spectroscopy measurements [S22].

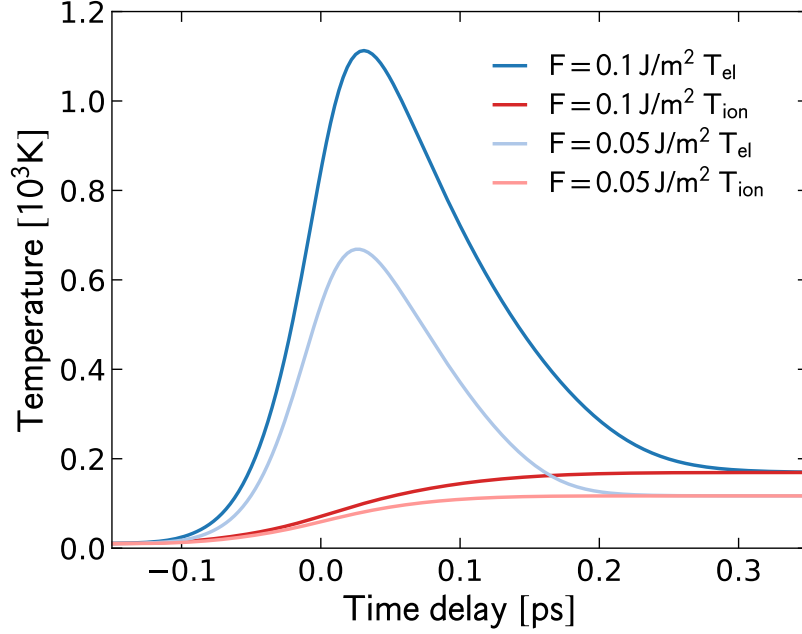


Figure S9. The results of the TTM calculations in NbSe₂ when two different fluences are applied.

-
- [S1] C. Ben Mahmoud, F. Grasselli, and M. Ceriotti, Predicting hot-electron free energies from ground-state data, *Phys. Rev. B* **106**, L121116 (2022).
- [S2] N. D. Mermin, Thermal Properties of the Inhomogeneous Electron Gas, *Phys. Rev.* **137**, A1441 (1965).
- [S3] F. Giustino, *Materials Modelling Using Density Functional Theory: Properties and Predictions* (Oxford University Press, 2014).
- [S4] J. F. Janak, Proof that $\frac{\partial E}{\partial n_i} = \epsilon$ in density-functional theory, *Phys. Rev. B* **18**, 7165 (1978).
- [S5] J. Chihara and M. Yamagiwa, Thermodynamics in Density-Functional Theory and Force Theorems, *Progress of Theoretical Physics* **111**, 339 (2004), <https://academic.oup.com/ptp/article-pdf/111/3/339/5199393/111-3-339.pdf>.
- [S6] P. Hohenberg and W. Kohn, Inhomogeneous Electron Gas, *Phys. Rev.* **136**, B864 (1964).
- [S7] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, *Optuna: A Next-generation Hyperparameter Optimization Framework* (2019), [arXiv:1907.10902 \[cs.LG\]](https://arxiv.org/abs/1907.10902).
- [S8] S. Watanabe, *Tree-Structured Parzen Estimator: Understanding Its Algorithm Components and Their Roles for Better Empirical Performance* (2023).
- [S9] A. P. Bartók, R. Kondor, and G. Csányi, On representing chemical environments, *Phys. Rev. B* **87**, 184115 (2013).
- [S10] M. J. Willatt, F. Musil, and M. Ceriotti, Feature optimization for atomistic machine learning yields a data-driven construction of the periodic table of the elements, *Phys. Chem. Chem. Phys.* **20**, 29661 (2018).
- [S11] V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti, and G. Csányi, Gaussian Process Regression for Materials and Molecules, *Chemical Reviews* **121**, 10073 (2021), pMID: 34398616, <https://doi.org/10.1021/acs.chemrev.1c00022>.
- [S12] S. K. Natarajan and M. A. Caro, *Particle Swarm Based Hyper-Parameter Optimization for Machine Learned Interatomic Potentials* (2020), [arXiv:2101.00049 \[cond-mat.mtrl-sci\]](https://arxiv.org/abs/2101.00049).
- [S13] F. Caruso and D. Novko, Ultrafast dynamics of electrons and phonons: from the two-temperature model to the time-dependent Boltzmann equation, *Advances in Physics: X* **7**, 2095925 (2022).
- [S14] D. Novko, F. Caruso, C. Draxl, and E. Cappelluti, Ultrafast Hot Phonon Dynamics in MgB₂

- Driven by Anisotropic Electron-Phonon Coupling, *Phys. Rev. Lett.* **124**, 077001 (2020).
- [S15] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, and et al., Advanced capabilities for materials modelling with QUANTUM ESPRESSO, *J. Phys. Condens. Matter* **29**, 465901 (2017).
- [S16] N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, Maximally localized Wannier functions: Theory and applications, *Rev. Mod. Phys.* **84**, 1419 (2012).
- [S17] J. Noffsinger, F. Giustino, B. D. Malone, C.-H. Park, S. G. Louie, and M. L. Cohen, EPW: A program for calculating the electron–phonon coupling using maximally localized Wannier functions, *Comput. Phys. Commun.* **181**, 2140 (2010).
- [S18] S. Poncé, E. Margine, C. Verdi, and F. Giustino, EPW: Electron–phonon coupling, transport and superconducting properties using maximally localized Wannier functions, *Comput. Phys. Commun.* **209**, 116 (2016).
- [S19] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [S20] T. Valla, A. V. Fedorov, P. D. Johnson, P.-A. Glans, C. McGuinness, K. E. Smith, E. Y. Andrei, and H. Berger, Quasiparticle Spectra, Charge-Density Waves, Superconductivity, and Electron-Phonon Coupling in $2H-NbSe_2$, *Phys. Rev. Lett.* **92**, 086401 (2004).
- [S21] A. Anikin, R. D. Schaller, G. P. Wiederrecht, E. R. Margine, I. I. Mazin, and G. Karapetrov, Ultrafast dynamics in the high-symmetry and in the charge density wave phase of $2H-NbSe_2$, *Phys. Rev. B* **102**, 205139 (2020).
- [S22] R. Venturini, A. Sarkar, P. Sutar, Z. Jagličić, Y. Vaskivskyi, E. Goreshnik, D. Mihailovic, and T. Mertelj, Unconventional photoinduced charge density wave dynamics in $2H-NbSe_2$, *Phys. Rev. B* **108**, 235160 (2023).