npj computational materials

www.nature.com/npjcompumats

ARTICLE OPEN

Electron–phonon physics from first principles using the EPW code

Hyungjun Lee[1,2], Samuel Poncé[3], Kyle Bushick[4], Samad Hajinazar[5,11], Jon Lafuente-Bartolome[1,2], Joshua Leveillee[1,2], Chao Lian[1,2], Jae-Mo Lihm[6], Francesco Macheda[7,8], Hitoshi Mori[5], Hari Paudyal[5], Weng Hong Sio[1,9], Sabyasachi Tiwari[1,2], Marios Zacharias[10], Xiao Zhang[4], Nicola Bonini[7], Emmanouil Kioupakis[4], Elena R. Margine[5] and Feliciano Giustino[1,2]✉

EPW is an open-source software for ab initio calculations of electron–phonon interactions and related materials properties. The code combines density functional perturbation theory and maximally localized Wannier functions to efficiently compute electron–phonon coupling matrix elements, and to perform predictive calculations of temperature-dependent properties and phonon-assisted quantum processes in bulk solids and low-dimensional materials. Here, we report on significant developments in the code since 2016, namely: a transport module for the calculation of charge carrier mobility under electric and magnetic fields using the Boltzmann transport equation; a superconductivity module for calculations of phonon-mediated superconductors using the anisotropic multi-band Eliashberg theory; an optics module for calculations of phonon-assisted indirect transitions; a module for the calculation of small and large polarons without supercells; and a module for calculating band structure renormalization and temperature-dependent optical spectra using the special displacement method. For each capability, we outline the methodology and implementation and provide example calculations.

npj Computational Materials (2023)9:156 ; https://doi.org/10.1038/s41524-023-01107-3

INTRODUCTION

The coupling between electrons and phonons is one of the most intensively studied fermion–boson interactions in condensed matter physics. It is responsible for a number of physical phenomena in solids such as conventional superconductivity, temperature-dependent resistivity in metals and mobility in semiconductors, the formation of polarons, and phonon-assisted optical processes, to name a few[1,2]. Predictive non-empirical calculations of electron–phonon interactions play an essential role in elucidating a variety of materials properties and their temperature dependence.

During the past three decades, first-principles calculations of electron–phonon interactions have been made possible by the development of density functional theory (DFT)[3,4] and density-functional perturbation theory (DFPT)[5–7]. However, the computational cost of these calculations is high, as they involve the evaluation of Brillouin zone integrals which typically require a very fine sampling of the crystal momenta of electrons and phonons. For instance, in DFPT calculations, every single phonon wavevector and vibrational mode requires the solution of Sternheimer-type equations that carry a computational complexity comparable to a DFT total energy calculation[8]. As a result, calculating properties related to the electron–phonon interaction usually entails a computational workload that is orders of magnitude more costly than standard DFT calculations, thus rendering direct ab initio calculations of these quantities challenging.

Several methods have been proposed to address this challenge[1]. Among those, one of the authors proposed to combine DFPT with maximally localized Wannier functions (MLWFs)[9] to perform a physics-based interpolation via a generalized Fourier transformation[10]. This method exploits the spatial localization of the electron-phonon matrix elements in the Wannier representation and enables efficient calculations of electron–phonon matrix elements on ultra-dense momentum grids while retaining the accuracy of DFPT.

The `EPW` code builds upon this methodology to compute a number of properties relating to electron–phonon interactions and temperature-dependent materials properties. `EPW` is the first open-source ab initio software devoted to electron–phonon interactions and has been actively developed for over 16 years. It was publicly released in 2010 under a GNU General Public License (GPL)[11] and is being distributed within the `Quantum ESPRESSO` materials simulation suite[12] since 2016. In addition to `EPW`, several software projects have been developed in recent years to address electron–phonon interactions, including: `Perturbo`[13], `elphbolt`[14], `Phoebe`[15], and `EPIq`[16]. All these packages rely on the same interpolation methodology employed by `EPW`[10].

Two prior manuscripts[17,18] describe the status of the `EPW` code until 2016. The present manuscript aims to describe progress on the code that has occurred since 2016 as well as ongoing efforts, leading to the latest `EPW` release. Added functionalities involve: a magneto-transport module for the calculation of the conductivity

[1]Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX 78712, USA. [2]Department of Physics, The University of Texas at Austin, Austin, TX 78712, USA. [3]European Theoretical Spectroscopy Facility, Institute of Condensed Matter and Nanosciences, Université catholique de Louvain, BE-1348 Louvain-la-Neuve, Belgium. [4]Department of Materials Science and Engineering, University of Michigan, Ann Arbor, MI 48109, USA. [5]Department of Physics, Applied Physics and Astronomy, Binghamton University-SUNY, Binghamton, NY 13902, USA. [6]Department of Physics and Astronomy, Seoul National University, Seoul 08826, Korea. [7]Department of Physics, King's College London, Strand, London WC2R 2LS, UK. [8]Istituto Italiano di Tecnologia, Graphene Labs, Via Morego 30, I-16163 Genova, Italy. [9]Institute of Applied Physics and Materials Engineering, University of Macau, Macao, SAR 999078, P. R. China. [10]University of Rennes, INSA Rennes, CNRS, Institut FOTON-UMR 6082, F-35000 Rennes, France. [11]Present address: Department of Chemistry, University at Buffalo, Buffalo, NY 14260, USA. We dedicate this manuscript to our dearest friend and beloved collaborator, Professor Nicola Bonini, who passed away in October 2022. Nicola was a source of inspiration and joy to us all and played a key role in the development of EPW.
✉email: fgiustino@oden.utexas.edu

in metals as well as the drift and Hall mobility in semiconductors; a module for the calculation of small and large polarons without using supercells; a module for the solution of the full-band Eliashberg equations for superconductors; a module for calculations of phonon-assisted indirect optical processes; and a module for calculations of finite-temperature electronic and optical properties via the special displacement method. The code has been refactored to keep up with contemporary developments in high-performance computing (HPC) architectures, for example via hybrid two-level message passing interface (MPI[19]) and Open Multi-Processing (OpenMP[20]) parallelization, and parallel I/O via parallel Hierarchical Data Format 5 (HDF5[21]).

The manuscript is organized as follows. In the 'Methods' section (located at the end of the manuscript), we describe the notation employed throughout this manuscript, we review the conceptual basis of electron–phonon interpolation using DFPT and MLWFs, and we discuss the extension of this methodology to the case of polar materials with long-range Coulomb interactions. In the 'Results' section, we describe the additional or expanded functionalities available in the latest release of the EPW code. In particular, we begin by discussing calculations of carrier transport within the ab initio Boltzmann transport equation. We consider both electric and magnetic fields, as well as phonon-limited and charged defect-limited transport. Then we move to calculations of the superconducting critical temperature and superconducting gap function. Here, we discuss elementary calculations based on the semi-empirical McMillan equation, the isotropic Eliashberg theory, and the fully-anisotropic full-band Eliashberg theory. Next, we outline the methodology for computing small and large polarons from DFPT without resorting to large supercells. We then proceed with the formalism for the calculation of optical absorption spectra including phonon-assisted indirect transitions. After these applications which are based on Wannier–Fourier interpolation, we describe the special displacement method, which offers an alternative and complementary strategy to computing electron–phonon properties without employing Wannier–Fourier interpolation. At the end of this section, we discuss recent advances in computational algorithms, parallelization, and I/O, and we report on parallel scaling benchmarks. In the 'Discussion' section we discuss possible future directions for the EPW software project.

RESULTS

In this section, we provide an overview of additional or expanded capabilities of the EPW code. In particular, we discuss how EPW calculates carrier transport properties via the ab initio Boltzmann transport equation (*ai*BTE), including both carrier-phonon and carrier-impurity scattering; how the superconducting critical temperature and superconducting gap are computed via the solution of the Eliashberg equations; how we solve the ab initio polaron equations to investigate small and large polarons without using large supercells; and how we perform calculations of optical absorption spectra in indirect-gap semiconductors by including phonon-assisted optical transitions. We also describe an alternative to Wannier interpolation to study electron–phonon interactions, based on the special displacement method. Of these features, the superconducting module has been enhanced and expanded with respect to the previous release in 2016[18]; all the other modules described in the following have been developed ex novo since the 2016 release. The notations and conventions employed below are described in detail in the 'Methods' section. In the same section, we provide an overview of the Wannier–Fourier interpolation of electron–phonon matrix elements.

Phonon-limited carrier transport using the ab initio Boltzmann transport equation

Background and formalism. The calculation of the electronic transport properties of metals and semiconductors is conveniently dealt with by the ab initio Boltzmann transport equation (*ai*BTE). The Boltzmann equation describes the non-equilibrium distribution function of electrons and holes in the presence of external electric or magnetic fields[2]. It carries strong predictive power as recently demonstrated for many common semiconductors[22]. Although the Boltzmann formalism is usually derived within the semi-classical approximation, the theory can rigorously be derived from a non-equilibrium many-body Green's function formalism and is understood as the quasiparticle approximation to the Kadanoff–Baym theory[23–26].

The EPW code implements the linearized Boltzmann transport equation, which describes the first-order response of the distribution function to external fields and is suitable for investigating the conductivity of metals and the low-field mobility of semiconductors. In this section, we focus on the case of phonon-limited transport properties. The extension to include charged defects is discussed in the next section.

We consider homogeneous extended solids held at uniform temperature, and carrier scattering by electron–phonon interactions only, for now. Within the *ai*BTE, the linear response $\partial_{E_\beta} f_{n\mathbf{k}}$ of the carrier distribution function to an external electric field $\mathbf{E}$ is given by[22,27–41]

$$\begin{aligned}\partial_{E_\beta} f_{n\mathbf{k}} = e v_{n\mathbf{k}\beta} \tfrac{\partial f^0_{n\mathbf{k}}}{\partial \varepsilon_{n\mathbf{k}}} \tau_{n\mathbf{k}} + \tfrac{2\pi\tau_{n\mathbf{k}}}{\hbar} \sum_{m\nu} \int \frac{\mathrm{d}\mathbf{q}}{\Omega_{\mathrm{BZ}}} |g_{mn\nu}(\mathbf{k},\mathbf{q})|^2 \partial_{E_\beta} f_{m\mathbf{k}+\mathbf{q}} \\ \times \left[(n_{\mathbf{q}\nu} + 1 - f^0_{n\mathbf{k}}) \delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} + \hbar\omega_{\mathbf{q}\nu}) \right. \\ \left. + (n_{\mathbf{q}\nu} + f^0_{n\mathbf{k}}) \delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} - \hbar\omega_{\mathbf{q}\nu}) \right].\end{aligned} \tag{1}$$

In this expression, $f_{n\mathbf{k}}$ is the electron distribution function, n denotes a Kohn–Sham band index, and $\mathbf{k}$ is the electron wavevector. $\partial_{E_\beta} f_{n\mathbf{k}}$ is a short-hand notation for $(\partial f_{n\mathbf{k}}/\partial E_\beta)|_{\mathbf{E}=0}$, $v_{n\mathbf{k}\alpha} = \hbar^{-1}\partial\varepsilon_{n\mathbf{k}}/\partial k_\alpha$ is the intra-band velocity matrix element for the Kohn–Sham eigenvalue $\varepsilon_{n\mathbf{k}}$, and δ denotes the Dirac delta function. $g_{mn\nu}(\mathbf{k},\mathbf{q})$ is the electron–phonon matrix element describing the scattering from the state $n\mathbf{k}$ to the state $m\mathbf{k}+\mathbf{q}$ via the phonon of branch ν and wavevector $\mathbf{q}$ (see the "Methods" section for more details). The integral is over the Brillouin zone. The temperature enters this equation via the Fermi–Dirac and Bose–Einstein equilibrium distribution functions $f^0_{n\mathbf{k}}$ and $n_{\mathbf{q}\nu}$, respectively. The quantity $\tau_{n\mathbf{k}}$ in Eq. (1) is the carrier relaxation time, and is obtained from Fermi's golden rule:

$$\begin{aligned}\tau^{-1}_{n\mathbf{k}} = \tfrac{2\pi}{\hbar} \sum_{m\nu} \int \frac{\mathrm{d}\mathbf{q}}{\Omega_{\mathrm{BZ}}} |g_{mn\nu}(\mathbf{k},\mathbf{q})|^2 \\ \times \left[(n_{\mathbf{q}\nu} + 1 - f^0_{m\mathbf{k}+\mathbf{q}}) \delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} - \hbar\omega_{\mathbf{q}\nu}) \right. \\ \left. + (n_{\mathbf{q}\nu} + f^0_{m\mathbf{k}+\mathbf{q}}) \delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} + \hbar\omega_{\mathbf{q}\nu}) \right].\end{aligned} \tag{2}$$

The electrical conductivity tensor $\sigma_{\alpha\beta}$, which corresponds to the variation of the current density with respect to the electric field, is computed from $\partial_{E_\beta} f_{n\mathbf{k}}$ using:

$$\sigma_{\alpha\beta} = -\frac{e}{\Omega} \sum_n \int \frac{\mathrm{d}\mathbf{k}}{\Omega_{\mathrm{BZ}}} v_{n\mathbf{k}\alpha} \partial_{E_\beta} f_{n\mathbf{k}}, \tag{3}$$

and the drift mobility is obtained from this expression upon dividing by the carrier density n_c:

$$\mu_{\alpha\beta} = \frac{\sigma_{\alpha\beta}}{e\, n_c}. \tag{4}$$

These equations are valid for metals and for electrons or holes in semiconductors taken separately. A common approximation to the *ai*BTE is the relaxation time approximation, which consists of

neglecting the second term on the right-hand side of Eq. (1). In this case the mobility takes the simpler form:

$$\mu_{\alpha\beta}^{\mathrm{SERTA}} = -\frac{e}{n_c\Omega}\sum_n \int \frac{\mathrm{d}\mathbf{k}}{\Omega_{\mathrm{BZ}}} \frac{\partial f_{n\mathbf{k}}^0}{\partial \varepsilon_{n\mathbf{k}}} v_{n\mathbf{k}\alpha} v_{n\mathbf{k}\beta} \tau_{n\mathbf{k}}. \tag{5}$$

We refer to this approximation as the "self-energy relaxation time approximation" (SERTA) since the transport lifetime in Eq. (2) corresponds to the quasiparticle lifetime, which is proportional to the imaginary part of the electron-phonon self-energy[1,22,24].

The electrical conductivity and the drift mobility described by Eq. (1) can be measured by time-of-flight measurements or by THz photo-conductivity measurements. In the case of Hall and De Haas–Van Alphen measurements, an additional magnetic field $\mathbf{B}$ is applied, and the resulting Lorentz force must be taken into account in the *ai*BTE. In these cases, EPW solves the following *ai*BTE equation[22,24,32,42]:

$$\begin{aligned}&\left[1-\tfrac{e}{\hbar}\tau_{n\mathbf{k}}(\mathbf{v}_{n\mathbf{k}}\times\mathbf{B})\cdot\nabla_{\mathbf{k}}\right]\partial_{E_\beta} f_{n\mathbf{k}}(\mathbf{B}) = e v_{n\mathbf{k}\beta}\tfrac{\partial f_{n\mathbf{k}}^0}{\partial\varepsilon_{n\mathbf{k}}}\tau_{n\mathbf{k}}\\ &+\tfrac{2\pi\tau_{n\mathbf{k}}}{\hbar}\sum_{m\nu}\int\frac{\mathrm{d}\mathbf{q}}{\Omega_{\mathrm{BZ}}}|g_{mn\nu}(\mathbf{k},\mathbf{q})|^2\partial_{E_\beta} f_{m\mathbf{k}+\mathbf{q}}(\mathbf{B})\\ &\times\left[(n_{\mathbf{q}\nu}+1-f_{n\mathbf{k}}^0)\delta(\varepsilon_{n\mathbf{k}}-\varepsilon_{m\mathbf{k}+\mathbf{q}}+\hbar\omega_{\mathbf{q}\nu})+(n_{\mathbf{q}\nu}+f_{n\mathbf{k}}^0)\delta(\varepsilon_{n\mathbf{k}}-\varepsilon_{m\mathbf{k}+\mathbf{q}}-\hbar\omega_{\mathbf{q}\nu})\right].\end{aligned} \tag{6}$$

The variation $\partial_{E_\beta} f_{n\mathbf{k}}(\mathbf{B})$ is computed for magnetic fields sufficiently small that their effect on the electronic and vibrational properties can be ignored; accordingly, the Kohn–Sham energies, phonons, and their couplings in the above equation are all evaluated for $\mathbf{B}=0$. The application of a magnetic field in the direction $\hat{\mathbf{B}}$ results in an orthogonal flow of the charge carriers which can be described by the linear response of the mobility to the field:

$$\mu_{\alpha\beta}^{(2)}(\hat{\mathbf{B}}) = -\lim_{\mathbf{B}\to\mathbf{0}}\frac{1}{|\mathbf{B}|}\left[\frac{1}{n_c\Omega}\sum_n\int\frac{\mathrm{d}\mathbf{k}}{\Omega_{\mathrm{BZ}}} v_{n\mathbf{k}\alpha}\partial_{E_\beta} f_{n\mathbf{k}}(\mathbf{B}) - \mu_{\alpha\beta}\right]. \tag{7}$$

From this expression, we obtain the Hall mobility in terms of the drift mobility as follows:

$$\mu_{\alpha\beta}^{\mathrm{H}}(\hat{\mathbf{B}}) = \sum_\gamma r_{\alpha\gamma}^{\mathrm{H}}(\hat{\mathbf{B}})\mu_{\gamma\beta}\,, \tag{8}$$

where we have introduced the dimensionless Hall factor as[22]

$$r_{\alpha\gamma}^{\mathrm{H}}(\hat{\mathbf{B}}) = \sum_{\gamma\delta}\mu_{\alpha\gamma}^{-1}\mu_{\gamma\delta}^{(2)}(\hat{\mathbf{B}})\mu_{\delta\beta}^{-1}. \tag{9}$$

Also in this case, one can simplify the solution of Eq. (6) by ignoring the second term on the right-hand side. This choice leads to the SERTA approximation in the presence of a magnetic field.

Computational considerations. EPW implements an iterative solver to obtain $\partial_{E_\beta} f_{n\mathbf{k}}$ from Eq. (1), with the possibility of Broyden mixing[43] to accelerate the convergence. The accurate evaluation of Eqs. (3) and (4) is computationally demanding because it requires the knowledge of $\partial_{E_\beta} f_{n\mathbf{k}}$ for a dense set of $\mathbf{k}$-points in an energy window of the order of a few tens of meV around the Fermi energy. In particular, since the scattering integral on the right-hand side of Eq. (1) couples linear response coefficients at $\mathbf{k}$ and $\mathbf{k}+\mathbf{q}$ points, the $\mathbf{k}$-grid and the $\mathbf{q}$-grid must be commensurate. As this term is not evaluated in the SERTA approximation, SERTA calculations can instead be performed using incommensurate grids, including for example random and quasi-random sampling[31]. It should be noted, however, that the SERTA approximation tends to underestimate transport coefficients by up to 50% as compared to the *ai*BTE[22].

Regardless of the approximation chosen to calculate transport coefficients, EPW exploits crystal symmetry operations to eliminate symmetry-equivalent wavevectors in the solution of Eq. (1); accordingly, $\partial_{E_\beta} f_{n\mathbf{k}}$ is evaluated within the irreducible wedge of the Brillouin zone. Furthermore, the solution of Eq. (1) is restricted to the set of wavevectors for which both the initial and the final electronic state lie within a user-defined energy window around reference energy. In particular, the wavevectors $\mathbf{k}$ and $\mathbf{q}$ are retained only if there exists a pair of bands m and n such that $|\varepsilon_{n\mathbf{k}}-E^{\mathrm{ref}}|<\Delta$ and $|\varepsilon_{m\mathbf{k}+\mathbf{q}}-E^{\mathrm{ref}}|<\Delta$, where E^{ref} is reference energy and Δ is the width of the window. In the case of metals, the reference energy is set to the Fermi energy. In the case of semiconductors, the Fermi energy is calculated from the user-specified temperature and carrier concentration, using the bisection method.

Magneto-transport calculations using Eq. (6) are harder to converge than calculations without magnetic fields. The iterative solution is initialized by using the result of Eq. (1) as a first approximation for $\partial_{E_\beta} f_{n\mathbf{k}}(\mathbf{B})$, and the quantity $\nabla_{\mathbf{k}}\partial_{E_\beta} f_{n\mathbf{k}}(\mathbf{B})$ appearing on the left-hand side of Eq. (6) is evaluated via finite differences. The correctness of the solution is tested by checking for the conservation of the carrier density, which corresponds to the condition $\sum_{n\mathbf{k}}\partial_{E_\beta} f_{n\mathbf{k}}(\mathbf{B})=0$.

To ensure numerically accurate evaluation of transport coefficients with EPW, it is important to verify that results be converged with respect to (i) the density of the Brillouin zone grids employed in Eqs. (1) and (6); (ii) the density of the coarse grids used for the Wannier–Fourier interpolation of the DFT and DFPT data[22]; (iii) the Gaussian smearing employed to compute the Dirac delta functions appearing in Eq. (1); and (iv) to check for the sensitivity of results to lattice parameters, exchange and correlation functionals, and pseudopotentials. To facilitate the convergence with respect to the Gaussian smearing and the $\mathbf{k}$-point sampling, EPW offers the possibility of using the adaptive smearing method of ref. [44].

Application example. To demonstrate the implementation of the transport module in EPW, we investigate the electron and hole mobilities of cubic boron nitride, as well as the associated Hall factors. We use the relativistic Perdew–Burke–Ernzerhof (PBE) parametrization[45] of the generalized gradient approximation to DFT. The pseudopotentials are norm-conserving, generated using the ONCVPSP code[46], and optimized via the PseudoDojo initiative[47]. We consider room temperature and low carrier concentrations of $10^{13}\,\mathrm{cm}^{-3}$. c-BN is a polar wide gap semiconductor with isotropic Born effective charges $Z_{\mathrm{B}}^*=1.91\,e$ and $Z_{\mathrm{N}}^*=-Z_{\mathrm{B}}^*$; isotropic dynamical quadrupoles $Q_{\kappa,\alpha\beta\gamma}=Q_\kappa|\epsilon_{\alpha\beta\gamma}|$ where $\epsilon_{\alpha\beta\gamma}$ is the Levi–Civita tensor and Q_κ an atom-dependent scalar with values $Q_{\mathrm{B}}=3.46ea_0$ and $Q_{\mathrm{N}}=-0.63ea_0$, where a_0 is the Bohr radius; and an isotropic high-frequency relative dielectric constant $\epsilon^\infty=4.54$[22]. As will be discussed in Fig. 9, long-range electron–phonon interactions are important in the case of c-BN. Wannier functions are calculated separately for the valence and the conduction band manifold in order to reduce the computational cost.

Figure 1a shows the convergence of calculated mobilities with the coarse grid of $\mathbf{k}$-points used in the Wannierization procedure. For this test, we employ a coarse $\mathbf{q}$-point grid with half the number of the $\mathbf{k}$-points; and we employ identical fine grids with 50^3 $\mathbf{k}$- and $\mathbf{q}$-points for the solution of Eq. (1); we use an energy window $\Delta=0.3$ eV around the reference energy E^{ref} set to the band edge, and we employ adaptive smearing. This panel shows that the mobilities are converged with an accuracy of 1% when coarse grids with 18^3 and 14^3 points are employed for electrons and holes, respectively. Figure 1b shows the convergence of calculated mobilities with respect to the fine grids. We see that both electron and hole mobilities are almost converged for grids consisting of 250^3 points. Generally speaking, the fine grid required to achieve converged results depends on the effective masses, with smaller masses requiring denser grids. As an example, accurate calculations of the electron mobility of GaAs require homogeneous grids of order 400^3 points, while in the case of AlAs or GaP 80^3 points are sufficient[22] (see Table 1).

In Fig. 1c we show convergence curves for the mobility evaluated in the SERTA approximation. In this case, the

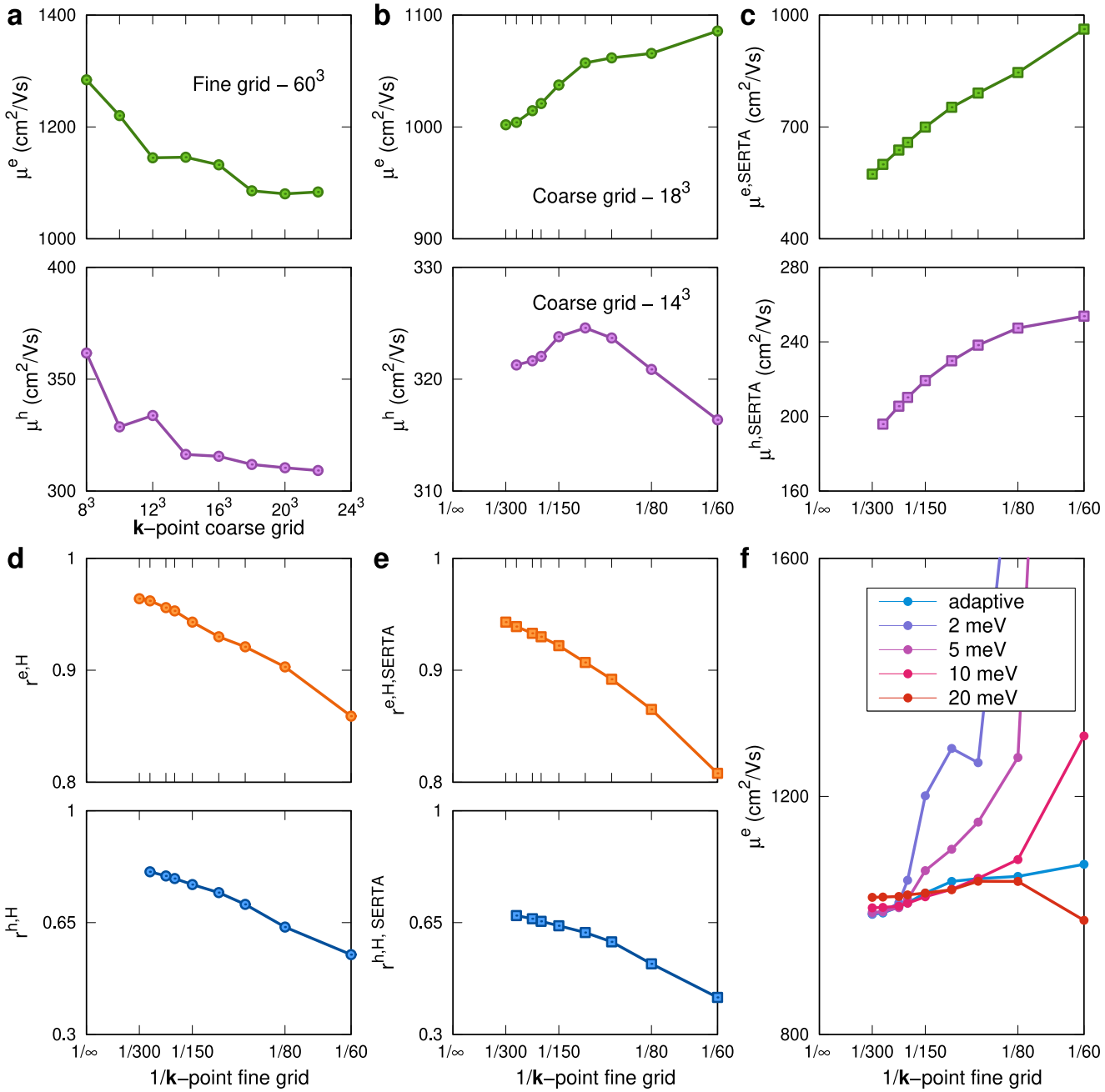

Fig. 1 Calculations of phonon-limited carrier mobility using `EPW`**. a** Convergence test for the electron and hole mobilities of c-BN with respect to the number of **k**-points on the coarse grid. The number of **q**-points is set to half the number of **k**-points. **b** Convergence of mobilities w.r.t. the number of grid points on the fine grids. The number of **q**-points is set to be the same as the number of **k**-points. **c** Same as in **b**, but this time for calculations within the SERTA approximation. **d** and **e** Hall factors of c-BN for electrons and holes, as a function of the number of points on the fine grids, for the full *ai*BTE and the SERTA approximation, respectively. The number of **q**-points is set to be the same as the number of **k**-points. **f** Dependence of the electron drift mobility of c-BN with the Gaussian smearing.

convergence with the density of points in the fine grids is much slower than for complete *ai*BTE calculations. More importantly, SERTA mobilities can differ by up to a factor of two from the corresponding *ai*BTE results. The large difference between full *ai*BTE calculations and SERTA seems to be a common trend in polar materials[22]. For the same **k**- and **q**-point grids, the computational saving afforded by the SERTA approximation is minimal; therefore, we recommend using SERTA only when the use of commensurate grids is too demanding, for example in the case of systems with many atoms in the unit cell. Figure 1d, e show the computed Hall factors for electrons and holes in c-BN, respectively. The Hall factor in the SERTA approximation appears to be close to the full *ai*BTE result, which is expected since it is defined as the ratio of two mobilities (see Eq. (9))[48].

In Fig. 1(f) we analyze the role of the Gaussian smearing parameter in the calculations. To this end, we compute the electron mobility of c-BN as a function of grid size, for varying smearing parameter. It is seen that the size of the grid necessary for convergence increases when decreasing the smearing in the few meV's range. On the other hand, the use of adaptive smearing affords fast convergence even for smaller grid sizes. Based on this comparison, adaptive smearing is the computationally most convenient strategy.

Table 1. Typical size of fine grids employed in EPW calculations.

EPW module	Typical size of fine Brillouin zone grid	
	Minimum	Maximum
Phonon-limited transport	$80\times80\times80$	$400\times400\times400$
Impurity-limited transport	$100\times100\times100$	$180\times180\times180$
Superconductivity	$60\times60\times60$	$120\times120\times120$
Polarons	$3\times3\times3$	$50\times50\times50$
Phonon-assisted absorption	$32\times32\times32$	$48\times48\times48$
Special displacement method	$3\times3\times3$	$10\times10\times10$

In the case where the fine grids for **k**- and **q**-points differ, we indicate the size of the denser grid. These values are only indicative and correspond to the examples mentioned in the 'Results' section, which refer to simple semiconductors, metals, and insulators with 2–4 atoms per unit cell. For systems with larger unit cells, the fine grids would have to be downscaled accordingly.

Defect-limited carrier transport using the ab initio Boltzmann transport equation

Background and formalism. Electron–phonon interactions are the dominant scattering mechanism in high-purity single crystals with low defect concentration, typically up to $10^{15}\,\mathrm{cm}^{-3}$ [49]. In tetrahedral semiconductors and a host of other materials, charged defects are ubiquitous since dopant elements are employed to introduce free carriers[50]. Donors release electrons to the conduction band, thus becoming positively charged defects; similarly, acceptors release holes into the valence band, and become negatively charged defects. In either case, the impurity generates a long-ranged Coulomb potential that scatters charge carriers. This scattering mechanism tends to dominate over electron–phonon processes at high doping concentrations.

There exist popular semi-empirical relations to estimate the effect of ionized impurity scattering on carrier transport in semiconductors, such as the Brooks–Herring[51] and Conwell–Weisskopf[52] formulas. However, these expressions rely on simplified parabolic band models and do not carry predictive power in the case of materials with multi-band or multi-valley band extrema. First-principles calculations offer a modern alternative to these earlier approaches, and have achieved considerable success in recent years[53–56]. The EPW code implements a module for charged defect scattering based on a randomized distribution of point charges. This approach is described in detail in ref. [57] and summarized below.

For convenience we rewrite the *ai*BTE from Eq. (1) in the more compact form:

$$-ev_{n\mathbf{k}\beta}\frac{\partial f^0_{n\mathbf{k}}}{\partial\varepsilon_{n\mathbf{k}}}=\sum_m\int\frac{\mathrm{d}\mathbf{q}}{\Omega_{\mathrm{BZ}}}\left(\tau^{-1}_{m\mathbf{k}+\mathbf{q}\to n\mathbf{k}}\,\partial_{E_\beta}f_{m\mathbf{k}+\mathbf{q}}-\tau^{-1}_{n\mathbf{k}\to m\mathbf{k}+\mathbf{q}}\,\partial_{E_\beta}f_{n\mathbf{k}}\right),\tag{10}$$

where the quantity $\tau^{-1}_{n\mathbf{k}\to m\mathbf{k}+\mathbf{q}}$ is the partial scattering rate from the Kohn–Sham state $n\mathbf{k}$ to the state $m\mathbf{k}+\mathbf{q}$. In the case of electron–phonon scattering, $\tau^{-1}_{n\mathbf{k}\to m\mathbf{k}+\mathbf{q}}$ is obtained from Eq. (2) by removing the summation over m and the integral over the Brillouin zone, i.e. $\sum_m\int d\mathbf{q}/\Omega_{\mathrm{BZ}}$. When both electron–phonon scattering and charged impurity scattering are taken into account, the partial scattering rate is written as the sum of the individual partial rates

$$\frac{1}{\tau_{n\mathbf{k}\to m\mathbf{k}+\mathbf{q}}}=\frac{1}{\tau^{\mathrm{ph}}_{n\mathbf{k}\to m\mathbf{k}+\mathbf{q}}}+\frac{1}{\tau^{\mathrm{imp}}_{n\mathbf{k}\to m\mathbf{k}+\mathbf{q}}},\tag{11}$$

where the superscripts refer to carrier-phonon (ph) and carrier-impurity (imp) scattering, respectively. In EPW, the carrier-impurity partial scattering rate $1/\tau^{\mathrm{imp}}_{n\mathbf{k}\to m\mathbf{k}+\mathbf{q}}$ is calculated under the following simplifying approximations: (i) each impurity is described by an idealized point charge, embedded in the dielectric continuum of the host material; (ii) the scattering rate from each impurity is evaluated within the first Born approximation; (iii) defects are sufficiently diluted that the scattering rates from different impurities are additive; (iv) impurities are randomly distributed, and this random distribution is formally taken into account by using the Kohn and Luttinger ensemble average[58]. Within these approximations, the charged impurity scattering rate reads[57]:

$$\frac{1}{\tau^{\mathrm{imp}}_{n\mathbf{k}\to m\mathbf{k}+\mathbf{q}}}=N_{\mathrm{imp}}\frac{2\pi}{\hbar}\left[\frac{e^2}{4\pi\epsilon_0}\frac{4\pi Z}{\Omega}\right]^2\sum_{\mathbf{G}\neq-\mathbf{q}}\frac{|\langle\psi_{m\mathbf{k}+\mathbf{q}}|e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}|\psi_{n\mathbf{k}}\rangle|^2}{|(\mathbf{q}+\mathbf{G})\cdot\boldsymbol{\epsilon}^0\cdot(\mathbf{q}+\mathbf{G})|^2}\delta(\varepsilon_{n\mathbf{k}}-\varepsilon_{m\mathbf{k}+\mathbf{q}}),\tag{12}$$

where N_{imp} is the number of impurities per crystal unit cell (dimensionless), Ze is the charge of each impurity, and $\boldsymbol{\epsilon}^0$ is the static relative dielectric constant tensor. The scattering rate in Eq. (12) takes into account the DFT electronic band structure and the lattice screening, including possible dielectric anisotropy. The main advantage of this model as compared to explicit calculations of charged defects in supercells is that it can be used systematically without requiring detailed knowledge of the defect physics and energetics in each material.

Computational considerations. In the EPW code, the scattering rate given in Eq. (12) is added to the electron–phonon scattering rate using Eqs. (11) and (10), and the *ai*BTE is solved as already described for the phonon-limited mobility. The computational overhead as compared to phonon-only calculations is minimal.

In Eq. (12), the sum over **G**-vectors is handled via a Gaussian filter, as described in the "Methods" section; the reader is referred to that section for more details. The overlap integrals between the initial and final Kohn–Sham states appearing in Eq. (12), $\langle\psi_{m\mathbf{k}+\mathbf{q}}|e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}|\psi_{n\mathbf{k}}\rangle$, are evaluated by means of the unitary Wannier function matrices described in the Methods.

The integral over the scattering wavevectors **q** in Eq. (10) of the impurity scattering rates given in Eq. (12) contains a singular $|\mathbf{q}|^{-4}$ term that is not integrable. This issue is resolved by introducing the screening of the defect potential by free carriers released upon ionization. To this end, we replace $\boldsymbol{\epsilon}^0$ in Eq. (12) with the total dielectric function:

$$\boldsymbol{\epsilon}^0_{\mathrm{tot}}=\boldsymbol{\epsilon}^0+\frac{q^2_{\mathrm{TF}}}{q^2}\boldsymbol{I},\tag{13}$$

where $\boldsymbol{I}$ denotes the identity matrix, and q_{TF} is the Thomas–Fermi wavenumber obtained from the long-wavelength limit of the Lindhard function[49,56,59]:

$$q^2_{\mathrm{TF}}=\frac{e^2}{4\pi\epsilon_0}\frac{4\pi}{\Omega}2\sum_n\int\frac{\mathrm{d}\mathbf{k}}{\Omega_{\mathrm{BZ}}}\left|\frac{\partial f^0_{n\mathbf{k}}}{\partial\varepsilon_{n\mathbf{k}}}\right|.\tag{14}$$

Temperature enters this expression via the equilibrium Fermi–Dirac distribution of the electrons or holes, $f^0_{n\mathbf{k}}$.

The concentration of charged defects is an external input parameter in these calculations. In the case of ionized impurities in semiconductors, it is also possible to use a simple estimate for the fraction f_{ii} of ionized impurities at a given temperature by using the impurity energy level in the gap ε_{d}[49,56,60].

$$f^2_{\mathrm{ii}}=\max\left(1,\frac{1}{n_d\,\Omega}\sum_n\int\frac{d\mathbf{k}}{\Omega_{\mathrm{BZ}}}f(\varepsilon_{n\mathbf{k}}-\varepsilon_{\mathrm{d}})\right),\tag{15}$$

where f is the Fermi–Dirac distribution and n_d is the total concentration of defects that may thermally ionize.

Owing to the $|\mathbf{q}|^{-4}$ scaling of the ionized-impurity scattering matrix elements, achieving convergence in the presence of impurities is slightly more demanding than for calculations of

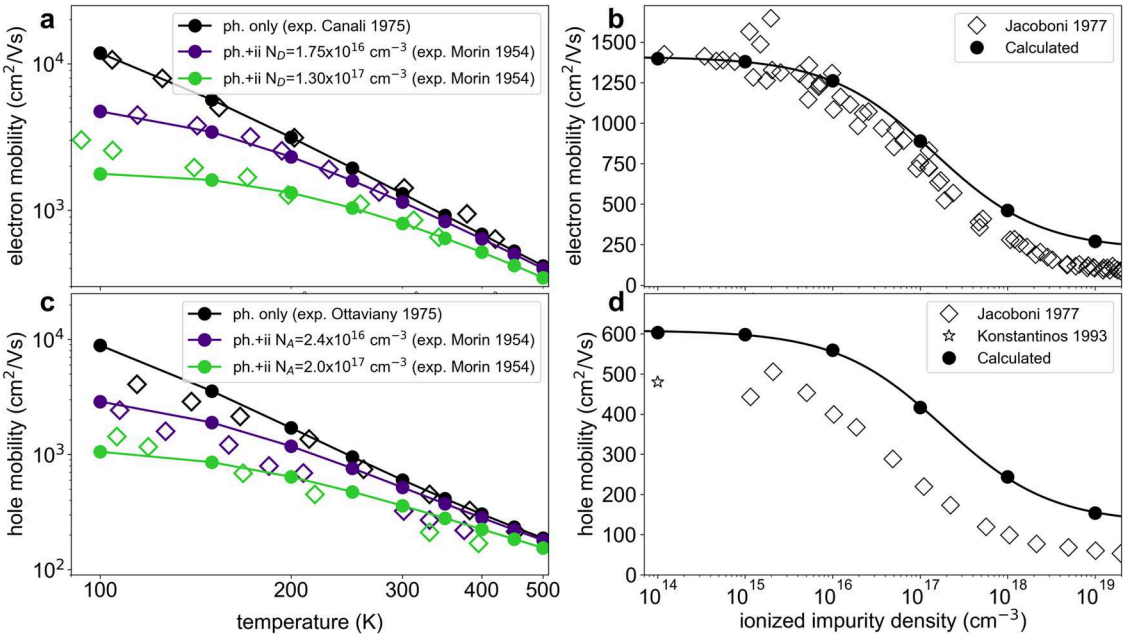

Fig. 2 Calculations of phonon- and impurity-limited carrier mobility using EPW. a Mobility of silicon as a function of temperature, for the following three cases: phonon scattering only (black disks and lines); phonon and ionized impurity scattering, with an impurity concentration of 1.75×10^{16} cm^{-3} (indigo disks and lines), and with a concentration of 1.3×10^{17} cm^{-3} (green disks and lines). Experimental data for the same concentrations are shown as diamonds of the same color[62,63,219]. **b** Same as in (**a**) but for the hole mobility of silicon. The impurity concentrations are 2.4×10^{16} cm^{-3} (indigo) and 2.0×10^{17} cm^{-3} (green), respectively. **c** Electron mobility of silicon at 300 K, as a function of ionized impurity concentration: black disks and lines are calculations, black diamonds are experiments[64,66]. **d** Same as in (**c**) but for the hole mobility.

phonon-limited mobility. For example, in recent calculations of carrier mobilities in Si, GaP, and SiC, fine grids with 100^3, 100^3, and 180^3 points were necessary, respectively[57] (see Table 1).

Application example. To illustrate the impurity scattering capability of EPW, we calculate the mobility of electrons and holes in silicon as a function of both temperature and dopant concentration. We consider coarse Brillouin zone grids with 12^3**k**-points and 6^3**q**-points, and fine grids with 100^3 points. We use the PBE functional[45] and ONCV pseudopotentials[46,61], and we include spin–orbit coupling for the valence bands. We account for quadrupole corrections using the procedure described in the Methods section.

Figure 2a shows the calculated electron mobility of silicon as a function of temperature. When considering phonon scattering only, our calculations are in very good with measurements on high-purity silicon[62]. Upon introducing ionized impurity scattering, the electron mobility at 100 K decreases from 11,813 to 4725 cm^2 V^{-1} s^{-1} for a dopant concentration of 1.75×10^{16} cm^{-3}, and to 1769 cm^2 V^{-1} s^{-1} for a dopant concentration of 1.3×10^{17} cm^{-3}. These calculations agree well with experimental data[63]. At higher temperatures, the reduction in the mobility due to impurity scattering is less significant since phonons provide the dominant scattering mechanism. The hole mobility in Fig. 2b exhibits a similar trend. Upon introducing ionized impurity scattering, the hole mobility at 100 K decreases from 8877 to 2884 cm^2 V^{-1} s^{-1} for a dopant concentration of 2.4×10^{16} cm^{-3}, and to 1056 cm^2 V^{-1} s^{-1} for a dopant concentration of 2.0×10^{17} cm^{-3}. Also in this case, the agreement with experimental data is very good[63].

Figure 2c shows the calculated electron mobility of silicon at 300 K, as a function of ionized impurity concentration. Up to a dopant concentration of 10^{16} cm^{-3}, phonon scattering dominates and the mobility is relatively insensitive to impurity scattering. Beyond this concentration, the mobility decreases sharply and approximately as the inverse of the impurity density, in line with Eq. (12). The calculations agree well with experiments up to an impurity concentration around 10^{18} cm^{-3}[64]. Beyond this concentration, it is expected that additional mechanisms such as two-impurity scattering and plasmon scattering[65] will further reduce the mobility. The hole mobility, which is shown in Fig. 2d, follows a similar trend. The calculated hole mobility at low doping, 603 cm^2 V^{-1} s^{-1}, slightly overestimates the experimental range 450–500 cm^2 V^{-1} s^{-1}; this effect can be traced to the underestimation of the heavy hole mass by DFT[31]. Upon increasing the impurity concentration, the mobility decreases following the same trend as for the electrons. These calculations agree well with experiments[64,66] and with previous first-principles calculations[56].

Phonon-mediated superconductivity using the ab initio Eliashberg theory

Background and formalism. First-principles calculations of phonon-mediated superconductors are primarily based on three approaches, namely semi-empirical methods based on the McMillan formula[67], the ab initio Eliashberg theory[68,69], and the density-functional theory for superconductors[70–73]. The EPW code implements the former two approaches. In this section, we briefly review methods based on the McMillan formula, and then we describe the Eliashberg formalism and its implementation.

The superconducting critical temperature T_c can be estimated using the semi-empirical McMillan equation in Allen–Dynes form[74]:

$$k_B T_c^{AD} = \frac{\hbar \omega_{log}}{1.2} \exp\left[\frac{-1.04(1+\lambda)}{\lambda - \mu^*(1+0.62\lambda)}\right], \quad (16)$$

where k_B is Boltzmann's constant, μ^* is the semi-empirical Coulomb pseudopotential, ω_{log} is the logarithmic average of the phonon frequencies, and λ is the electron–phonon coupling constant, as defined in ref. [1]. The parameter μ^* is largely material independent, for a large class of superconductors a value in the 0.1-0.15 range being typically required to fit the experimental

T_c[74–76]. Note that for very strong coupling ($\lambda > 1.5$), in ref. [74] two correction prefactors f_1 and f_2 are introduced in Eq. (16) to better estimate the critical temperature in this regime.

In addition, the code allows the user to calculate the critical temperature using a more recent prescription based on a machine learning approach[77]:

$$T_c^{\mathrm{ML}} = f_\omega f_\mu T_c^{\mathrm{AD}}, \tag{17}$$

where the correction factors f_ω and f_μ are given by

$$f_\omega = 1.92\frac{\lambda + \omega_{\log}/\overline{\omega}_2 - (\mu^*)^{1/3}}{\lambda^{1/2}\exp(\omega_{\log}/\overline{\omega}_2)} - 0.08, \qquad f_\mu = \frac{6.86\exp(-\lambda/\mu^*)}{\lambda^{-1} - \mu^* - \omega_{\log}/\overline{\omega}_2} + 1. \tag{18}$$

In these expressions, $\overline{\omega}_2$ is the square root of the second moment of the normalized weight function $g(\omega) = 2\alpha^2F(\omega)/(\omega\lambda)$, with α^2F being the Eliashberg spectral function, as defined in ref. [1]. Equations (16) and (17) are useful for preliminary calculations but do not carry the predictive power of the ab initio Eliashberg theory.

The Eliashberg theory[68,69] describes the superconducting phase transition by means of finite-temperature Green's functions. In this theory, superconducting pairing arises from an attractive electron–electron interaction mediated by phonons, which is partly compensated by the inter-electron Coulomb repulsion. The Eliashberg theory can be formulated as a Dyson equation for a generalized 2×2 matrix Green's function via the Nambu-Gor'kov formalism[78,79]. The off-diagonal elements of this matrix describe Cooper-pair amplitudes in the superconducting state and are related to the superconducting gap function. These elements become nonzero below the critical temperature, marking the transition to the superconducting state. It is standard practice to expand the pairing self-energy using Pauli matrices $\hat{\tau}_i$ ($i = 0, \ldots, 3$) as follows[75,80–86]:

$$\hat{\Sigma}_{n\mathbf{k}}(i\omega_j) = i\hbar\omega_j\left[1 - Z_{n\mathbf{k}}(i\omega_j)\right]\hat{\tau}_0 + \chi_{n\mathbf{k}}(i\omega_j)\hat{\tau}_3 + \phi_{n\mathbf{k}}(i\omega_j)\hat{\tau}_1, \tag{19}$$

where $i\omega_j = i(2j+1)\pi T$ is the fermionic Matsubara frequency with j being an integer, T is the absolute temperature, $Z_{n\mathbf{k}}(i\omega_j)$ is the mass renormalization function, $\chi_{n\mathbf{k}}(i\omega_j)$ is the energy shift, and $\phi_{n\mathbf{k}}(i\omega_j)$ is the order parameter. This self-energy is expressed in terms of the electron Green's function using the Migdal approximation[87] for the electron–phonon contribution, and the GW approximation for the electron–electron contribution[88,89]. Using Eq. (19) inside the Dyson equation for the electron Green's function yields a set of coupled equations for $Z_{n\mathbf{k}}$, $\chi_{n\mathbf{k}}$, and $\phi_{n\mathbf{k}}$:

$$Z_{n\mathbf{k}}(i\omega_j) = 1 + \frac{k_BT}{\omega_j N(\varepsilon_F)}\sum_{mj'}\int\frac{\mathrm{d}\mathbf{q}}{\Omega_{\mathrm{BZ}}}\frac{\omega_{j'}Z_{m\mathbf{k}+\mathbf{q}}(i\omega_{j'})}{\theta_{m\mathbf{k}+\mathbf{q}}(i\omega_{j'})}\lambda(n\mathbf{k}, m\mathbf{k}+\mathbf{q}, \omega_j - \omega_{j'}), \tag{20}$$

$$\chi_{n\mathbf{k}}(i\omega_j) = -\frac{k_BT}{N(\varepsilon_F)}\sum_{mj'}\int\frac{\mathrm{d}\mathbf{q}}{\Omega_{\mathrm{BZ}}}\frac{\varepsilon_{m\mathbf{k}+\mathbf{q}} - \mu_F + \chi_{m\mathbf{k}+\mathbf{q}}(i\omega_{j'})}{\theta_{m\mathbf{k}+\mathbf{q}}(i\omega_{j'})}\lambda(n\mathbf{k}, m\mathbf{k}+\mathbf{q}, \omega_j - \omega_{j'}), \tag{21}$$

$$\phi_{n\mathbf{k}}(i\omega_j) = \frac{k_BT}{N(\varepsilon_F)}\sum_{mj'}\int\frac{\mathrm{d}\mathbf{q}}{\Omega_{\mathrm{BZ}}}\frac{\phi_{m\mathbf{k}+\mathbf{q}}(i\omega_{j'})}{\theta_{m\mathbf{k}+\mathbf{q}}(i\omega_{j'})}\left[\lambda(n\mathbf{k}, m\mathbf{k}+\mathbf{q}, \omega_j - \omega_{j'}) - N(\varepsilon_F)V_{n\mathbf{k},m\mathbf{k}+\mathbf{q}}\right], \tag{22}$$

having defined the auxiliary function:

$$\theta_{n\mathbf{k}}(i\omega_j) = \left[\hbar\omega_j Z_{n\mathbf{k}}(i\omega_j)\right]^2 + \left[\varepsilon_{n\mathbf{k}} - \mu_F + \chi_{n\mathbf{k}}(i\omega_j)\right]^2 + \left[\phi_{n\mathbf{k}}(i\omega_j)\right]^2. \tag{23}$$

In Eqs. (20)–(23), $N(\varepsilon_F)$ is the density of states (DOS) per spin at the Fermi level, μ_F is the chemical potential, and the quantities $V_{n\mathbf{k},m\mathbf{k}+\mathbf{q}}$ denote the matrix elements of the screened Coulomb interaction W between electron pairs, as given in refs. [90,91]. The anisotropic electron–phonon coupling parameters $\lambda(n\mathbf{k}, m\mathbf{k}+\mathbf{q}, \omega_j - \omega_{j'})$ appearing in these equations are calculated as

$$\lambda(n\mathbf{k}, m\mathbf{k}+\mathbf{q}, \omega_j - \omega_{j'}) = \frac{N(\varepsilon_F)}{\hbar}\sum_\nu |g_{mn\nu}(\mathbf{k}, \mathbf{q})|^2\frac{2\omega_{\mathbf{q}\nu}}{(\omega_j - \omega_{j'})^2 + \omega_{\mathbf{q}\nu}^2}. \tag{24}$$

Equations (20)–(23) are supplemented by a statement of particle number conservation, which determines the chemical potential[84]:

$$N_e = \sum_n\int\frac{\mathrm{d}\mathbf{k}}{\Omega_{\mathrm{BZ}}}\left[1 - 2k_BT\sum_j\frac{\varepsilon_{n\mathbf{k}} - \mu_F + \chi_{n\mathbf{k}}(i\omega_j)}{\theta_{n\mathbf{k}}(i\omega_j)}\right], \tag{25}$$

where N_e is the number of electrons per unit cell. Equations (20)–(25) are referred to as the anisotropic full-bandwidth (FBW) Eliashberg equations[92] since they explicitly take into account scattering processes involving electrons with energies and momenta that are not restricted to the vicinity of the Fermi surface. To find the temperature-dependent superconducting gap, in EPW these equations are solved iteratively for different temperatures; the highest temperature for which a non-trivial solution exists ($\phi_{n\mathbf{k}} \neq 0$) is the superconducting critical temperature T_c. Below this temperature, the superconducting gap $\Delta_{n\mathbf{k}}$ is given by

$$\Delta_{n\mathbf{k}}(i\omega_j) = \phi_{n\mathbf{k}}(i\omega_j)/Z_{n\mathbf{k}}(i\omega_j). \tag{26}$$

The numerical solution of Eqs. (20)–(25) is computationally demanding. A common simplification of these equations consists of restricting the energy range close to the Fermi level[75,80–86,93]. In this approach, it is assumed that the DOS within this energy window is constant. It can be shown that, within these approximations, the energy shift $\chi_{n\mathbf{k}}$ vanishes and the requirement in Eq. (25) is automatically satisfied. As a result, only two equations for $Z_{n\mathbf{k}}$ and $\phi_{n\mathbf{k}}$ need to be solved self-consistently:

$$Z_{n\mathbf{k}}(i\omega_j) = 1 + \frac{\pi k_BT}{N(\varepsilon_F)\omega_j}\sum_{mj'}\int\frac{\mathrm{d}\mathbf{q}}{\Omega_{\mathrm{BZ}}}\frac{\omega_{j'}}{\sqrt{\hbar^2\omega_{j'}^2 + \Delta_{m\mathbf{k}+\mathbf{q}}^2(i\omega_{j'})}} \times\lambda(n\mathbf{k}, m\mathbf{k}+\mathbf{q}, \omega_j - \omega_{j'})\delta(\varepsilon_{m\mathbf{k}+\mathbf{q}} - \varepsilon_F), \tag{27}$$

$$Z_{n\mathbf{k}}(i\omega_j)\Delta_{n\mathbf{k}}(i\omega_j) = \frac{\pi k_BT}{N(\varepsilon_F)}\sum_{mj'}\int\frac{\mathrm{d}\mathbf{q}}{\Omega_{\mathrm{BZ}}}\frac{\Delta_{m\mathbf{k}+\mathbf{q}}(i\omega_{j'})}{\sqrt{\hbar^2\omega_{j'}^2 + \Delta_{m\mathbf{k}+\mathbf{q}}^2(i\omega_{j'})}} \times\left[\lambda(n\mathbf{k}, m\mathbf{k}+\mathbf{q}, \omega_j - \omega_{j'}) - N(\varepsilon_F)V_{n\mathbf{k},m\mathbf{k}+\mathbf{q}}\right]\delta(\varepsilon_{m\mathbf{k}+\mathbf{q}} - \varepsilon_F). \tag{28}$$

These equations are referred to as the anisotropic Fermi surface restricted (FSR) Eliashberg equations[18,85].

To extract physical quantities of interest, such as the tunneling density of states and the heat capacity, the gap function on the real frequency axis is required[75,81,82]. In the EPW code, the continuation of $\Delta_{n\mathbf{k}}(i\omega_j)$ from the imaginary to the real axis is performed either using Padé approximants[94,95], or using the exact analytic continuation of ref. [96], as described in ref. [85].

Accounting for the Coulomb repulsion in the Eliashberg equations requires the evaluation of the matrix elements $V_{n\mathbf{k},m\mathbf{k}+\mathbf{q}}$. While this is feasible in principle, it is common practice to replace the product $N(\varepsilon_F)V_{n\mathbf{k},m\mathbf{k}+\mathbf{q}}$ with the semi-empirical Morel–Anderson pseudopotential μ_c^*[97]. In this approach, the value of μ_c^* is linked to the cutoff frequency ω_c for the Matsubara summation on the imaginary axis, with ω_c typically set to ten times the maximum phonon frequency ω_{ph}. Note that when converting between the μ^* in Allen–Dynes formula in Eq. (16) and μ_c^*, a ratio $\omega_c/\omega_{\mathrm{ph}}$ of ten is often used[74]. In EPW, μ_c^* is specified by the user as an external parameter, and in many applications, values in the range $\mu_c^* = 0.1$-0.2 yield reasonable agreement with experiments. More recently, first-principles calculations of μ_c^* as a Fermi surface average of $V_{n\mathbf{k},m\mathbf{k}+\mathbf{q}}$ have been used in conjunction with EPW[91,98–100].

In the case of simple superconductors which do not exhibit significant anisotropy, it may be a good approximation to neglect the band and momentum dependence of the superconducting gap. In

these cases, instead of Eqs. (20)–(25), it is possible to solve a simplified version of the equations where all properties are averaged over the Fermi surface[76,101–103]. The relations are referred to as the isotropic Eliashberg equations:

$$Z(i\omega_j) = 1 + \frac{k_B T}{N(\varepsilon_F)\omega_j} \int d\varepsilon N(\varepsilon) \sum_{j'} \frac{\omega_{j'} Z(i\omega_{j'})}{\theta(\varepsilon, i\omega_{j'})} \lambda(\omega_j - \omega_{j'}), \tag{29}$$

$$\chi(i\omega_j) = -\frac{k_B T}{N(\varepsilon_F)} \int d\varepsilon N(\varepsilon) \sum_{j'} \frac{\varepsilon - \mu_F + \chi(i\omega_{j'})}{\theta(\varepsilon, i\omega_{j'})} \lambda(\omega_j - \omega_{j'}), \tag{30}$$

$$\phi(i\omega_j) = \frac{k_B T}{N(\varepsilon_F)} \int d\varepsilon N(\varepsilon) \sum_{j'} \frac{\phi(i\omega_{j'})}{\theta(\varepsilon, i\omega_{j'})} \left[\lambda(\omega_j - \omega_{j'}) - \mu_c^*\right], \tag{31}$$

$$N_e = \int d\varepsilon N(\varepsilon) \left[1 - 2k_B T \sum_j \frac{\varepsilon - \mu_F + \chi(i\omega_j)}{\theta(\varepsilon, i\omega_j)}\right], \tag{32}$$

where the counterpart of Eq. (23) is

$$\theta(\varepsilon, i\omega_j) = \left[\hbar\omega_j Z(i\omega_j)\right]^2 + \left[\varepsilon - \mu_F + \chi(i\omega_j)\right]^2 + \left[\phi(i\omega_j)\right]^2. \tag{33}$$

In Eqs. (29)–(31), $\lambda(\omega_j - \omega_{j'})$ is a momentum-averaged version of Eq. (24):

$$\lambda(\omega_j - \omega_{j'}) = \frac{1}{[N(\varepsilon_F)]^2} \sum_{m,n} \int \frac{d\mathbf{k}}{\Omega_{BZ}} \int \frac{d\mathbf{q}}{\Omega_{BZ}} \lambda(n\mathbf{k}, m\mathbf{k} + \mathbf{q}, \omega_j - \omega_{j'}) \delta(\varepsilon_{n\mathbf{k}} - \varepsilon_F) \delta(\varepsilon_{m\mathbf{k}+\mathbf{q}} - \varepsilon_F). \tag{34}$$

Similar to the anisotropic case, Eqs. (29)–(31) can also be reduced to a set of two equations for $Z(i\omega_j)$ and $\phi(i\omega_j)$, under the same conditions as discussed for Eqs. (27) and (28).

Computational considerations. Solving Eqs. (20)–(25) or the simplified version Eqs. (27) and (28) is computationally demanding because a fine sampling of electron-phonon processes near the Fermi surface is required. In addition, due to the implicit form of these equations, the **k**- and **q**-point grids must be commensurate. To accelerate calculations, EPW exploits crystal symmetries so that the gap function and related quantities are only computed for **k**-points in the irreducible wedge of the Brillouin zone, and only electronic states within a small energy window centered at the Fermi energy are considered. Numerical tests show that converged results are obtained when this energy window is of the order of a few times the maximum phonon energy. In general, grids with an average point density $\Delta\mathbf{k} \sim 0.02$ Å^{-1} and $\Delta\mathbf{q} \sim 0.04$ Å^{-1} are sufficient to converge the superconducting gap[85,91,104], but denser meshes may be required for superconductors with complex Fermi surfaces. As an example, 12^2 and 40^3**k**-point grids and 6^2 and 20^3**q**-point grids were used to describe the two-gap structure of bilayer C_6CaC_6 and the single-gap structure of bulk CaC_6, respectively[91] (see Table 1).

One further complication in the solution of the Eliashberg equations is that all quantities depend on Matsubara frequencies. The Matsubara frequencies are proportional to the absolute temperature, therefore superconductors with low T_c require a larger number of frequencies within the same energy range. While it is common to set the Matsubara frequency cutoff to ten times the maximum phonon frequency[85,86,101,105,106], we recommend always performing convergence tests with respect to this parameter. In EPW, the number of Matsubara frequencies can be reduced using a sparse sampling scheme whereby only a subset of frequencies is retained with a density that decreases logarithmically with the Matsubara index. The grid density is controlled by an adjustable parameter; using the default setting for this parameter, EPW yields a ~30% reduction of Matsubara frequencies, while keeping all the lowest ~40% of frequencies. Numerical tests show that this approach preserves the accuracy of more expensive full-grid calculations.

FBW Eliashberg calculations are computationally more demanding than FSR calculations. In particular, to determine the chemical potential from Eq. (25) one needs a Matsubara frequency cutoff of at least a few times the Kohn–Sham energy window, leading to a considerable increase in computational cost. To circumvent this issue, the default behavior of EPW is to keep the chemical potential fixed at the Fermi level unless it is instructed by the user to solve Eq. (25). In this case, careful convergence tests with respect to the Matsubara frequency cutoff are warranted.

Application example. To demonstrate the implementation of the superconducting module in EPW, we investigate the superconducting properties of hexagonal niobium diselenide (2H-$NbSe_2$). 2H-$NbSe_2$ exhibits a superconducting phase transition and a charge density wave instability below 7.2 and 33 K, respectively; both phases are driven by a strong momentum-dependent electron-phonon coupling[107–112]. For brevity, in this section, we focus on superconductivity and ignore the lattice instability.

We employ the PBE exchange and correlation functional[45] and the optB86b-vdW van der Waals functional[113–116], as well as ONCV pseudopotentials[46,61] including semi-core electrons for Nb. We use a plane waves kinetic energy cutoff of 80 Ry, Γ-centered $24 \times 24 \times 12$ **k**-point and $6 \times 6 \times 4$ **q**-point coarse grids, and Methfessel–Paxton smearing[117] of 0.025 Ry. The optimized lattice parameters are $a = 3.46$ Å and $c = 12.57$ Å, in good agreement with the experimental data $a = 3.43$ Å and $c = 12.55$ Å[108].

The Eliashberg equations are solved on a uniform $60 \times 60 \times 40$ **k**-point and $30 \times 30 \times 20$ **q**-point fine grids. We consider energy windows of 0.4 and 0.8 eV, Matsubara frequency cutoffs of 0.4, 0.8, and 1.2 eV, and a Coulomb pseudopotential of $\mu_c^* = 0.2$. The Dirac deltas in the FSR approach are replaced by Gaussians of width 50 meV.

Figure 3a shows the band structure and DOS of 2H-$NbSe_2$. Three bands cross the Fermi level. This system exhibits soft phonons along the Γ*M* direction[109,111,112], which are stabilized by anharmonic effects[109]. Here, for simplicity, we avoid soft phonons by slightly increasing the electronic smearing to 0.025 Ry in phonon calculations, as shown in Fig. 3b. The same panel also shows the distribution of the electron–phonon coupling strength as described by the Eliashberg spectral function α^2F. Figure 3c shows the anisotropic gap function color-coded on the Fermi surface of 2H-$NbSe_2$. We find a highly anisotropic two-gap structure: the high-gap region corresponds to the Fermi arcs around the K point, and the low-gap region corresponds to the Γ-centered pockets. These results are similar to previous calculations for the related compound NbS_2[98]. Figure 3d shows the energy distribution of the superconducting gap function $\Delta_{n\mathbf{k}}$ for various temperatures. The gap is seen to close gradually with increasing temperature, vanishing at the critical temperature $T_c = 19$ K. The calculated critical temperature overestimates the experimental value of 7.2 K, consistent with prior calculations for bulk and monolayer transition metal dichalcogenides[98,109,112,118–120]. The calculations in Fig. 3d are performed within the FBW Eliashberg approach, with the chemical potential set to the Fermi energy. Calculated gap values for energy windows of 0.4 and 0.8 eV are very similar, therefore in the following we only consider the former. Figure 3e shows the sensitivity of the superconducting gap to the chemical potential: by requiring particle number conservation via Eq. (25), the chemical potential shifts by 25 meV with respect to the Fermi energy, and the critical temperature slightly increases, by <1 K. In Fig. 3f we perform the same test, but this time with a wider energy window, obtaining similar results.

In Fig. 3g we also compare the FBW Eliashberg approach with the FSR approach, which is approximately 30% faster for the system under consideration. The critical temperatures calculated from these approaches agree within 0.5 K. This finding relates to the DOS of 2H-$NbSe_2$ being slowly varying within a few hundred meV's around the Fermi energy. Larger differences are expected for systems with rapidly varying DOS, such as for example the high-T_c hydride superconductors[105].

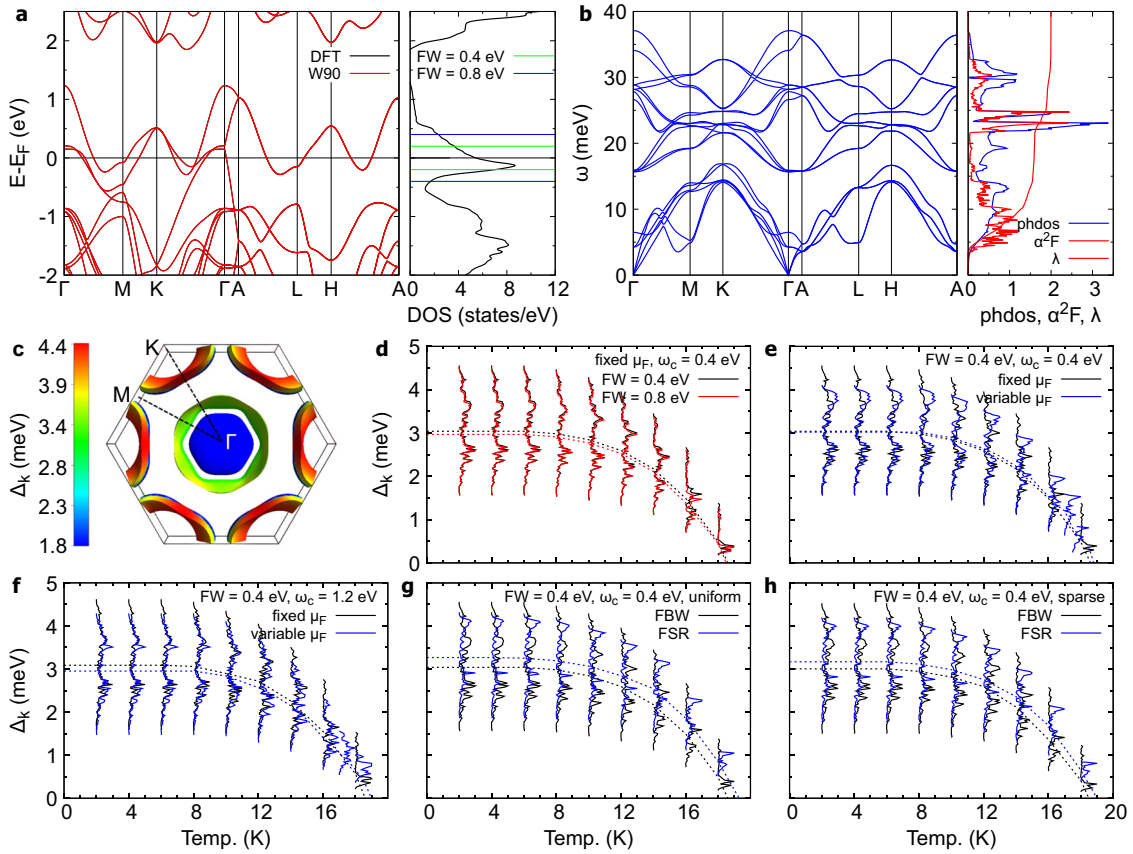

Fig. 3 Eliashberg calculations for phonon-mediated superconductors in EPW. **a** Band structure and DOS of 2H-$NbSe_2$. **b** Phonon dispersion relations, phonon density of states (blue line), Eliashberg spectral function α^2F (solid red line), and cumulative electron-phonon coupling strength λ (dashed red line) of 2H-$NbSe_2$. In (**a**), the bands obtained by Wannier interpolation are shown in dashed red lines. **c** Momentum-resolved superconducting gap of 2H-$NbSe_2$, color-coded on the Fermi surface[220] and evaluated at 2 K [same dataset as in (**d**)]. **d** Energy distribution of the superconducting gap of 2H-$NbSe_2$ as a function of temperature, calculated using the FBW Eliashberg approach, chemical potential set to the Fermi energy, and a Matsubara frequency cutoff of 0.4 eV. The energy window is 0.4 eV for the black line, and 0.8 eV for the red line. **e** Gap function of 2H-$NbSe_2$ obtained from the FBW approach using fixed chemical potential (black line) and variable chemical potential (blue line). The energy window and Matsubara cutoff are both 0.4 eV. **f** Same as in (**e**), but using a Matsubara cutoff of 1.2 eV. **g** Comparison between the gap functions obtained with the FBW approach (black line) and the FSR approach (blue). Both the energy window and the Matsubara cutoff are 0.4 eV. **h** Same as in (**g**), but using a sparse Matsubara frequency grid. The dashed lines in **d**–**h** represent averages of the gap distributions.

When using a logarithmic Matsubara frequency grid, the results are essentially unaffected as shown in Fig. 3h. At the same time, the sparse Matsubara grids afford a significant computational saving of almost a factor of two.

Polarons and electron self-trapping

Background and formalism. An important manifestation of electron-phonon interactions is the formation of polarons. The polaron is a quasiparticle consisting of an electron or a hole dressed by a distortion of the crystal lattice. The lattice distortion can be interpreted as a cloud of virtual phonons accompanying the electron or hole. In the presence of strong electron–phonon coupling, the electron or hole can become trapped in the lattice distortion that it produced; this configuration is referred to as a self-trapped polaron[121–123].

In the context of first-principles calculations, the formation of polarons has generally been investigated by means of DFT calculations of an excess charge added to a large supercell. When the calculation is initialized with a localized distortion of the crystal lattice, it may be energetically favorable for the electron or hole to localize around this distortion, leading to the formation of a polaron[124–130]. This approach faces two challenges: first, the existence of polaronic states and their formation energy are very sensitive to the choice of the exchange and correlation functional; second, the size of the supercells required to obtain converged energies and wavefunctions may be prohibitive.

In the EPW code, these challenges are overcome by recasting the polaron problem into the solution of a coupled nonlinear system of equations for the energy, wavefunction, and atomic displacements associated with the polaron[131–134]. The founding principle of this methodology is that the DFT formation energy of the polaron can be expressed as a self-interaction-free functional of the polaron wavefunction $\psi(\mathbf{r})$ and the atomic displacements in the polaronic state $\Delta\tau_{\kappa ap}$[132] (all symbols are defined in the 'Methods' section):

$$\Delta E_f = \int d\mathbf{r}\, \psi^*(\mathbf{r})\hat{H}^0_{KS}\psi(\mathbf{r}) + \sum_{\kappa ap}\int d\mathbf{r}\, \frac{\partial V^0_{KS}}{\partial \tau_{\kappa ap}}|\psi(\mathbf{r})|^2\Delta\tau_{\kappa ap} + \frac{1}{2}\sum_{\substack{\kappa ap\\ \kappa' a' p'}} C^0_{\kappa ap,\kappa' a' p'}\Delta\tau_{\kappa ap}\Delta\tau_{\kappa' a' p'}, \quad (35)$$

where $\hat{H}^0_{KS}$ and $C^0_{\kappa ap,\kappa' a' p'}$ are the Kohn–Sham Hamiltonian and the matrix of interatomic force constants in the ground-state structure without the polaron, and the integrals are over the Born–von Kármán (BvK) supercell. Variational minimization of this energy

functional leads to the coupled system of equations:

$$\hat{H}^0_{\mathrm{KS}}\,\psi(\mathbf{r}) + \sum_{\kappa a p} \frac{\partial V^0_{\mathrm{KS}}(\mathbf{r})}{\partial \tau_{\kappa a p}} \Delta\tau_{\kappa a p}\,\psi(\mathbf{r}) = \varepsilon\,\psi(\mathbf{r})\,, \tag{36}$$

$$\Delta\tau_{\kappa a p} = -\sum_{\kappa' a' p'} (C^0)^{-1}_{\kappa a p, \kappa' a' p'} \int \mathrm{d}\mathbf{r}\, \frac{\partial V^0_{\mathrm{KS}(\mathbf{r})}}{\partial \tau_{\kappa' a' p'}} |\psi(\mathbf{r})|^2\,, \tag{37}$$

where V^0_{KS} is the Kohn–Sham potential in the ground-state structure without polaron, and ε represents the quasiparticle excitation energy of the polaron[133,134]. To avoid performing calculations in large supercells, in EPW Eqs. (36) and (37) are rewritten more conveniently in terms of Kohn–Sham states, phonons, and electron–phonon matrix elements evaluated in the primitive unit cell of the crystal. To this aim, the wavefunction is expanded on the basis of Kohn–Sham states:

$$\psi(\mathbf{r}) = \frac{1}{\sqrt{N_{\mathrm{uc}}}} \sum_{n\mathbf{k}} A_{n\mathbf{k}} \psi_{n\mathbf{k}}\,, \tag{38}$$

and the atomic displacements are expanded on the basis of lattice vibrational eigenmodes:

$$\Delta\tau_{\kappa a p} = -\frac{2}{N_{\mathrm{uc}}} \sum_{\mathbf{q}\nu} B^*_{\mathbf{q}\nu} \sqrt{\frac{\hbar}{2M_\kappa \omega_{\mathbf{q}\nu}}}\, \mathrm{e}_{\kappa a,\nu}(\mathbf{q}) \mathrm{e}^{i\mathbf{q}\cdot\mathbf{R}_p}. \tag{39}$$

With these definitions, Eqs. (36) and (37) become a nonlinear system of equations for the polaron coefficients $A_{n\mathbf{k}}$ and $B_{\mathbf{q}\nu}$:

$$\frac{2}{N_{\mathrm{uc}}} \sum_{\mathbf{q}m\nu} B_{\mathbf{q}\nu}\, g^*_{mn\nu}(\mathbf{k},\mathbf{q})\, A_{m\mathbf{k}+\mathbf{q}} = (\varepsilon_{n\mathbf{k}} - \varepsilon) A_{n\mathbf{k}}\,, \tag{40}$$

$$B_{\mathbf{q}\nu} = \frac{1}{N_{\mathrm{uc}}} \sum_{mn\mathbf{k}} A^*_{m\mathbf{k}+\mathbf{q}} \frac{g_{mn\nu}(\mathbf{k},\mathbf{q})}{\hbar\omega_{\mathbf{q}\nu}} A_{n\mathbf{k}}\,. \tag{41}$$

These relations are referred to as the ab initio polaron equations. The polaron formation energy in Eq. (35) can be expressed in terms of the polaron coefficients $A_{n\mathbf{k}}$ and $B_{\mathbf{q}\nu}$ as[132]

$$\Delta E_{\mathrm{f}} = \frac{1}{N_{\mathrm{uc}}} \sum_{n\mathbf{k}} |A_{n\mathbf{k}}|^2 (\varepsilon_{n\mathbf{k}} - \varepsilon_{\mathrm{CBM}}) - \frac{1}{N_{\mathrm{uc}}} \sum_{\mathbf{q}\nu} |B_{\mathbf{q}\nu}|^2 \hbar\omega_{\mathbf{q}\nu}\,, \tag{42}$$

where $\varepsilon_{\mathrm{CBM}}$ is the Kohn–Sham eigenvalue of the conduction band minimum. This expression holds for electron polarons; in the case of hole polarons, $\varepsilon_{\mathrm{CBM}}$ is replaced by the eigenvalue of the valence band maximum, and the first term on the right-hand side of Eq. (42) acquires a minus sign. Equations (38), (39), and (42) provide the real-space wavefunction of the polaron, the accompanying atomic displacements, and its formation energy.

Computational considerations. The solution of Eqs. (40) and (41) require the knowledge of the Kohn-Sham energies $\varepsilon_{n\mathbf{k}}$, the phonon frequencies $\omega_{\mathbf{q}\nu}$, and the electron-phonon matrix elements $g_{mn\nu}(\mathbf{k},\mathbf{q})$. All of these quantities are calculated by EPW using Wannier–Fourier interpolation as described in the Methods. The solution of the coupled system of equations is performed using an iterative procedure as follows. First, $A_{n\mathbf{k}}$ is initialized in order to compute $B_{\mathbf{q}\nu}$ via Eq. (41). Then, Eq. (40) is solved by constructing and diagonalizing the effective Hamiltonian

$$H_{n\mathbf{k},n'\mathbf{k}'} = \varepsilon_{n\mathbf{k}}\delta_{n\mathbf{k},n'\mathbf{k}'} - \frac{2}{N_{\mathrm{uc}}} \sum\nolimits_{\nu} B_{\mathbf{k}'-k\nu}\, g^*_{n'n\nu}(\mathbf{k},\mathbf{k}'-k)\,. \tag{43}$$

The set of solution coefficients $A_{n\mathbf{k}}$ is used again in Eq. (41), and the process is repeated until convergence is achieved.

The initialization of the coefficients $A_{n\mathbf{k}}$ is achieved by using a Gaussian function in reciprocal space centered around the band extremum. Alternatively, a polaron solution from a previous calculation can be used. In the construction of the effective Hamiltonian matrix in Eq. (43), we set the gauge condition $e_{\kappa a,\nu}(-\mathbf{q}) = e^*_{\kappa a,\nu}(\mathbf{q})$[135]. This condition ensures that $B_{-\mathbf{q}\nu} = B^*_{\mathbf{q}\nu}$, so only half of the $B_{\mathbf{q}\nu}$ coefficients need to be evaluated at each iteration. In addition, for parallel execution, we set a global gauge for all interpolated Kohn–Sham wavefunctions and vibrational eigenmodes. The diagonalization of the effective Hamiltonian is performed via the Davidson method[136] as implemented in Quantum ESPRESSO. The convergence of the iterative procedure is tested by evaluating the atomic displacements in real space via Eq. (39) and comparing these real-valued vectors between successive iterations. In the case of polaron calculations, choosing a fine Brillouin zone grid with $N\times N\times N$ points implies that one is studying polarons in a periodic BvK supercell consisting of $N\times N\times N$ unit cells. The formation energy of an isolated polaron is obtained by taking the limit of $N\to\infty$. For large enough N, we observe that the formation energy scales as $1/N$, which is consistent with the Makov–Payne asymptotic expansion of the energy of a point charge in a periodic cell[132,137]. Based on this observation, to obtain the energy of an isolated polaron, we perform a Makov-Payne extrapolation to the limit of infinite supercell size. Generally speaking, we find that for small polarons the wavefunction shape is converged already for BvK supercells of the order of $5\times5\times5$ unit cells, while for large polarons it is necessary to use supercells that are large enough to fully contain the polaron density. While there is no formal upper limit to this size, the largest BvK supercell sizes that we employed are up to $50\times50\times50$ for the large electron polaron in NaCl (see Table 1).

To visualize the polaron, EPW expresses the wavefunction as a linear combination of maximally localized Wannier functions:

$$\psi(\mathbf{r}) = \sum_{mp} A_{mp} \mathrm{w}_m(\mathbf{r} - \mathbf{R}_p), \tag{44}$$

where the coefficients A_{mp} are obtained from the solution vectors $A_{n\mathbf{k}}$ via the Wannier–Fourier transformation:

$$A_{mp} = \frac{1}{N_{\mathrm{uc}}} \sum_{n\mathbf{k}} \mathrm{e}^{i\mathbf{k}\cdot\mathbf{R}_p} U^\dagger_{mn\mathbf{k}} A_{n\mathbf{k}}\,, \tag{45}$$

and $U_{mn\mathbf{k}}$ is the unitary matrix that generates the smooth Bloch gauge [Eq. (64) of the "Methods" section]. The inversion of this last relation also allows one to interpolate the coefficients $A_{n\mathbf{k}}$ throughout the Brillouin zone, which is useful to visualize how specific bands contribute to the polaron wavefunction (Fig. 4).

Application example. To demonstrate the implementation of the polaron module in EPW, we investigate the electron and hole polarons in rocksalt NaCl, a prototypical polar insulator that hosts small hole polarons called V_K centers[138,139]. We perform calculations using the PBE exchange and correlation functional[45], ONCV pseudopotentials[47], and a planewaves kinetic energy cutoff of 150 Ry. Ground-state and lattice-dynamical calculations are performed using a $12\times12\times12$ uniform and unshifted Brillouin zone mesh for both **k**- and **q**-points. The Kohn–Sham states, phonons, and electron-phonon matrix elements that are needed to solve Eqs. (40) and (41) are generated on Brillouin zone grids with up to $50\times50\times50$ points by Wannier–Fourier interpolation. To this end, we use three Wannier functions to describe the Cl-3*p* states in the valence band, and one Wannier function to describe a single conduction band formed by the Na-3*s* states. The use of additional valence or conduction bands only brings negligible changes to the polaron formation energy. In order to obtain the polaron formation energies in the limit of infinite supercell size, we solve Eqs. (40) and (41) for $N\times N\times N$ **k**- and **q**-point grids with increasing N, and we perform a linear extrapolation of the energy vs. $1/N$ curve. Figure 4a shows an isosurface of the calculated wavefunction of the electron polaron in NaCl. The formation energy of this polaron in the infinite supercell limit is 37 meV; with a radius of 147 Å, this polaron spans several tens of crystalline unit cells. The wavefunction is isotropic and can be described as a combination of Na-3*s* orbitals modulated by an approximately Gaussian envelope.

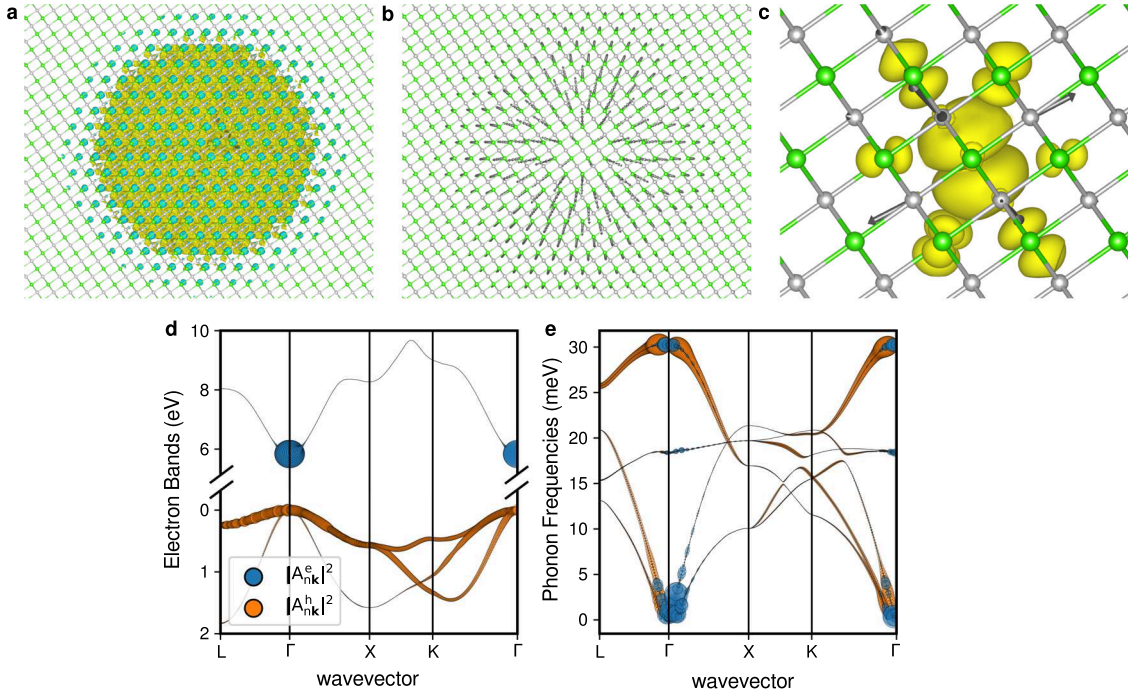

Fig. 4 Calculation of small and large polarons with EPW. a Isosurface plot of the wavefunction of the electron polaron in NaCl. Na atoms are in silver, Cl atoms are in green. **b** Displacements of the Cl atoms associated with the electron polaron shown in (**a**). The displacements are exaggerated for clarity. **c** Isosurface plot of the hole polaron in NaCl, and associated atomic displacements. The displacements are exaggerated for clarity. **d** Spectral weights of the electron (blue) and hole (orange) polaron of NaCl, superimposed to the band structure. The size of the circles is proportional to $|A_{n\mathbf{k}}|^2$. **e** Spectral weights of the electron (blue) and hole (orange) polaron of NaCl, superimposed to the phonon dispersion relations. The size of the circles is proportional to $|B_{\mathbf{q}\nu}|^2$.

In Fig. 4b we show the atomic displacements associated with the electron polaron of Fig. 4a. For clarity we only show the displacement of the Cl ions. The ions tend to move away from the center of the electron wavefunction, consistent with the fact that the electron polaron tends to repel anions. Figure 4c shows an isosurface of the calculated hole polaron of NaCl. The formation energy in the dilute limit is 820 meV, indicating a strongly bound polaron. Consistent with the large formation energy, we find a very small polaron of radius 6 Å, which is comparable to the lattice parameter of 5.69 Å. This polaron consists primarily of a single 3*p* orbital centered on Cl and oriented along the [100] direction. The accompanying atomic displacements are shown in the same panel. As expected from the negative charge of the polaron, the largest displacements are found for the Na cations which tend to move away from the polaron center.

In Fig. 4d we overlay the coefficients $|A_{n\mathbf{k}}|^2$ with the band structure in order to determine which electronic states contribute to the electron (blue) and hole (orange) polaron. In the case of the electron polaron, only electrons near the bottom of the conduction band contribute; the narrow distribution of $A_{n\mathbf{k}}$ near $\mathbf{k}=0$ is consistent with the large spatial extent of the electron polaron in real space. Conversely, the hole polaron draws weight from the entire Brillouin zone, in line with the strong localization in real space. Similarly, in Fig. 4e we show the coefficients $|B_{\mathbf{q}\nu}|^2$ on the phonon dispersion relations. The electron polaron is primarily driven by long-wavelength longitudinal-optical phonons and longitudinal acoustic phonons, while the hole polaron is driven mostly by short-range optical modes. These spectral weight distributions suggest that the electron polaron in NaCl should be classified as a Fröhlich-type[140] polaron, while the hole polaron should be classified as a Holstein-type polaron[141].

Phonon-assisted indirect absorption

Background and formalism. The EPW code has the capability to compute, from first principles, optical absorption spectra in indirect band gap semiconductors by including phonon-assisted optical transitions within second-order time-dependent perturbation theory. Phonon-assisted transitions involve two virtual processes, namely the absorption of a photon, and the absorption or emission of a phonon, in either order. In this section, we outline the general formalism to describe these processes[142,143].

We consider a linearly polarized electromagnetic wave with vector potential of amplitude A_0, frequency ω, and polarization vector $\mathbf{e}$. In the following equations, the long-wavelength limit and the corresponding electric dipole approximation are understood. Second-order time-dependent perturbation theory states that the transition rate of an electron from an initial Kohn–Sham state $n\mathbf{k}$ to a final state $m\mathbf{k}+\mathbf{q}$ involving a photon and a phonon $\mathbf{q}\nu$ is[142]:

$$W_{mn\nu}(\mathbf{k},\mathbf{q};\omega) = \tfrac{2\pi}{\hbar}e^2A_0^2\sum_{\beta=\pm1}\left|\mathbf{e}\cdot[\mathbf{S}_{1,mn\nu}(\mathbf{k},\mathbf{q})+\mathbf{S}_{2,mn\nu\beta}(\mathbf{k},\mathbf{q})]\right|^2 \times\delta(\varepsilon_{m\mathbf{k}+\mathbf{q}}-\varepsilon_{n\mathbf{k}}-\hbar\omega+\beta\hbar\omega_{\mathbf{q}\nu}). \tag{46}$$

In this equation, $\beta=+1$ and -1 represent phonon emission and absorption processes, respectively. $\mathbf{S}_{1,mn\nu}(\mathbf{k},\mathbf{q})$ and $\mathbf{S}_{2,mn\nu\beta}(\mathbf{k},\mathbf{q})$ are the transition amplitudes for the processes illustrated in Fig. 5a: $\mathbf{S}_{1,mn\nu}(\mathbf{k},\mathbf{q})$ refers to a process whereby the electron absorbs a photon, and then absorbs or emits a phonon; $\mathbf{S}_{2,mn\nu\beta}(\mathbf{k},\mathbf{q})$ describes a process whereby phonon absorption/emission takes place, followed by the absorption of a photon. Explicit expressions for these amplitudes are:

$$\mathbf{S}_{1,mn\nu}(\mathbf{k},\mathbf{q}) = \sum_j \frac{g_{mj\nu}(\mathbf{k},\mathbf{q})\mathbf{v}_{jn}(\mathbf{k})}{\varepsilon_{j\mathbf{k}}-\varepsilon_{n\mathbf{k}}-\hbar\omega+i\eta}, \tag{47}$$

$$\mathbf{S}_{2,mn\nu\beta}(\mathbf{k},\mathbf{q}) = \sum_j \frac{\mathbf{v}_{mj}(\mathbf{k}+\mathbf{q})g_{jn\nu}(\mathbf{k},\mathbf{q})}{\varepsilon_{j\mathbf{k}+\mathbf{q}}-\varepsilon_{n\mathbf{k}}+\beta\hbar\omega_{\mathbf{q}\nu}+i\eta}, \tag{48}$$

where $\mathbf{v}_{mn}$ denotes velocity matrix elements between the Kohn–Sham states and the sum extends to all possible occupied

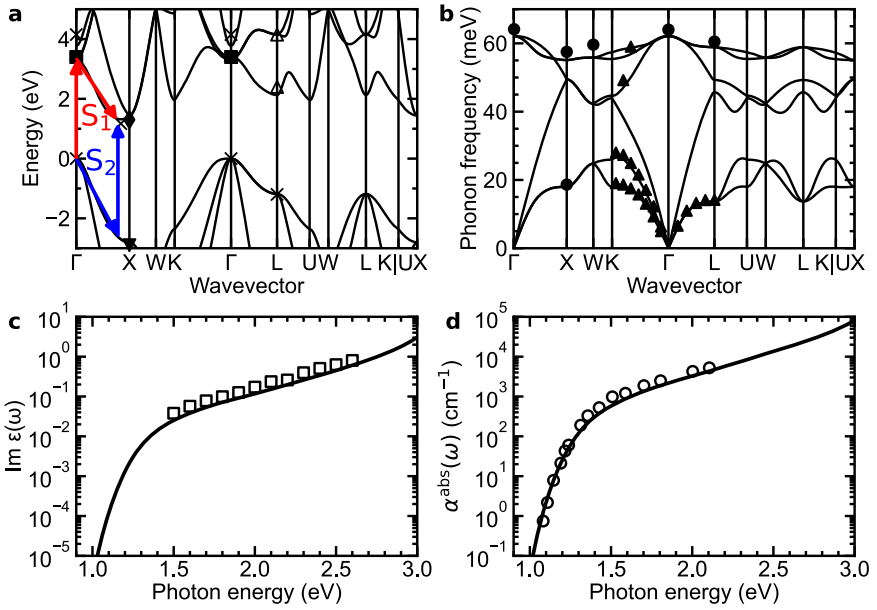

Fig. 5 Indirect phonon-assisted optical absorption spectra with EPW. a Calculated band structure of silicon including GW quasiparticle corrections (lines). The energy is referred to the top of the valence bands. The marks are measured energies of critical points: cross: ref. [158]; diamond: ref. [89]; filled triangle: ref. [221]; filled square: ref. [222]; open triangle: ref. [157]. **b** Calculated phonon dispersion relations of silicon (lines). Neutron scattering data are shown as disks[160] and filled triangles[161]. **c** Calculated imaginary part of the dielectric function of silicon (line) including phonon-assisted processes. The squares represent experimental data from ref. [223]. **d** Calculated absorption coefficient of silicon. The circles are measurements from ref. [224].

and unoccupied states. The energy η is a small parameter to avoid singular denominators. From the transition rates $W_{mn\nu}(\mathbf{k},\mathbf{q};\omega)$, we obtain the imaginary part of the dielectric function by summing over all possible transitions in the Brillouin zone and in the band manifold[142,143]:

$$\mathrm{Im}\left[\epsilon(\omega)\right] = 2\frac{\pi e^2}{\epsilon_0 \Omega}\frac{1}{\omega^2}\sum_{mn\nu,\beta=\pm 1}\int\frac{\mathrm{d}\mathbf{k}}{\Omega_{\mathrm{BZ}}}\int\frac{\mathrm{d}\mathbf{q}}{\Omega_{\mathrm{BZ}}}\left|\mathbf{e}\cdot\left[\mathbf{S}_{1,mn\nu}(\mathbf{k},\mathbf{q})+\mathbf{S}_{2,mn\nu\beta}(\mathbf{k},\mathbf{q})\right]\right|^2 \times P_{mn\nu\beta}(\mathbf{k},\mathbf{q})\delta(\varepsilon_{m\mathbf{k}+\mathbf{q}}-\varepsilon_{n\mathbf{k}}-\hbar\omega+\beta\hbar\omega_{\mathbf{q}\nu}), \tag{49}$$

where the factor of two results from the electron spin in the case of spin-unpolarized systems; this factor is omitted in the case of calculations including spin–orbit coupling. The quantities $P_{mn\nu\beta}(\mathbf{k},\mathbf{q})$ contain the temperature-dependent Fermi–Dirac and Bose–Einstein distributions as follows:

$$P_{mn\nu\beta}(\mathbf{k},\mathbf{q}) = \left(n_{\mathbf{q}\nu}+\frac{1+\beta}{2}\right)f_{n\mathbf{k}}(1-f_{m\mathbf{k}+\mathbf{q}}) - \left(n_{\mathbf{q}\nu}+\frac{1-\beta}{2}\right)(1-f_{n\mathbf{k}})f_{m\mathbf{k}+\mathbf{q}}. \tag{50}$$

In Eq. (49), the scalar quantity $\epsilon(\omega)$ represents the projection of the dielectric tensor along the polarization direction $\mathbf{e}$ of the electric field. For easier comparison with experimental measurements, after we obtain the imaginary part of the dielectric function from EPW, we calculate the absorption coefficient using[142]:

$$\alpha^{\mathrm{abs}}(\omega) = \frac{\omega\,\mathrm{Im}[\epsilon(\omega)]}{c\,n(\omega)}, \tag{51}$$

where $n(\omega)$ is the real part of the refractive index and c is the speed of light. $n(\omega)$ can be calculated from the standard relations between the dielectric function and the refractive index[144]. In this case, the real part of the dielectric function can be obtained by performing a Kramers–Kronig transformation of $\mathrm{Im}[\epsilon(\omega)]$ and applying a rigid shift to match the value of $\epsilon(\omega=0)$ computed from DFPT using Quantum ESPRESSO. Alternatively, $n(\omega)$ can be taken from experiments[143].

Computational considerations. In the optics module of EPW, the imaginary part of the dielectric function is calculated using Eq. (49). The summation over virtual states in Eqs. (47) and (48) are restricted to the manifold of valence and conduction bands included in the Wannierization procedure. The Dirac delta functions appearing in Eq. (49) are replaced by Gaussian functions or Lorentzian functions with a finite broadening. The small parameter η in Eqs. (47) and (48) are used to avoid singular denominators which arise if, for a given photon energy $\hbar\omega$, direct transitions are resonant with indirect transitions. These situations are encountered, for example, at the onset of direct transitions, i.e. when the photon energy matches the direct gap. In these situations, the dielectric function and the absorption spectrum become sensitive to the choice of η: too small an η leads to a divergence of the spectrum, and too large an η leads to an excessive broadening. To probe the sensitivity of the spectra to this parameter, the optics module of EPW calculates $\mathrm{Im}[\epsilon(\omega)]$ for a range of broadening parameters between 1 meV and 0.5 eV. Efforts are currently ongoing to eliminate these spurious singularities: we believe that they arise from an intrinsic limitation of second-order perturbation theory when direct and indirect transitions are in resonance; a more general theory that correctly describes these resonances is under development[145].

In Eqs. (47) and (48), the velocity matrix elements are sensitive to the details of the electronic structure. For accurate calculations, it is preferable to include GW quasiparticle corrections to the Kohn–Sham eigenvalues[89]. These corrections require a renormalization of the velocity matrix elements to preserve the *f*-sum rule of optical transitions[146,147]. In the EPW code, this renormalization is carried out by evaluating the velocity via finite differences. To this end, we consider quasiparticle energies and Kohn–Sham eigenvalues evaluated on **k**-point grids slightly offset along the Cartesian directions. The renormalized velocity matrix elements are given by[147]

$$v^{\mathrm{QP}}_{mn,\alpha}(\mathbf{k}) = \frac{\varepsilon^{\mathrm{QP}}_{m\mathbf{k}+\delta\mathbf{q}_\alpha}-\varepsilon^{\mathrm{QP}}_{n\mathbf{k}-\delta\mathbf{q}_\alpha}}{\varepsilon^{\mathrm{KS}}_{m\mathbf{k}+\delta\mathbf{q}_\alpha}-\varepsilon^{\mathrm{KS}}_{n\mathbf{k}-\delta\mathbf{q}_\alpha}}v^{\mathrm{KS}}_{mn,\alpha}(\mathbf{k}), \tag{52}$$

where $\delta\mathbf{q}_\alpha$ is a small wavevector along the Cartesian direction α, of magnitude $10^{-3}\times 2\pi/a$ (a is the lattice parameter).

To achieve convergence in calculations of indirect phonon-

assisted optical absorption spectra, Brillouin zone grids with $40 \times 40 \times 40$ points have been employed in previous work on silicon[143], as well as in more recent work on silicon and diamond[145] (see Table 1). In addition, convergence of the phonon-assisted optical spectra for cubic boron arsenide[148] and cubic silicon carbide[149] was achieved with Brillouin zone grids of $32 \times 32 \times 32$. When free-carrier absorption is considered, denser grids are required for converged spectra. For example, a grid of $48 \times 48 \times 48$ is needed for converged phonon-assisted free-carrier absorption spectra in silicon[150].

Application example. To demonstrate the implementation of the optics module in `EPW`, we examine phonon-assisted optical absorption in silicon. Gound-state DFT calculations are performed using the PBE exchange and correlation functional[45], ONCV pseudopotentials[46,61], a planewaves kinetic energy cutoff of 60 Ry, and a $6 \times 6 \times 6$ Brillouin zone sampling of both the **k**-point grid and the **q**-point grid. The optimized DFT lattice parameter is $a = 5.478$ Å, similar to previous studies[151–153].

We evaluate quasiparticle corrections on a $6 \times 6 \times 6$ **k**-point grid within the GW method using the `BerkeleyGW` code[154]. The GW-corrected band gap is 1.31 eV, which slightly overestimates the experimental value of 1.12 eV at room temperature[155], but agrees well with prior work using similar settings[89]. The resulting quasiparticle band structure is shown in Fig. 5a. Our calculations agree well with experimental measurements at several critical points[89,156–158], shown by marks in the figure. In Fig. 5a we also show representative optical transitions leading to the amplitudes in Eqs. (47) and (48). The fundamental gap is indirect, between the top of the valence band at the Γ point and the bottom of the conduction band near the *X* point. To correctly describe optical transitions near the fundamental gap, it is essential to include phonon-assisted processes. Figure 5b shows calculated phonon dispersion relations, which agree well with prior calculations by us[18] and other groups[159], as well as experimental data[160,161] (solid symbols). Figure 5c shows the imaginary part of the dielectric function, as obtained from Eq. (49) using $32 \times 32 \times 32$ fine **k**- and **q**-point grids. The temperature is set to 300 K, and the Dirac delta functions in Eq. (49) are approximated via Gaussians of width 50 meV. In this plot, a rigid shift of −0.19 eV is applied in order to match the GW band gap to the measured gap of silicon at room temperature (1.12 eV)[155]. Figure 5d shows the related optical absorption coefficient from Eq. (51). This calculation requires the real part of the frequency-dependent refractive index, for which we used experimentally measured values from ref. [162] for simplicity.

Our calculated spectra agree very well with experiments in the photon energy range between the indirect gap and the direct gap at 3.3 eV. Beyond the lineshape, the magnitude also agrees well with the experiment over several orders of magnitude, with the theory underestimating the experimental data by 50% at most. This residual underestimation might be related to the rigid shift of the band gap, which we did not include in the velocity renormalization expressed by Eq. (52), and to the fact that electron–phonon matrix elements are slightly too weak in DFT as a result of the band gap problem[1,163].

Special displacement method

Background and formalism. The `EPW` package contains a set of standalone modules, the `ZG` toolkit, for calculating finite-temperature properties including quantum zero-point effects via the special displacement method (SDM)[164–166]. The special displacement method is a supercell-based approach that is distinct from and complementary to the Wannier–Fourier interpolation method employed in all previous sections. The founding principle of this method is that the effects of electron–phonon couplings on the electronic and optical properties of extended solids can be captured by performing calculations for a large supercell where the atoms have been displaced away from their equilibrium crystallographic sites. The displacements are chosen in such a way that the corresponding atomic configuration represents the best single-point approximant to the quantum thermal distribution of the atomic coordinates[165].

Under the approximations of adiabatic Born–Oppenheimer decoupling and harmonic lattice, the quantum thermal average of an electronic or optical property described by the observable *O* can be written as[165]

$$O(T) = \prod_{\mathbf{q}\nu} \int \frac{\mathrm{d}x_{\mathbf{q}\nu}\mathrm{d}y_{\mathbf{q}\nu}}{\pi u^2_{\mathbf{q}\nu}} e^{-|z_{\mathbf{q}\nu}|^2/u^2_{\mathbf{q}\nu}} O(\{\tau_{\kappa a p}\}), \tag{53}$$

where **q** runs over the set of wavevectors in a uniform Brillouin zone grid which excludes time-reversal invariant points and time-reversal partners. In Eq. (53), $z_{\mathbf{q}\nu}$ denote normal mode coordinates with real part $x_{\mathbf{q}\nu}$ and imaginary part $y_{\mathbf{q}\nu}$. $u^2_{\mathbf{q}\nu} = (\hbar/2M_0\omega_{\mathbf{q}\nu})(2n_{\mathbf{q}\nu}+1)$ is the mean-square displacement for the oscillator $\mathbf{q}\nu$ with Bose–Einstein occupation $n_{\mathbf{q}\nu}$; $O(\{\tau\})$ is the property of interest, such as for example the Kohn–Sham eigenvalues, density of states, or optical absorption spectrum, calculated for the set of atomic coordinates $\{\tau_{\kappa a p}\}$. The relation between these coordinates and the normal coordinates $z_{\mathbf{q}\nu}$, which is required to carry out the integral in Eq. (53), is provided in ref. [165].

In the special displacement method, the configurational average expressed by Eq. (53) is approximated by a single calculation for an optimum configuration:

$$O(T) \simeq O(\{\tau^0_{\kappa a p} + \Delta\tau^{\mathrm{ZG}}_{\kappa a p}\}), \tag{54}$$

where $\tau^0_{\kappa a p}$ represent atomic coordinates in the DFT ground-state at zero temperature, and the optimum 'ZG' displacement is given by[165]

$$\Delta\tau^{\mathrm{ZG}}_{\kappa a p} = \sqrt{\frac{M_0}{N_p M_\kappa}} 2 \sum_{\mathbf{q}\nu} S_{\mathbf{q}\nu} u_{\mathbf{q}\nu} \,\mathrm{Re}\left[e^{i\mathbf{q}\cdot\mathbf{R}_p} e_{\kappa a,\nu}(\mathbf{q})\right]. \tag{55}$$

As in Eq. (53), the summation is restricted to **q**-points from a uniform Brillouin zone grid which are not time-reversal invariant and are not time-reversal partners. This partitioning is described in Appendix B of ref. [1]. The quantities $S_{\mathbf{q}\nu}$ appearing in Eq. (55) are signs (±1) determined by the `ZG` module so as to guarantee that the resulting displacements make $O(\{\tau^0_{\kappa a p} + \Delta\tau^{\mathrm{ZG}}_{\kappa a p}\})$ the best possible approximant to Eq. (53). In the thermodynamic limit of a large supercell, the ZG displacements reproduce the exact mean-square anisotropic displacement tensors, which are given by[167]

$$U_{\kappa,aa'}(T) = \frac{\hbar}{2M_\kappa \omega_{\mathbf{q}\nu}} \sum_\nu \int \frac{\mathrm{d}\mathbf{q}}{\Omega_{\mathrm{BZ}}} e_{\kappa a,\nu}(\mathbf{q}) e^*_{\kappa a',\nu}(\mathbf{q})\,(2n_{\mathbf{q}\nu}+1). \tag{56}$$

In the same limit, a single evaluation of the property $O(\{\tau^0_{\kappa a p} + \Delta\tau^{\mathrm{ZG}}_{\kappa a p}\})$ tends to the exact thermal average in Eq. (53). In the case of non-periodic systems, such as for example nanocrystals and quantum dots, these equations are replaced by their Γ-point only versions[164,168,169].

The special displacement method is similar in spirit to computing thermodynamic averages using path-integral molecular dynamics[170], but it differs insofar as a single calculation is required to evaluate the average instead of many molecular dynamics snapshots. In this method, electron–phonon couplings are included non-perturbatively through the changes of the Kohn–Sham energies and wavefunctions caused by the ZG displacements. The method can be applied to compute any property that can be expressed by means of a Fermi Golden Rule, such as for example temperature-dependent band structures, density of states, and optical spectra. The main limitation of this approach as compared to the Wannier–Fourier interpolation method employed in the preceding sections is that, being an adiabatic theory, fine spectral features on the scale of the phonon energy are averaged out. Conversely, its main advantage is that it is easy to use as it requires a single DFT calculation.

The Kohn–Sham energy eigenvalues generated by the special displacement method capture temperature renormalization and quantum zero-point effects at the same level as the adiabatic Allen–Heine theory of temperature-dependent band structures[171]. The calculation of the imaginary part of the dielectric function including temperature effects, zero-point corrections, and phonon-assisted indirect processes, is performed by evaluating:

$$\mathrm{Im}[\epsilon(\omega)] = \frac{2\pi e^2}{\epsilon_0 m_e^2 \Omega_{sc}} \frac{1}{\omega^2} \sum_{cv} \int \frac{d\mathbf{K}}{\Omega_{BZ,sc}} |\langle \psi_{c\mathbf{K}} | \mathbf{e} \cdot \hat{\mathbf{p}} | \psi_{v\mathbf{K}} \rangle|^2 \delta(\varepsilon_{c\mathbf{K}} - \varepsilon_{v\mathbf{K}} - \hbar\omega), \tag{57}$$

where Ω_{sc} and $\Omega_{BZ,sc}$ are the volumes of the supercell and the supercell Brillouin zone, respectively, the summations over v, c refer to valence and conduction states, and $\mathbf{K}$ is a wavevector of the supercell Brillouin zone. This expression contains both direct and indirect optical transitions: the indirect transitions are hidden in the dependence of the wavefunctions $\psi_{v\mathbf{K}}$ and $\psi_{c\mathbf{K}}$ on the ZG displacement. Similarly, this expression includes temperature renormalization via the dependence of the energies $\varepsilon_{v\mathbf{K}}$ and $\varepsilon_{c\mathbf{K}}$ on the ZG displacements[164].

Computational considerations. Calculations using the special displacement method proceed as follows. First, one performs phonon calculations for the crystalline unit cell (not the supercell) using standard DFPT on a coarse uniform Brillouin zone grid. From this calculation, the phonon frequencies and eigenmodes are computed on a finer Brillouin-zone grid with $N_1 \times N_2 \times N_3 \mathbf{q}$-points using standard interpolation of the force constant matrix[8]. Then the ZG displacement within a supercell consisting of $N_1 \times N_2 \times N_3$ unit cells is evaluated at the temperature T via Eq. (55). The desired property is finally computed with this supercell, with the atoms displaced according to $\Delta\tau^{ZG}_{\kappa a p}$. The special displacements are generated by the `ZG` module; the procedure is computationally inexpensive and is performed serially.

The `ZG` toolkit provides several codes to analyze the results of supercell calculations performed with special displacements, for example, Brillouin-zone unfolding of temperature-dependent band structures from the supercell to the unit cell, the density of states, electronic spectra, and vibrational spectra.

Band unfolding is performed using the procedure outlined in ref. [172], which consists of determining the spectral function in the primitive Brillouin zone by projecting the wavefunctions of the supercell into the wavefunctions of the unit cell. In practice, the spectral density for the wavevector $\mathbf{k}$ at the energy ε is obtained as

$$A_{\mathbf{k}}(\varepsilon) = \sum_{m\mathbf{K}} P_{m\mathbf{K},\mathbf{k}}\, \delta(\varepsilon - \varepsilon_{m\mathbf{K}}), \tag{58}$$

where the spectral weights $P_{m\mathbf{K},\mathbf{k}}$ are given by

$$P_{m\mathbf{K},\mathbf{k}} = \sum_{\mathbf{G}_{sc}} |c_{m\mathbf{K}}(\mathbf{G}_{sc} + \mathbf{k} - \mathbf{K})|^2. \tag{59}$$

Here, $c_{m\mathbf{K}}$ denotes planewaves coefficients of supercell wavefunctions, and $\mathbf{G}_{sc}$ denotes reciprocal lattice vectors of the supercell. Equation (59) refers to norm-conserving pseudopotential implementations. The `ZG` module includes additional terms that are required in this expression when using ultrasoft and PAW (projector augmented-wave) pseudopotentials[165,169]. A similar unfolding strategy is employed to analyze lattice dynamics, as well as X-ray and neutron diffuse scattering intensities accounting for multiphonon interactions[173].

The supercell size required to achieve converged results in calculations employing the special displacement method depends on the target property. For example, to compute optical spectra over a broad photon energy range, supercells of size $4 \times 4 \times 4$ unit cells are typically sufficient. However, to obtain accurate phonon-induced band gap renormalization, say within a 10 meV accuracy, supercells as large as $10 \times 10 \times 10$ may be required[165] (see Table 1).

Application example. To demonstrate the implementation of the `ZG` module, we investigate the temperature-dependent band structure renormalization and phonon-assisted optical absorption spectra of silicon and $BaSnO_3$. We perform calculations using the local density approximation (LDA) for the exchange and correlation[174,175] and ONCV pseudopotentials[46,61]. We use a planewaves kinetic energy cutoff of 40 Ry for silicon and 120 Ry for $BaSnO_3$, and $6 \times 6 \times 6$ uniform $\mathbf{k}$-point grids for ground-state calculations. With these settings, we obtain indirect and direct band gaps of 0.49 and 2.56 eV for silicon, respectively, and indirect and direct band gaps of 1.06 and 1.55 eV for $BaSnO_3$, respectively. We perform calculations of dynamical matrices on $4 \times 4 \times 4$ $\mathbf{q}$-point grids in both cases and use the `ZG` module to generate displacements in $3 \times 3 \times 3$ supercells. The signs $S_{\mathbf{q}\nu}$ appearing in Eq. (55) are determined by minimizing the error descriptor in Eq. (54) of ref. [165] with a dimensionless threshold $\eta = 0.1$, after enforcing a smooth Berry connection between vibrational eigenmodes across the Brillouin zone[165]. Kohn–Sham energies in the structures with ZG displacements are calculated using a $12 \times 12 \times 12$ uniform $\mathbf{K}$-point grid in the supercell Brillouin zone, and 108 unoccupied states. For calculations of dielectric functions, we use up to 200 randomly generated $\mathbf{K}$-points in the supercell Brillouin zone, as well as 27 and 135 conduction bands for silicon and $BaSnO_3$, respectively. All Dirac deltas are replaced by Gaussians of width 30 meV.

Figure 6a shows the joint density of states (JDOS) of silicon at 0 K (red) and 300 K (blue), as calculated using the special displacement method. For comparison, the JDOS computed for the DFT ground state structure is also shown in green. We see that, upon including electron-phonon interactions via the special displacements, even at 0 K the JDOS is red-shifted with respect to the DFT ground state. This is a manifestation of the zero-point band gap renormalization[176]. Upon increasing temperature, the onset of the JDOS further red-shifts. This temperature-induced band gap narrowing is referred to as the Varshni effect[177]. The horizontal offset between the JDOS for the DFT ground state structure and the JDOS computed with special displacements at 0 K yields a zero-point renormalization of 50 meV, in good agreement with prior work[178,179]. More accurate values can be calculated by increasing the supercell size[164,165]. The band gap renormalization can also be computed without using the JDOS; to this end, one needs to evaluate the Kohn–Sham eigenvalues at the supercell $\mathbf{K}$-points that unfold onto the wavevectors of the band extrema in the Brillouin zone of the unit cell.

Figure 6b and c report color maps of the electronic spectral functions of silicon and $BaSnO_3$ at 0 K calculated using the special displacement method. For comparison, the band structures in the DFT ground state are overlaid with these color maps. These maps can directly be compared to angle-resolved photoelectron spectroscopy data. From these images, we identify numerically the quasiparticle band structures by extracting the spectral peaks. In the case of silicon, the valence band maximum blue-shifts by 32 meV with respect to ground-state DFT, and the conduction band red-shifts by 18 meV. The resulting gap renormalization of 50 meV is in agreement with the value determined in Fig. 6a via the JDOS. In the case of $BaSnO_3$, Fig. 6 shows a zero-point renormalization of the band gap of 10 meV; however, we emphasize that this value is not fully converged, and larger supercells, as well as corrections for Fröhlich couplings[180], are necessary to obtain accurate data.

Figure 6d reports a convergence test of the imaginary part of the dielectric function of silicon, as computed with the special displacement method. In this calculation, we keep the supercell fixed, and we increase the number of random $\mathbf{K}$-points in the supercell Brillouin zone. It is seen that the dielectric function converges relatively rapidly with the number of points and full convergence is achieved with 200 points when the Dirac delta functions in Eq. (57) are replaced by Gaussians with 30 meV width.

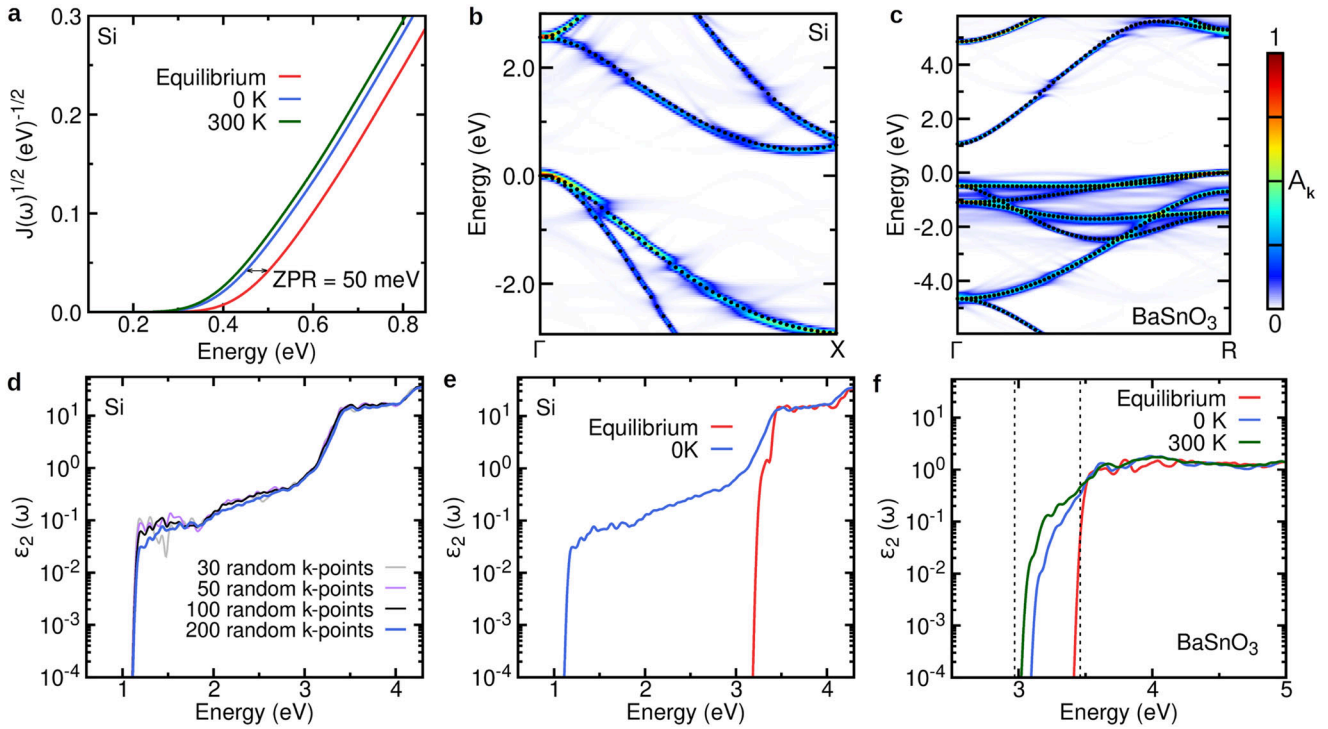

Fig. 6 Applications of the special displacement method implemented in the `ZG` module. a Square root of the JDOS of silicon evaluated with atoms in the DFT ground-state structure (red); using special displacements at 0 K (blue); and using special displacements at 300 K (green). The red-shift of the JDOS with respect to the ground-state structure signals the quantum zero-point correction to the band gap (0 K curve) and its temperature renormalization (300 K curve), respectively. **b** Electron spectral function of silicon calculated with the special displacement method at 0 K, along the ΓX path in the Brillouin zone. The black disks indicate the band structures calculated in the crystalline unit cell with the DFT ground-state geometry. **c** Electron spectral function of $BaSnO_3$ was calculated with the special displacement method at 0 K, along the ΓR path in the Brillouin zone. The black disks indicate the band structures calculated in the crystalline unit cell with the DFT ground-state geometry. **d** Convergence of the imaginary part of the dielectric function of silicon at 0 K with respect to the number of random **K**-points used to sample the Brillouin zone of the supercell. **e** Imaginary part of the dielectric function of silicon calculated with the atoms in the DFT ground-state structure (red), and by using special displacements at 0 K (blue). A scissor shift of 0.64 eV is employed to match the experimental gap[225]. **f** Imaginary part of the dielectric function of $BaSnO_3$ calculated with the atoms in the DFT ground-state structure (red), and by using special displacements at 0 K (blue) and 300 K (green). The vertical dashed lines indicate the direct and indirect band gap energies. A scissor shift of 1.91 eV is employed to match the experimental gap[226].

Clearly, a larger smearing would require fewer **K**-points.

In Fig. 6e we compare the imaginary part of the dielectric function of silicon computed in the DFT ground state structure (red) with the special displacement method (blue). We see that the special displacements correctly capture phonon-assisted indirect optical transitions in the energy range between the indirect gap and direct gap of silicon. This approach provides an alternative strategy for computing optical spectra to Eq. (49) implemented in the indirect optics module of `EPW`. In the special displacement method, both temperature-dependent band structure renormalization and phonon-assisted processes are included on the same footing. Figure 6f shows the imaginary part of the dielectric function of $BaSnO_3$ at 0 K (blue) and 300 K (green), including phonon-assisted processes. For comparison, we also show the calculation using the DFT ground-state structure (red), which misses phonon-assisted transitions. The spectra compare well with prior work using the special displacement method[181].

Implementation and HPC benchmarks

In this section, we describe some of the implementation and parallel programming models of `EPW`. We first outline the computational workflow and basic capabilities of `EPW`. Then we describe our recent efforts to prepare `EPW` for exascale HPC systems with a focus on an explorative implementation of a highly scalable parallelization scheme, hybrid two-level MPI, and OpenMP parallelization. We present benchmarking results which demonstrate that, with this improved parallelization strategy, `EPW` can run at nearly full scale on the pre-exascale HPC system Frontera at the Texas Advanced Computing Center (TACC). Finally, we describe the I/O strategy employed in this explorative implementation, which is based on XML (eXtensible Markup Language) and HDF5 file formats. We emphasize that `EPW` continues to evolve with changes and improvements, and the reader is referred to the `EPW` website for its most up-to-date features and functionalities[182].

Computational workflow and basic capabilities. The `EPW` code is logically partitioned into two components. The first component serves as an interface to the `PWscf` and `PHonon` codes of `Quantum ESPRESSO`. The code imports ground-state charge density and wavefunctions with computational parameters from prior runs with `pw.x`, and dynamical matrices, variations of the Kohn–Sham potentials, and vibrational mode pattern files from a prior phonon calculation with `ph.x`. All these quantities are evaluated on coarse **k**- and **q**-points grids, and the phonon wavevectors are restricted to the irreducible wedge of the Brillouin zone. Then the code calls `Wannier90` in library mode to calculate the unitary rotation matrices $U_{mn\mathbf{k}}$ [Eq. (64) of the "Methods" section] needed to transform Bloch states into maximally localized Wannier functions. Subsequently, the code unfolds the irreducible wedge of the phonon Brillouin zone into

the full zone, evaluates electron–phonon matrix elements on coarse **k**- and **q**-point grids, and it transforms the Hamiltonian, dynamical matrices, and electron–phonon matrix elements into the Wannier representation. At this point, the code can print relevant quantities for subsequent runs with `epw.x`, and it can continue to the interpolation part. The interpolation engine of `EPW` carries out Wannier interpolation of the Hamiltonian, dynamical matrices, and electron–phonon matrix elements on fine **k**- and **q**-point grids after reading these quantities in the Wannier representation, and then it calculates electron–phonon-related properties on these grids.

The current snapshot of the code supports spin-unpolarized and non-magnetic spin–orbit calculations; work is currently in progress to extend the code to the cases of spin-polarized and non-collinear magnetic cases. Regarding pseudopotentials (PPs), norm-conserving (NC) PPs are supported. Since both ultrasoft PPs and the PAW method violate fundamental symmetry relations that the electron–phonon matrix elements must fulfill[183], extending the current NCPP implementation to these cases poses some challenges. We will proceed to these extensions in future releases.

Parallelization. At the time of writing of this manuscript, we witness the launch of the first supercomputer entering the exascale computing era, namely Frontier at Oak Ridge National Laboratory; additional exascale systems are currently being deployed worldwide. Exascale HPC systems are characterized by many-core and heterogeneous architectures. Since 2020, we have made efforts to ready `EPW` for the exascale transition. In the following, we describe one strategy that we are currently exploring, a hybrid two-level MPI/OpenMP parallelization scheme, which makes `EPW` work seamlessly and efficiently on many-core architectures.

Historically `EPW` has adopted a one-level parallelization over the electron wavevectors (**k**-point parallelization) via MPI. This strategy has three shortcomings, which lead to an early saturation of the speedup with the number of cores in large-scale calculations: (i) There is an upper bound for the number of total MPI tasks which guarantees good scalability; this bound is determined by the number of **k**-points. (ii) As the number of MPI tasks increases, the overhead cost associated with MPI communications, in particular collective communications, increases. (iii) Due to the internal buffers of MPI, the memory overhead also increases with the number of MPI tasks. To address these issues, we are exploring the extension of the one-level parallelization over **k**-points to the hybrid two-level MPI and OpenMP parallelization over both **k**- and **q**-points.

In the hierarchical two-level MPI and OpenMP parallelization strategy, the total MPI tasks are partitioned into two levels of MPI groups, each consisting of a set of MPI tasks; in the lowest level, OpenMP parallelization is employed within each MPI task. For example, the total MPI tasks are first divided into **k**- or **q**-point pools, and each pool is further divided into **q**- and **k**-point pools, respectively. Each **q**- or **k**-point pool contains a set of MPI tasks, and OpenMP parallelization is employed within each MPI task. The parallelization option employed in the upper-level pool (over **k**-points or **q**-points) is chosen so as to maximize the calculation efficiency.

As compared to the one-level parallelization, the hierarchical two-level MPI and OpenMP parallelization can reduce the number of MPI tasks in each MPI communicator group, thereby reducing the overhead cost for collective MPI communication as well as the memory footprint due to the internal buffers of the MPI library. OpenMP parallelization can further reduce the memory footprint associated with replicated memory allocations in the code. This strategy enables increased flexibility in distributing the computational workload over a large number of cores.

To demonstrate the scaling behavior of `EPW`, we perform strong-scaling tests for the two-gap superconductor MgB_2. We evaluate electron-phonon matrix elements on $150 \times 150 \times 150$ **k**- and **q**-point grids (Fig. 7). These calculations are performed on the Frontera supercomputer at TACC. Frontera consists of 8368 nodes, each equipped with two Intel Xeon Platinum 8280 ("Cascade Lake") processors with 28 cores per processor. To minimize statistical fluctuations in these benchmarks, we execute each calculation four times, and we average the resulting Wall times. Figure 7 shows that using this parallelization scheme, we achieve ~92% of the ideal speedup up to 448,000 cores in the evaluation of the electron–phonon matrix elements.

The hybrid two-level MPI and OpenMP parallelization scheme is controlled by three parameters that can be optimized to achieve maximum parallel scaling efficiency: the number of upper-level MPI tasks, the number of lower-level MPI tasks, and the number of OpenMP threads per MPI task. In the benchmarks shown in Fig. 7, we fix the number of **k**-point pools to 250, by varying the number of **q**-point pools depending on the number of cores. With this choice, we are able to achieve near-ideal speedup on extreme scaling tests; we expect that a speedup even superior to 92% could be achieved upon further optimization of the parallelization parameters We also point out that, in large-scale runs, I/O time dominates the total Wall time. To overcome this I/O bottleneck, we employ a low-I/O mode whereby memory is exploited as much as possible and minimum usage of storage media is made.

Our strategy for OpenMP parallelization is to use threaded versions of scientific libraries such as Intel MKL and Cray LibSci, rather than explicitly using OpenMP directives throughout the code. For the benchmarks reported in Fig. 7, we find that exceeding 7 OpenMP threads per MPI task worsens the performance. Therefore we use 4 OpenMP threads per MPI task, and we make all threads belonging to the same MPI task reside in the same non-uniform memory access (NUMA) domain.

Structured I/O. Recently we have been investigating the use of XML and the HDF5 data formats. Both formats can describe hierarchical data and are self-describing, flexible, and portable. For small to medium-sized data, the XML format is used; for large binary data, the HDF5 file format is used in parallel. In previous versions of `EPW`, several large arrays were read or written via per-process I/O or parallel MPI I/O. The former quickly overwhelms the file systems in large-scale calculations by generating a large number of files; the latter is not easily portable. Both issues can be addressed for large-size binary data by employing parallel I/O via parallel HDF5.

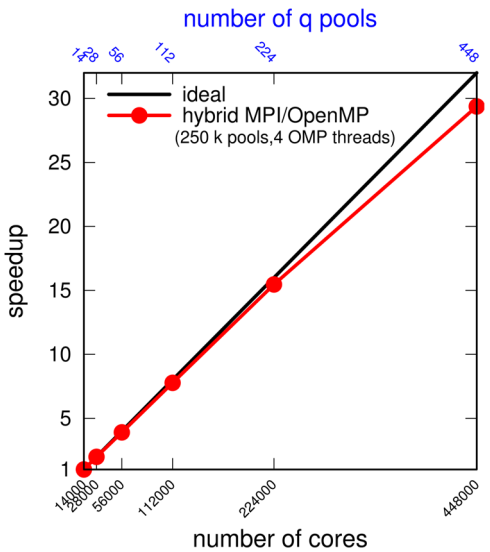

Fig. 7 Extreme scaling benchmarks of `EPW`. Strong-scaling test for the evaluation of the electron–phonon matrix elements of MgB_2. These tests were performed using `EPW v6-alpha` available at ref. [227].

We note that parallel I/O performance heavily depends on the details of file systems and the parallelization settings. For instance, in the case of the LUSTRE file system, the number and size of stripes play an essential role in achieving high performance. In the case of the ROMIO implementation of MPI-IO, identifying optimum parameters in large-scale runs is nontrivial and requires systematic experimentation on HPC systems.

DISCUSSION

Since the last technical paper in 2016[18], the EPW code has considerably expanded in scope, to the point of becoming a robust and efficient software platform for developing, testing, and deploying methods addressing electron–phonon physics and related materials properties. It is then natural to ask which components and functionalities are still in need of improvement, what are emerging directions in this area, and more generally what comes next for the EPW project.

On the functionality front, we expect to continue expanding the capabilities for transport, superconductivity, polarons, and optics described in the 'Results' section. For example, the transport module is based on the steady-state linearized Boltzmann transport equation; here, it would be desirable to generalize the methodology to high-field transport and to time-dependent driving fields. Similarly, in the superconductivity module, the treatment of electron–electron repulsion could be improved by seamlessly integrating EPW with standard GW codes. The study of polaron physics from an ab initio many-body perspective is just beginning, therefore we can expect a number of developments in this area, from the study of polaron transport to their optical properties and their relation with the theory of band structure renormalization. Calculations of phonon-assisted optical transitions are currently restricted to optical absorption spectra, but the formalism can just as well be employed for investigating related phenomena such as for example Auger–Meitner recombination[184]. Beyond these core modules of the EPW project, we anticipate growth in the calculation of electronic and optical properties using the special displacement method within the EPW/ZG code; for example, the current implementation focuses on harmonic systems, but generalizations to strongly anharmonic systems have recently been proposed[185].

On the accuracy front, we anticipate that future work will focus on improving the precision of the Wannier interpolation method that is the basis for EPW. For example, methods that take dimensionality into account in the interpolation of the electron–phonon matrix elements will be essential to perform predictive and reliable calculations of transport, optics, superconductivity, and polarons in 2D materials[186,187]. Furthermore, it would be highly desirable to improve the predictive power of the DFPT electron–phonon matrix elements before even proceeding to Wannier–Fourier interpolation; a recent proposal to employ the GW method to calculate many-body corrections to the DFPT matrix elements offers a promising path toward this goal[188].

On the front of HPC, we envision continuing the current effort to enable EPW for exascale computing architectures. In addition to the hybrid MPI/OpenMP parallelization, it will be necessary to leverage diverse GPU architectures and to enable the code for large-scale runs that will become possible with the exascale supercomputers that are being deployed worldwide.

Another important direction will be to enhance the interoperability of EPW with other major electronic structure software packages such as Abinit[189], VASP[190], Siesta[191], and GPAW[192]. As several codes currently use the information generated by EPW for a variety of post-processing tasks, it will make sense to develop standardized data structures that adhere to the FAIR data principles.

Regarding the programming model, transitioning toward object-oriented programming (OOP) will be useful due to the potential for high modularity, extensibility, and reusability, which will facilitate flexible and sustained software development[193]. We expect that targeted use of OOP will greatly increase the efficiency and productivity in the development and maintenance of code[194].

Alongside these developments, we anticipate increased attention to the issue of automation and the capability of performing electron–phonon calculations at scale for high-throughput approaches, data science, and artificial intelligence/machine learning (AI/ML) applications.

In summary, in this manuscript, we provided a comprehensive update on the current status, functionalities, and performance of the EPW code, we offered an overview of possible future directions for this and similar codes. We hope that, beyond capturing a snapshot of the current status of the EPW project, the present manuscript will contribute to making advanced electron–phonon calculations more accessible and more widely used in computational materials discovery and design.

METHODS

The key element of any calculation of electron-phonon interactions and related materials properties is the electron–phonon matrix element. The EPW code employs physics-based Wannier–Fourier interpolation to compute electron-phonon matrix elements accurately and efficiently and uses these matrix elements as the starting point for advanced many-body calculations of electronic, optical, and transport properties of solids including temperature dependence and phonon-assisted quantum processes.

In the following we introduce the notations and conventions employed throughout the manuscript, we provide a brief overview of the general principles of Wannier–Fourier interpolation of electron–phonon matrix elements, and we outline the extension of this interpolation scheme to the case of polar materials. A more comprehensive description of this methodology including detailed derivations can be found in ref. [10].

Notation and definitions

In the study of the electronic structure and lattice dynamics of crystalline solids, infinitely extended crystals are described by considering a periodic BvK supercell. This supercell consists of $N_{\mathrm{uc}} = N_1 \times N_2 \times N_3$ primitive unit cells of the crystal and is subject to periodic boundary conditions. The pth unit cell within the BvK supercell is identified by the direct lattice vector $\mathbf{R}_p$, with $p = 1, \ldots, N_{\mathrm{uc}}$. The dual lattice of the direct lattice vectors $\mathbf{R}_p$ in reciprocal space defines a uniform grid of $N_{\mathrm{uc}} = N_1 \times N_2 \times N_3$ wavevectors in the Brillouin zone. Throughout the manuscript, we use $\mathbf{k}$ and $\mathbf{q}$ to denote electron and phonon wavevectors in the crystal Brillouin zone, respectively, and $\mathbf{K}$ to denote electron wavevectors in the Brillouin zone of the BvK supercell.

We use $\psi_{n\mathbf{k}}$ to indicate the eigenfunctions of the DFT Kohn-Sham Hamiltonian $\hat{H}_{\mathrm{KS}}$, and $\varepsilon_{n\mathbf{k}}$ the corresponding eigenvalues. This wavefunction is periodic and normalized over the BvK supercell, and can be expressed in the Bloch form:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N_{\mathrm{uc}}}} u_{n\mathbf{k}}(\mathbf{r}) \mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}, \tag{60}$$

with $u_{n\mathbf{k}}$ being a lattice-periodic function normalized in the primitive unit cell of volume Ω. The position vector of the atomic nucleus κ belonging to the unit cell identified by $\mathbf{R}_p$ is $\boldsymbol{\tau}_{\kappa p} = \mathbf{R}_p + \boldsymbol{\tau}_\kappa$, with $\boldsymbol{\tau}_\kappa$ being the position vector within the primitive unit cell. The components of this vector along the Cartesian directions $\alpha = 1, 2, 3$ are denoted by $\tau_{\kappa\alpha p}$.

Vibrational eigenmodes and eigenfrequencies are obtained by diagonalizing the dynamical matrix, which is computed via DFPT using Quantum ESPRESSO[12]. In turn, the dynamical matrix is the Fourier transform of the interatomic force constants $C_{\kappa\alpha p,\kappa'\alpha' p'}$, that

is the Hessian of the DFT total energy in the atomic coordinates. The relation between interatomic force constants and dynamical matrix is[135]

$$D_{\kappa\alpha,\kappa'\alpha'}(\mathbf{q}) = \frac{1}{\sqrt{M_\kappa M_{\kappa'}}} \sum_p e^{i\mathbf{q}\cdot\mathbf{R}_p} C_{\kappa\alpha 0,\kappa'\alpha' p}, \tag{61}$$

where M_κ is the mass of the κth nucleus. The eigenvalues and eigenvectors of the dynamical matrix are denoted by $\omega^2_{\mathbf{q}\nu}$ and $e_{\kappa\alpha,\nu}(\mathbf{q})$, respectively, and the index ν runs from 1 to $3M$ where M is the number of atoms in the primitive cell. $\omega_{\mathbf{q}\nu}$ corresponds to the vibrational frequency, and $e_{\kappa\alpha,\nu}(\mathbf{q})$ is the normal mode of vibration or polarization vector.

The electron–phonon matrix element is defined by[1,10]

$$g_{mn\nu}(\mathbf{k},\mathbf{q}) = \langle u_{m\mathbf{k}+\mathbf{q}} | \Delta_{\mathbf{q}\nu} v^{\mathrm{KS}} | u_{n\mathbf{k}} \rangle, \tag{62}$$

where the integral is evaluated over the unit cell, and the lattice-periodic component of the variation of the Kohn–Sham potential is given by

$$\Delta_{\mathbf{q}\nu} v^{\mathrm{KS}} = \sqrt{\frac{\hbar}{2M_0\omega_{\mathbf{q}\nu}}} \sum_{\kappa\alpha p} e^{-i\mathbf{q}\cdot(\mathbf{r}-\mathbf{R}_p)} \sqrt{\frac{M_0}{M_\kappa}}\, e_{\kappa\alpha,\nu}(\mathbf{q})\, \frac{\partial V^{\mathrm{KS}}(\mathbf{r})}{\partial \tau_{\kappa\alpha p}}. \tag{63}$$

In this expression, M_0 is an arbitrary reference mass introduced for convenience. MLWFs are defined in terms of Bloch states as[9]

$$\mathrm{w}_{mp}(\mathbf{r}) = \mathrm{w}_{m0}(\mathbf{r}-\mathbf{R}_p) = \frac{1}{N_{\mathrm{uc}}} \sum_{n\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{R}_p)}\, U_{nm\mathbf{k}}\, u_{n\mathbf{k}}(\mathbf{r}), \tag{64}$$

where $U_{nm\mathbf{k}}$ is a unitary matrix in the indices m and n.

Throughout the manuscript, we used interchangeably the notation

$$\frac{1}{N_{\mathrm{uc}}} \sum_{\mathbf{k}} \qquad \text{or} \qquad \int \frac{d\mathbf{k}}{\Omega_{\mathrm{BZ}}} \tag{65}$$

to indicate summations or integrals over the Brillouin zone. The quantity Ω_{BZ} on the right indicates the Brillouin zone volume.

The unitary matrix $U_{nm\mathbf{k}}$ appearing in Eq. (64) is determined by requiring that the functions $\mathrm{w}_{mp}(\mathbf{r})$ be as localized as possible[195], i.e. MLWFs[9]. The computation of these unitary matrices is performed by minimizing the spatial spread of the resulting Wannier functions[9]. This procedure is notoriously challenging since the spread functional exhibits multiple local minima, especially in the case of entangled electron bands. In practice, the construction of optimal MLWFs often requires fine-tuning of several computational parameters, such as the inner and outer energy windows, and the projections used to initialize the minimization algorithm. Historically, this optimization has often been performed by using chemical intuition, and sometimes a trial-and-error approach.

In an effort to enable more systematic construction of MLWFs for high-throughput calculations and data-driven materials research, several strategies have been developed, such as for example those of refs. [196,197]. These strategies fall into two categories, one relying on the projectability concept[198], and the other on the selected-columns-of-the-density-matrix (SCDM) method[199,200]. The former approach relies on the projectability of Bloch states on pseudo-atomic orbitals associated with the pseudopotentials; the projectability metric of each Bloch state is used to downselect the band manifold for subsequent minimization of the spread functional. The latter approach does not rely on the pseudo-atomic orbitals or other choice of initial projections but generates an optimal set of states from the columns of the density matrix via QR decomposition with column pivoting.

Both approaches are very promising and bring Wannier functions one step closer to complete automation. Nevertheless, some challenges remain. For example, to automatically construct Wannier functions and achieve high-quality Wannier interpolation, both strategies require the use of all the occupied valence bands (except for low-lying semi-core states) as well as a few empty conduction bands for metallic or semi-metallic systems. In systems with medium-sized to large unit cells, the number of bands thus identified may be relatively large, and this will make EPW calculations more demanding because the number of operations required to obtain electron-phonon matrix elements scales with the square of the number of Wannier functions. Therefore, more work will be needed in the future to achieve automation whilst preserving calculation efficiency.

General principles of Wannier interpolation of electron–phonon matrix elements

The localized nature of MLWFs provides the basis for accurate and efficient interpolation of Kohn–Sham wavefunctions and energies[201]. This is achieved by expressing the Kohn–Sham Hamiltonian in the Wannier representation, and noting that the Hamiltonian matrix elements decrease rapidly with the distance between the Wannier function centers. The Hamiltonian matrix elements in the Wannier representation are[195]

$$H_{mn}(\mathbf{R}_p) = \langle \mathrm{w}_{m0} | \hat{H} | \mathrm{w}_{np} \rangle = \int_{\mathrm{sc}} d\mathbf{r}\, \mathrm{w}^*_{m0}(\mathbf{r}) \hat{H} \mathrm{w}_{n0}(\mathbf{r}-\mathbf{R}_p). \tag{66}$$

Using Eq. (64), these matrix elements can be expressed in terms of the Kohn–Sham Hamiltonian in the Bloch representation $H_{m'n'}(\mathbf{k})$:

$$H_{mn}(\mathbf{R}_p) = \frac{1}{N_{\mathrm{uc}}} \sum_{m'n'\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}_p} U^\dagger_{mm'\mathbf{k}} H_{m'n'}(\mathbf{k}) U_{n'n\mathbf{k}}, \tag{67}$$

where the matrix $U_{mn\mathbf{k}}$ is the same as in Eq. (64). Once $H_{mn}(\mathbf{R}_p)$ has been determined, Eq. (67) can be inverted to generate $H_{m'n'}(\mathbf{k})$ anywhere in the Brillouin zone.

Similarly to the electronic case, the dynamical matrix can be expressed in the phonon Wannier representation as[10]

$$D_{\kappa\alpha,\kappa'\alpha'}(\mathbf{R}_p) = \frac{1}{N_p} \sum_{\mathbf{q}\mu\nu} e^{-i\mathbf{q}\cdot\mathbf{R}_p} e^\dagger_{\kappa\alpha,\mu}(\mathbf{q}) D_{\mu\nu}(\mathbf{q}) e_{\kappa'\alpha',\nu}(\mathbf{q}), \tag{68}$$

where $D_{\mu\nu}(\mathbf{q})$ is the matrix element of the dynamical matrix in the Bloch representation for phonons. Once $D_{\kappa\alpha,\kappa'\alpha'}(\mathbf{R}_p)$ has been determined, this relation can be inverted to obtain the phonon eigenvectors and eigenvalues anywhere in the Brillouin zone; this is a standard procedure employed to compute phonon dispersion relations[7,202].

Equations (67) and (68) can be generalized to the case of electron–phonon matrix elements by considering Fourier transforms for both the Kohn–Sham states and for the vibrational eigenmodes. The resulting electron–phonon matrix elements in the Wannier representation are given by[10]:

$$g_{mn\kappa\alpha}(\mathbf{R}_p,\mathbf{R}_{p'}) = \frac{1}{N^2_{\mathrm{uc}}} \sum_{\mathbf{k},\mathbf{q}} e^{-i(\mathbf{k}\cdot\mathbf{R}_p+\mathbf{q}\cdot\mathbf{R}_{p'})} \sum_{m'n'\nu} \sqrt{\frac{2M_\kappa\omega_{\mathbf{q}\nu}}{\hbar}} e^\dagger_{\kappa\alpha,\nu}(\mathbf{q}) U^\dagger_{mm'\mathbf{k}+\mathbf{q}} g_{m'n'\nu}(\mathbf{k},\mathbf{q}) U_{n'n\mathbf{k}}. \tag{69}$$

Once the $g_{mn\kappa\alpha}(\mathbf{R}_p,\mathbf{R}_{p'})$ have been computed, this relation can be inverted to generate electron-phonon matrix elements anywhere in the Brillouin zone.

Equations (67)–(69) constitute the backbone of the interpolation engine of `EPW`. In practice, the interpolation module of `EPW` reads in the DFT electron density, dynamical matrices, and variations of the Kohn–Sham potential evaluated by `Quantum ESPRESSO` on a coarse Brillouin zone grid; computes the Kohn–Sham wavefunctions and electron–phonon matrix elements on this grid; calls the `Wannier90` code in library mode to obtain the Wannier matrices in Eq. (64); transforms Hamiltonian, dynamical matrix, and electron–phonon matrix elements in the Wannier representation; and interpolates all these quantities onto arbitrarily dense **k**- and **q**-point grids. A qualitative schematic of this process is shown in Fig. 8.

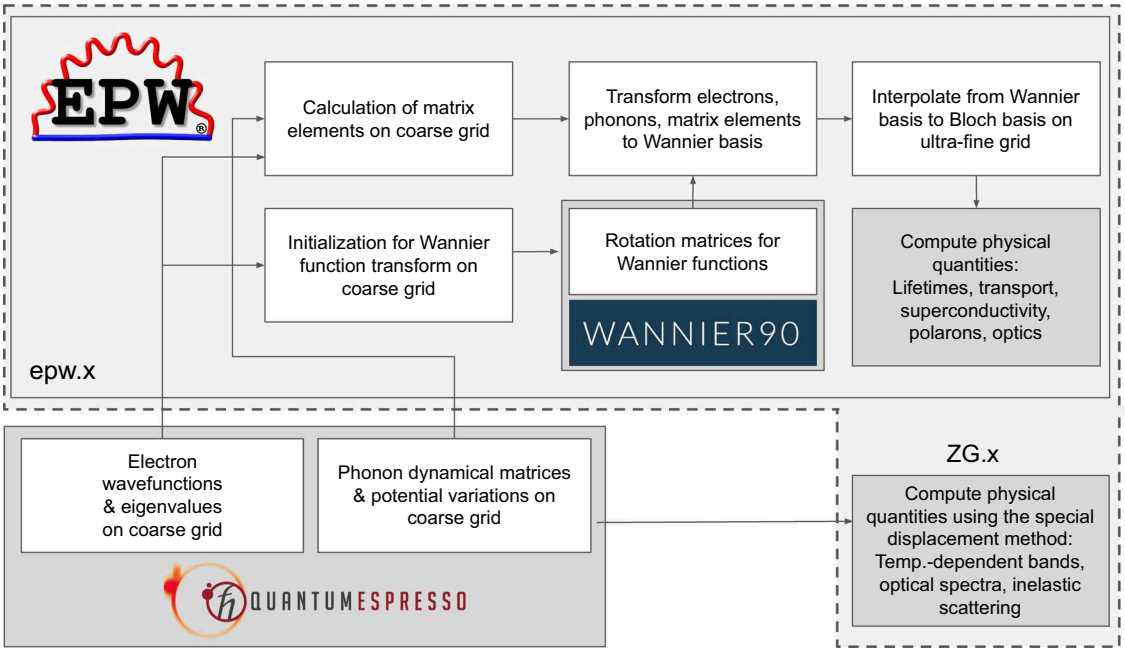

Fig. 8 Schematic overview of the interpolation engine of `EPW` and its relation to the `Quantum ESPRESSO` and `Wannier90` codes. The `EPW/ZG` module is a stand-alone supercell-based set of codes, as described in the "Results" section.

Treatment of long-range electron–phonon interactions

In semiconductors and insulators, the ionic displacements associated with a phonon can induce a variation of the Kohn–Sham potential that is long-ranged in nature. As a result, the electron–phonon matrix element $g_{mnv}(\mathbf{k},\mathbf{q})$ associated with longitudinal–optical (LO) phonons becomes singular in the long-wavelength limit $\mathbf{q}\to 0$. More specifically, in any material exhibiting non-vanishing Born effective charges, the matrix elements associated with LO phonons diverge as $\mathbf{q}/|\mathbf{q}|^2$ at small $\mathbf{q}$. In these cases, the assumption of locality that underpins the methodology described in the previous section ceases to hold, and the Wannier interpolation procedure needs to be modified to correctly capture the singularity.

The singularity in the matrix element can be dealt with by considering a multipole expansion of the Kohn–Sham potentials resulting from individual atomic displacements. The first order in this expansion is the dipole potential[203], which scales as $\mathbf{q}/|\mathbf{q}|^2$ and is responsible for the well-known Fröhlich electron–phonon interaction[204]. The modification of the electron–phonon interpolation method to include long-range effects consists of separating short-range and long-range parts in the matrix elements, as follows:

$$g_{mnv}(\mathbf{k},\mathbf{q}) = g^{\mathcal{S}}_{mnv}(\mathbf{k},\mathbf{q}) + g^{\mathcal{L},\mathrm{D}}_{mnv}(\mathbf{k},\mathbf{q}) + g^{\mathcal{L},\mathrm{Q}}_{mnv}(\mathbf{k},\mathbf{q}), \tag{70}$$

where the first term on the right-hand side is the short-range component, the second term is the dipole component, and the third term is the quadrupole component. Additional multipoles could be considered, but the dipole and quadrupole terms already lead to very accurate results. The dipole matrix element $g^{\mathcal{L},\mathrm{D}}_{mnv}(\mathbf{k},\mathbf{q})$ was derived in refs. [205,206] and reads:

$$g^{\mathcal{L},\mathrm{D}}_{mnv}(\mathbf{k},\mathbf{q}) = i\frac{4\pi}{\Omega}\frac{e^2}{4\pi\epsilon_0}\sum_{\kappa}\left(\frac{\hbar}{2M_\kappa\omega_{\mathbf{q}\nu}}\right)^{1/2}\sum_{\mathbf{G}\neq-\mathbf{q}} e^{-i(\mathbf{q}+\mathbf{G})\cdot\boldsymbol{\tau}_\kappa}\frac{\sum_p U_{mp\mathbf{k}+\mathbf{q}+\mathbf{G}}U^{\dagger}_{pn\mathbf{k}}}{(\mathbf{q}+\mathbf{G})\cdot\boldsymbol{\epsilon}^{\infty}\cdot(\mathbf{q}+\mathbf{G})}(\mathbf{q}+\mathbf{G})\cdot\mathbf{Z}^{*}_{\kappa}\cdot\mathbf{e}_{\kappa\nu}(\mathbf{q}). \tag{71}$$

In this expression, ϵ_0 is the vacuum permittivity, $\boldsymbol{\epsilon}^{\infty}$ is the high-frequency dielectric tensor of the material, $\mathbf{G}$ represents a reciprocal lattice vector, $\mathbf{Z}^{*}_{\kappa}$ is the Born effective charge tensor of the atom κ, the bracket indicates the integral over the BvK supercell, and we use the notation $\mathbf{a}\cdot\mathbf{B}\cdot\mathbf{c} = \sum_{\alpha\beta}a_\alpha B_{\alpha\beta}c_\beta$. The unitary matrices $U_{mn\mathbf{k}}$ in Eq. (71) are the Wannier transformation obtained following the standard procedure[207,208] and coming from the $\mathbf{q}+\mathbf{G}\to 0$ limit of the overlap integrals between Kohn–Sham wavefunctions $\langle\psi_{m\mathbf{k}+\mathbf{q}}|e^{i\mathbf{q}\cdot\mathbf{r}}|\psi_{n\mathbf{k}}\rangle = \sum_p U_{mp\mathbf{k}+\mathbf{q}}U^{\dagger}_{pn\mathbf{k}} + \mathcal{O}(\mathbf{q})$. This matrix element reduces to the standard Fröhlich interaction[204] when one considers parabolic electron bands and a dispersionless LO mode in a cubic material[186,205]. This term is of the order of $|\mathbf{q}|^{-1}$. The quadrupole contribution is the second term of the multipole expansion of the Kohn–Sham potential and is of the order of $|\mathbf{q}|^0$. The corresponding matrix element $g^{\mathcal{L},\mathrm{Q}}_{mnv}(\mathbf{k},\mathbf{q})$ was derived in refs. [40,41,187,209,210] and reads:

$$\begin{aligned} g^{\mathcal{L},\mathrm{Q}}_{mnv}(\mathbf{k},\mathbf{q}) = {} & \frac{4\pi}{\Omega}\frac{e^2}{4\pi\epsilon_0}\sum_{\kappa\alpha}\left(\frac{\hbar}{2M_\kappa\omega_{\mathbf{q}\nu}}\right)^{1/2}\sum_{\mathbf{G}\neq-\mathbf{q}}\frac{e^{-i(\mathbf{q}+\mathbf{G})\cdot\boldsymbol{\tau}_\kappa}e_{\kappa\alpha,\nu}(\mathbf{q})}{(\mathbf{q}+\mathbf{G})\cdot\boldsymbol{\epsilon}^{\infty}\cdot(\mathbf{q}+\mathbf{G})} \\ & \times\sum_{\beta\gamma}\Big[\tfrac{1}{2}Q_{\kappa,\alpha\beta\gamma}(q_\beta+G_\beta)(q_\gamma+G_\gamma)\sum_p U_{mp\mathbf{k}+\mathbf{q}+\mathbf{G}}U^{\dagger}_{pn\mathbf{k}} \\ & - Z^{*}_{\kappa,\alpha\beta}(q_\beta+G_\beta)\sum_{sp}U_{ms\mathbf{k}+\mathbf{q}+\mathbf{G}}(q_\gamma+G_\gamma)\left(A^{\mathrm{W}}_{sp\mathbf{k},\gamma}+\langle u^{\mathrm{W}}_{s\mathbf{k}}|V^{\mathrm{Hxc},\mathcal{E}_\gamma}|u^{\mathrm{W}}_{p\mathbf{k}}\rangle\right)U^{\dagger}_{pn\mathbf{k}}\Big]. \end{aligned} \tag{72}$$

In this expression, $\mathbf{Q}_\kappa$ is the dynamical quadrupole tensor[211], $V^{\mathrm{Hxc},\mathcal{E}_\gamma}(\mathbf{r})$ is the self-consistent potential induced by a uniform electric field $\mathcal{E}_\gamma$ along the Cartesian direction γ[41,187], $A^{\mathrm{W}}_{sp\mathbf{k},\gamma} = -i\langle\partial u^{\mathrm{W}}_{s\mathbf{k}}/\partial k_\gamma|u^{\mathrm{W}}_{p\mathbf{k}}\rangle$ is the Berry connection, and $u^{\mathrm{W}}_{p\mathbf{k}}$ is the smooth Bloch basis[187]. In `EPW`, the calculation of the Berry connection term is automatically activated for quadrupole calculations when the velocity matrix is computed in the Wannier representation, cf. Eq. (27) of ref. [22]. This term can be non-negligible in particular for materials where the quadrupole tensor $\mathbf{Q}_\kappa$ vanishes by symmetry, such as SrO[187]. The $V^{\mathrm{Hxc},\mathcal{E}}$ term has been implemented in `Abinit`[40,41,212] and `EPW`[187,213] and has been tested for Si, GaP, GaAs, SrO, h-BN monolayer, MoS_2 monolayer, and InSe monolayer. In all these cases, it was found that the contribution from this term represents <0.1% of the total quadrupole correction[41,187]. Since this term is very small and its implementation is rather cumbersome, it has not been included in public releases of these codes.

The calculation strategy employed by `EPW` is as follows. First, the complete matrix elements $g_{mnv}(\mathbf{k},\mathbf{q})$ are evaluated on coarse $\mathbf{k}$ and $\mathbf{q}$ grids using DFPT. Second, the long-range contributions $g^{\mathcal{L},\mathrm{D}}_{mnv}(\mathbf{k},\mathbf{q})$

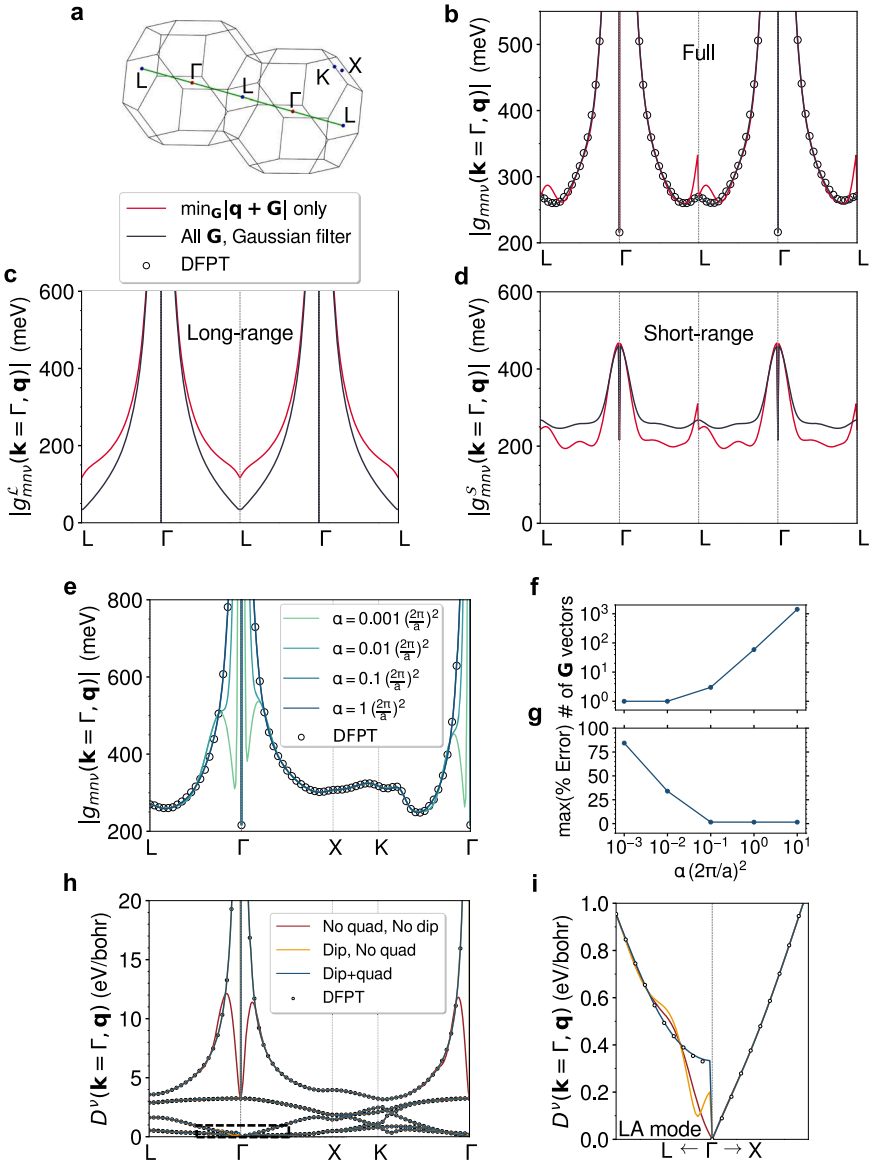

Fig. 9 Accurate computation of long-range electron–phonon matrix elements with EPW. **a** Brillouin zone of c-BN, the system used in this example. **b**–**d** Full, long-range and short-range electron–phonon matrix elements, respectively, for the LO phonon mode along the **q**-path shown by the green line in (**a**). The red lines are for the choice of including only the dominant reciprocal lattice vector in the **G** sums. The dark blue lines correspond to the choice of applying a Gaussian filter, the procedure implemented in EPW. Circles are the reference data from explicit DFPT calculations. **e** Sensitivity of the interpolation on the α parameter of the Gaussian filter. **f** Number of **G** vectors included in the sum as a function of α, and **g** corresponding maximum relative error with respect to DFPT reference data. **h** Comparison between the descriptor $D_\nu(\mathbf{k}=0,\mathbf{q})$ obtained from direct DFPT calculations (white circles) and from Wannier interpolation: without including long-range contributions (red line); including the dipole term (orange line); and including both dipole and quadrupole terms (blue line). In each case, the corresponding corrections to the phonon dynamical matrix are also included. **i** Zoom over the area highlighted in (**h**), focusing on the LA phonon mode.

and $g^{\mathcal{L},Q}_{mnv}(\mathbf{k},\mathbf{q})$ are subtracted from the DFPT matrix elements using Eqs. (71) and (72), leaving out the short-range component $g^{\mathcal{S}}_{mnv}(\mathbf{k},\mathbf{q})$ on the coarse grids. Third, the standard Wannier electron–phonon interpolation of ref. [10] is applied to the short-range component only. And fourth, the long-range contributions are added back using Eqs. (71) and (72) on the fine **k** and **q** grids.

The infinite sum over the **G** vectors in Eqs. (71) and (72) ensure the periodicity of the matrix elements in reciprocal space. In practical calculations, this sum might be computationally demanding to converge. In principle, periodicity could be enforced by including only one reciprocal lattice vector in the sum, namely $\tilde{\mathbf{G}}_{\mathbf{q}}$ such that $|\mathbf{q}+\tilde{\mathbf{G}}_{\mathbf{q}}| = \min_{\mathbf{G}}|\mathbf{q}+\mathbf{G}|$. However, this choice introduces derivative discontinuities in $g^{\mathcal{L},\mathrm{D}}_{mnv}(\mathbf{k},\mathbf{q})$ at the Brillouin zone boundaries, which in turn cause spurious oscillations in the interpolation of $g^{\mathcal{S}}_{mnv}(\mathbf{k},\mathbf{q})$. We illustrate this point in Fig. 9a–d, focusing on the longitudinal-optical (LO) phonon of cubic boron nitride (c-BN).

An alternative strategy to avoid the sum over **G** is to cut off the interaction range in Eqs. (71) and (72) using a Gaussian filter, $\exp[-(\mathbf{q}+\mathbf{G})\cdot\boldsymbol{\epsilon}^{\infty}\cdot(\mathbf{q}+\mathbf{G})/4\alpha]$. This choice finds motivation in the Ewald summation method that is commonly employed to evaluate the non-analytic contribution to the dynamical matrix in polar materials[7,8,202]. It was employed in previous versions of EPW as well as in other codes[13,40,206], and it avoids derivative discontinuity. However, it does not preserve the periodicity of the matrix elements in reciprocal space, which is required by the periodic gauge. In EPW,

to maintain the periodicity of the matrix elements, we sum over a shell of **G** vectors centered around $\tilde{\mathbf{G}}_{\mathbf{q}}$, and then apply the Gaussian filter. Figure 9c shows how this procedure yields periodic and smooth long-range matrix elements with the correct behavior next to the singularities, and enables accurate interpolation of the complete matrix elements, as shown in Fig. 9b.

The a parameter in the Gaussian filter is chosen in such a way as to ensure accurate interpolation (which requires large a) whilst using as few **G** vectors as possible in Eqs. (71) and (72) (which requires small a). In Fig. 9e we compare the interpolated matrix element $g_{mnv}(\mathbf{k},\mathbf{q})$ to explicit DFPT calculations, for different values of a, with the reciprocal space summation restricted to those **G** vectors such that $(\mathbf{q}+\mathbf{G})\cdot\boldsymbol{\epsilon}^{\infty}\cdot(\mathbf{q}+\mathbf{G})<56a$. With this choice, only **G**-vectors yielding a value of the Gaussian filter larger than $\exp(-14)\simeq 10^{-6}$ are included in the sum. We find that when a is chosen to match the size of the Brillouin zone [$a=1\,(2\pi/a)^2$, a is the lattice parameter], good interpolation is achieved. We note that, for layered materials, more **G** vectors are automatically included in the sum along the perpendicular direction due to the anisotropy of the dielectric tensor. For completeness, in Fig. 9f we show how the number of required **G** vectors increases with a, and in Fig. 9g we show the interpolation error in the matrix elements as a function of a. We emphasize that, since EPW implements this Gaussian filter, caution should be used when extracting dipolar and quadrupolar matrix elements from EPW for separate post-processing, because $g^{\mathcal{L},\mathrm{D}}_{mnv}(\mathbf{k},\mathbf{q})$ and $g^{\mathcal{L},\mathrm{Q}}_{mnv}(\mathbf{k},\mathbf{q})$ have the expected **q**-dependence only at long wavelength.

Figure 9h, i illustrate the importance of describing dipole and quadrupole interactions using Eqs. (71) and (72), for the case of c-BN. To facilitate the comparison between the interpolation results and explicit DFPT calculations for long-wavelength acoustic modes, we use the following descriptor which removes the factor $\omega_{\mathbf{q}\nu}^{-1/2}$ from the matrix element and averages over electronic degeneracies:

$$D_{\nu}(\mathbf{k},\mathbf{q})=\frac{1}{\hbar N_{\mathrm{W}}}\left[2\rho\Omega\hbar\omega_{\mathbf{q}\nu}\sum\nolimits_{mn}|g_{mnv}(\mathbf{k},\mathbf{q})|^2\right]^{1/2}. \tag{73}$$

Here, the sum over bands is carried over the N_{W} states in the Wannier manifold and ρ is the mass density of the crystal. $D_{\nu}(\mathbf{k},\mathbf{q})$ has units of energy divided by length and can be thought of as a deformation potential of sort. In Fig. 9h, i, we choose the manifold composed by the top three valence bands of c-BN, and set $\mathbf{k}=0$. We see that the dipole term is necessary to correctly describe the singular behavior of the LO phonon at long wavelength, and the quadrupole term is necessary to correctly describe the discontinuous behavior of the longitudinal acoustic (LA) phonons at a long wavelength [see expanded view in Fig. 9i]. The inclusion of both terms guarantees a high-quality interpolation of the electron-phonon matrix elements across the entire Brillouin zone. In particular, in this example, we find that the deformation potential calculated by including dipole and quadrupole terms agrees with direct DFPT calculations within 0.01 eV/Bohr (<1% of the maximum value) for the LA mode of c-BN along the ΓL line near the zone center. In refs. [22,187] a number of additional comparisons between the present method and direct DFPT calculations are discussed for standard three-dimensional (3D) and two-dimensional (2D) semiconductors.

The EPW code also implements long-range corrections for the phonon dynamical matrix by computing dipole–dipole, dipole–quadrupole, and quadrupole–quadrupole terms as discussed in refs. [7,8,202,211]. Systematic tests of the interpolation procedure described in this section are provided in ref. [22].

The expressions for the long-range dipole and quadrupole contributions to the electron–phonon matrix elements are given in Eqs. (71) and (72) are for 3D bulk crystals. Several generalizations of these expressions to the case of 2D materials and the transition from 3D to 2D have been proposed during the past few years[186,187,213–216]. In the first of such generalizations[214], the authors obtain the long-wavelength limit of polar electron–phonon couplings in 2D by evaluating the average dielectric polarization of a 2D slab enclosing the layer and removing the interaction between periodic replicas via Coulomb truncation. In ref. [215], the authors additionally consider the contributions to the polar matrix elements arising from out-of-plane dipoles and demonstrate an improvement in the interpolation accuracy. In addition to these studies, more recent work by some of us focused on including quadrupole corrections for 2D materials[187] and on establishing the connection between polar matrix elements in 3D and 2D materials[186]. In particular, in ref. [187], the 2D long-range matrix elements to quadrupolar order are given by

$$g^{\mathcal{L},2\mathrm{D}}_{mnv}(\mathbf{k},\mathbf{q})=\left(\frac{\hbar}{2\omega_{\mathbf{q}\nu}}\right)^{\frac{1}{2}}\sum_{\kappa\alpha}\frac{e_{\kappa\alpha,\nu}(\mathbf{q})}{\sqrt{M_{\kappa}}}\sum_{\mathbf{G}\neq-\mathbf{q}}\sum_{sp}U_{ms\mathbf{k}+\mathbf{q}+\mathbf{G}}\langle u^{\mathrm{W}}_{s\mathbf{k}+\mathbf{q}+\mathbf{G}}|V^{\mathcal{L},2\mathrm{D}}_{\mathbf{q}+\mathbf{G}\kappa\alpha}|u^{\mathrm{W}}_{p\mathbf{k}+\mathbf{q}}\rangle U^{\dagger}_{pn\mathbf{k}}, \tag{74}$$

where the 2D long-range scattering potential to first order in **q** is

$$V^{\mathcal{L},2\mathrm{D}}_{\mathbf{q}\kappa\alpha}(\mathbf{r})=\frac{\pi e}{S}\frac{f(|\mathbf{q}|)}{|\mathbf{q}|}\mathrm{e}^{-i\mathbf{q}\cdot\boldsymbol{\tau}_{\kappa}}\left[\frac{1}{\tilde{\epsilon}^{\parallel}(\mathbf{q})}\sum_{\beta}\left\{2iq_{\beta}Z^{*}_{\kappa,\alpha\beta}+\sum_{\gamma}q_{\beta}q_{\gamma}Q_{\kappa,\alpha\beta\gamma}-|\mathbf{q}|^2Q_{\kappa,\alpha zz}\right.\right.$$
$$\left.\left.-2q_{\beta}Z^{*}_{\kappa,\alpha\beta}q_{\gamma}V^{\mathrm{Hxc},\mathcal{E}_{\gamma}}(\mathbf{r})\right\}+\frac{1}{\tilde{\epsilon}^{\perp}(\mathbf{q})}\left\{2|\mathbf{q}|^2Z_{\kappa,\alpha z}\left[z+V^{\mathrm{Hxc},\mathcal{E}_z}(\mathbf{r})\right]\right\}\right], \tag{75}$$

where S is the unit-cell area, and the dielectric functions appearing at the denominators are

$$\tilde{\epsilon}^{\parallel}(\mathbf{q})=1+\frac{2\pi f(|\mathbf{q}|)}{|\mathbf{q}|}\mathbf{q}\cdot\boldsymbol{\alpha}^{\parallel}\cdot\mathbf{q}, \tag{76}$$

$$\tilde{\epsilon}^{\perp}(\mathbf{q})=1-2\pi|\mathbf{q}|f(|\mathbf{q}|)\,\alpha^{\perp}. \tag{77}$$

In these expressions, $\boldsymbol{\alpha}^{\parallel}$ and $\alpha^{\perp}$ are macroscopic polarizabilities in- and out-of-plane, respectively, and $f(q)=1-\tanh(qL/2)$ is a range-separation function and the distance L is determined according to Eq. (63) of ref. [187]. The 2D extension of the overlap integral is provided in Eqs. (46) and (47) of ref. [187]. This formulation also uses Coulomb truncation and generalizes the previous approaches of refs. [214,215] to the case of quadrupole corrections. In the approach of Ref. [186], the authors propose a unified formulation of polar electron–phonon interactions in 3D and 2D that reduces to the approach of refs. [206,217] in 3D, and to the approach of refs. [214,215] in 2D. The theory is developed along the same lines as in ref. [217], by evaluating the potential of point dipoles resulting from the atomic displacements. The key difference is that, in this case, the dipoles belong to a supercell consisting of a dielectric stack where one component is the material under consideration, and the other is a homogeneous dielectric (or vacuum). By solving this electrostatic problem, the authors obtain the following matrix element:

$$g_{mnv}(\mathbf{k},\mathbf{q})=\frac{e^2}{2\epsilon_0\epsilon_{\infty,\mathrm{env}}\Omega}\left(\frac{\hbar}{2\omega_{\mathbf{q}\nu}}\right)^{1/2}\sum_{\mathbf{G}\neq-\mathbf{q}}\frac{\langle u_{m\mathbf{k}+\mathbf{q}+\mathbf{G}}|u_{n\mathbf{k}}\rangle}{|\mathbf{q}_{\parallel}+\mathbf{G}_{\parallel}|}\sum_{\kappa}M_{\kappa}^{-1/2}\mathrm{e}^{-i(\mathbf{q}_{\parallel}+\mathbf{G}_{\parallel})\cdot\boldsymbol{\tau}_{\kappa\parallel}}$$
$$\times\sum_{\alpha\beta}Z^{*}_{\kappa,\alpha\beta}e_{\kappa\beta,\nu}(\mathbf{q})\left[\delta_{\alpha,\parallel}\,i(\mathbf{q}+\mathbf{G})_{\alpha}K(\mathbf{q}+\mathbf{G},\tau_{\kappa z})-\delta_{\alpha,z}\frac{\partial K(\mathbf{q}+\mathbf{G},\tau_{\kappa z})}{\partial\tau_{\kappa z}}\right], \tag{78}$$

where $\epsilon_{\infty,\mathrm{env}}$ is the dielectric constant of the homogeneous dielectric surrounding the 2D material, and $K(\mathbf{q}+\mathbf{G},\tau_{\kappa z})$ is a kernel that depends on the thickness of the 2D layer and the surrounding medium, as well as their dielectric constants. The expression for K is given by Eq. (A1) of ref. [186]. The main practical difference between this formulation and other approaches is that it can be used without the need for Coulomb truncation.

DATA AVAILABILITY

The data presented in each figure are available from the corresponding author upon reasonable request.

CODE AVAILABILITY

The codes used for the current study, namely Quantum ESPRESSO, Wannier90, and EPW, are all open-source software distributed via their corresponding websites. In particular, the EPW code described in this manuscript can be downloaded from https://epw-code.org. The data presented in each figure are available from the corresponding author upon reasonable request.

Received: 15 February 2023; Accepted: 6 August 2023;
Published online: 25 August 2023

REFERENCES

1. Giustino, F. Electron–phonon interactions from first principles. *Rev. Mod. Phys.* **89**, 015003 (2017).
2. Ziman, J. M. *Electrons and Phonons* (Oxford University Press, 1960).
3. Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **136**, B864–B871 (1964).
4. Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133–A1138 (1965).
5. Baroni, S., Giannozzi, P. & Testa, A. Green's-function approach to linear response in solids. *Phys. Rev. Lett.* **58**, 1861–1864 (1987).
6. Savrasov, S. Y. Linear response calculations of lattice dynamics using muffin-tin basis sets. *Phys. Rev. Lett.* **69**, 2819–2822 (1992).
7. Gonze, X. & Lee, C. Dynamical matrices, Born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory. *Phys. Rev. B* **55**, 10355–10368 (1997).
8. Baroni, S., de Gironcoli, S., Corso, A. D. & Giannozzi, P. Phonons and related properties of extended systems from density-functional perturbation theory. *Rev. Mod. Phys.* **73**, 515 (2001).
9. Marzari, N., Mostofi, A. A., Yates, J. R., Souza, I. & Vanderbilt, D. Maximally localized Wannier functions: theory and applications. *Rev. Mod. Phys.* **84**, 1419–1475 (2012).
10. Giustino, F., Cohen, M. L. & Louie, S. G. Electron–phonon interaction using Wannier functions. *Phys. Rev. B* **76**, 165108 (2007).
11. GNU General Public License v3, Free Software Foundation. Gnu general public license v3. https://www.gnu.org/licenses (2007).
12. Giannozzi, P. et al. Advanced capabilities for materials modelling with QUANTUM ESPRESSO. *J. Phys.: Condens. Matter* **29**, 465901 (2017).
13. Zhou, J.-J. et al. Perturbo: A software package for ab initio electron–phonon interactions, charge transport and ultrafast dynamics. *Comput. Phys. Commun.* **264**, 107970 (2021).
14. Protik, N. H., Li, C., Pruneda, M., Broido, D. & Ordejón, P. The elphbolt ab initio solver for the coupled electron–phonon Boltzmann transport equations. *npj Comput. Mater.* **8**, 28 (2022).
15. Cepellotti, A., Coulter, J., Johansson, A., Fedorova, N. S. & Kozinsky, B. Phoebe: a high-performance framework for solving phonon and electron Boltzmann transport equations. *J. Phys. Mater.* **5**, 035003 (2022).
16. Marchese, G. et al. Epiq (2022) (last accessed 27 January 2023); https://the-epiq-team.gitlab.io/epiq-site.
17. Noffsinger, J. et al. EPW: a program for calculating the electron–phonon coupling using maximally localized Wannier functions. *Comput. Phys. Commun.* **181**, 2140–2148 (2010).
18. Poncé, S., Margine, E. R., Verdi, C. & Giustino, F. EPW: electron–phonon coupling, transport and superconducting properties using maximally localized Wannier functions. *Comput. Phys. Commun.* **209**, 116–133 (2016).
19. Message Passing Interface Forum. MPI: A Message-Passing Interface Standard Version 5.0 (Message Passing Interface Forum, 2022).
20. Dagum, L. & Menon, R. Openmp: an industry standard API for shared-memory programming. *Comput. Sci. Eng.* **5**, 46–55 (1998).
21. The HDF Group. Hierarchical Data Format, version 5 (1997-NNNN). https://www.hdfgroup.org/HDF5/.
22. Poncé, S. et al. First-principles predictions of Hall and drift mobilities in semiconductors. *Phys. Rev. Res.* **3**, 043022 (2021).
23. Kadanoff, L. P. & Baym, G. *Quantum Statistical Mechanics: Green's Function Methods in Equilibrium and Nonequilibrium Problems*. Frontiers in Physics. *A Lecture Note and Reprint Series* (Benjamin, 1962).
24. Poncé, S., Li, W., Reichardt, S. & Giustino, F. First-principles calculations of charge carrier mobility and conductivity in bulk semiconductors and two-dimensional materials. *Rep. Prog. Phys.* **83**, 036501 (2020).
25. Mahan, G. D. Quantum transport equation for electric and magnetic fields. *Phys. Rep.* **145**, 251–318 (1987).
26. Macheda, F. *Ab Initio Prediction of Thermomagnetic and Thermoelectric Transport Phenomena in 3d and 2d Materials*. Ph.D. thesis, King's College London (2021).
27. Restrepo, O. D., Varga, K. & Pantelides, S. T. First-principles calculations of electron mobilities in silicon: phonon and Coulomb scattering. *Appl. Phys. Lett.* **94**, 212103 (2009).
28. Li, W. Electrical transport limited by electron–phonon coupling from Boltzmann transport equation: an ab initio study of Si, Al, and MoS_2. *Phys. Rev. B* **92**, 075405 (2015).
29. Fiorentini, M. & Bonini, N. Thermoelectric coefficients of *n*-doped silicon from first principles via the solution of the Boltzmann transport equation. *Phys. Rev. B* **94**, 085204 (2016).
30. Zhou, J.-J. & Bernardi, M. Ab initio electron mobility and polar phonon scattering in GaAs. *Phys. Rev. B* **94**, 201201 (2016).
31. Poncé, S., Margine, E. R. & Giustino, F. Towards predictive many-body calculations of phonon-limited carrier mobilities in semiconductors. *Phys. Rev. B* **97**, 121201(R) (2018).
32. Macheda, F. & Bonini, N. Magnetotransport phenomena in p-doped diamond from first principles. *Phys. Rev. B* **98**, 201201(R) (2018).
33. Sohier, T., Campi, D., Marzari, N. & Gibertini, M. Mobility of two-dimensional materials from first principles in an accurate and automated framework. *Phys. Rev. Mater.* **2**, 114010 (2018).
34. Ma, J., Nissimagoudar, A. S. & Li, W. First-principles study of electron and hole mobilities of Si and GaAs. *Phys. Rev. B* **97**, 045201 (2018).
35. Poncé, S., Schlipf, M. & Giustino, F. Origin of low carrier mobilities in halide perovskites. *ACS Energy Lett.* **4**, 456–463 (2019).
36. Poncé, S., Jena, D. & Giustino, F. Hole mobility of strained GaN from first principles. *Phys. Rev. B* **100**, 085204 (2019).
37. Poncé, S., Jena, D. & Giustino, F. Route to high hole mobility in GaN via reversal of crystal-field splitting. *Phys. Rev. Lett.* **123**, 096602 (2019).
38. Lee, N.-E., Zhou, J. J., Chen, H. Y. & Bernardi, M. Ab initio electron-two-phonon scattering in GaAs from next-to-leading order perturbation theory. *Nat. Commun.* **11**, 1607 (2020).
39. Poncé, S. & Giustino, F. Structural, electronic, elastic, power, and transport properties of β-Ga_2O_3 from first principles. *Phys. Rev. Res.* **2**, 033102 (2020).
40. Brunin, G. et al. Phonon-limited electron mobility in Si, GaAs, and GaP with exact treatment of dynamical quadrupoles. *Phys. Rev. B* **102**, 094308 (2020).
41. Brunin, G. et al. Electron-phonon beyond Fröhlich: dynamical quadrupoles in polar and covalent solids. *Phys. Rev. Lett.* **125**, 136601 (2020).
42. Macheda, F., Poncé, S., Giustino, F. & Bonini, N. Theory and computation of Hall scattering factor in graphene. *Nano Lett.* **20**, 8861–8865 (2020).
43. Broyden, C. G. A class of methods for solving nonlinear simultaneous equations. *Math. Comput.* **19**, 577 (1965).
44. Li, W., Carrete, J., Katcho, N. A. & Mingo, N. ShengBTE: a solver of the Boltzmann transport equation for phonons. *Comput. Phys. Commun.* **185**, 1747–1758 (2014).
45. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865 (1996).
46. Hamann, D. R. Optimized norm-conserving Vanderbilt pseudopotentials. *Phys. Rev. B* **88**, 085117 (2013).
47. van Setten, M. J. et al. The PseudoDojo: training and grading a 85 element optimized norm-conserving pseudopotential table. *Comput. Phys. Commun.* **226**, 39–54 (2018).
48. Di Paola, C., Macheda, F., Laricchia, S., Weber, C. & Bonini, N. First-principles study of electronic transport and structural properties of $Cu_{12}Sb_4S_{13}$ in its high-temperature phase. *Phys. Rev. Res.* **2**, 033055 (2020).
49. Ashcroft, N. W. & Mermin, N. D.*Solid State Physics* (Saunders College Publishing, 1976).
50. Lundstrom, M. *Fundamentals of Carrier Transport* 2nd edn (Cambridge University Press, 2000).
51. Brooks, H. Theory of the electrical properties of germanium and silicon. In *Advances in Electronics and Electron Physics* (ed., Marton, L.), Ch. 2, 85–182 (Academic Press, 1955).
52. Debye, P. P. & Conwell, E. M. Electrical properties of *n*-type germanium. *Phys. Rev.* **93**, 693–706 (1954).
53. Restrepo, O. D., Varga, K. & Pantelides, S. T. First-principles calculations of electron mobilities in silicon: phonon and Coulomb scattering. *Appl. Phys. Lett.* **94**, 212103 (2009).
54. Graziosi, P., Kumarasinghe, C. & Neophytou, N. Material descriptors for the discovery of efficient thermoelectrics. *ACS Appl. Energy Mater.* **3**, 5913–5926 (2020).
55. Ganose, A. M. et al. Efficient calculation of carrier scattering rates from first principles. *Nat. Commun.* **12**, 2222 (2021).
56. Lu, I.-T., Zhou, J.-J., Park, J. & Bernardi, M. First-principles ionized-impurity scattering and charge transport in doped materials. *Phys. Rev. Mater.* **6**, L010801 (2022).

57. Leveillee, J., Zhang, X., Kioupakis, E. & Giustino, F. Ab initio calculation of carrier mobility in semiconductors including ionized-impurity scattering. *Phys. Rev. B* **107**, 125207 (2023).
58. Kohn, W. & Luttinger, J. M. Quantum theory of electrical transport phenomena. *Phys. Rev.* **108**, 590–611 (1957).
59. Lindhard, J. On the properties of a gas of charged particles. *Dan. Mat. Fys. Medd.* **28**, 3–58 (1954).
60. Sanders, N. & Kioupakis, E. Phonon- and defect-limited electron and hole mobility of diamond and cubic boron nitride: a critical comparison. *Appl. Phys. Lett.* **119**, 062101 (2021).
61. Schlipf, M. & Gygi, F. Optimization algorithm for the generation of ONCV pseudopotentials. *Comput. Phys. Commun.* **196**, 36 (2015).
62. Canali, C., Jacoboni, C., Nava, F., Ottaviani, G. & Alberigi-Quaranta, A. Electron drift velocity in silicon. *Phys. Rev. B* **12**, 2265–2284 (1975).
63. Morin, F. J. & Maita, J. P. Electrical properties of silicon containing arsenic and boron. *Phys. Rev.* **96**, 28–35 (1954).
64. Jacoboni, C., Canali, C., Ottaviani, G. & Alberigi Quaranta, A. A review of some charge transport properties of silicon. *Solid State Electron.* **20**, 77–89 (1977).
65. Caruso, F. & Giustino, F. Theory of electron–plasmon coupling in semiconductors. *Phys. Rev. B* **94**, 115208 (2016).
66. Misiakos, K. & Tsamakis, D. Accurate measurements of the silicon intrinsic carrier density from 78 to 340 K. *J. Appl. Phys.* **74**, 3293–3297 (1993).
67. McMillan, W. L. Transition temperature of strong-coupled superconductors. *Phys. Rev.* **167**, 331 (1968).
68. Eliashberg, G. Interactions between electrons and lattice vibrations in a superconductor. *Sov. Phys.-JETP* **11**, 696 (1960).
69. Eliashberg, G. Temperature Green's function for electrons in a superconductor. *Sov. Phys.-JETP* **12**, 1000 (1961).
70. Oliveira, L. N., Gross, E. K. U. & Kohn, W. Density-functional theory for superconductors. *Phys. Rev. Lett.* **60**, 2430–2433 (1988).
71. Lüders, M. et al. Ab initio theory of superconductivity. I. Density functional formalism and approximate functionals. *Phys. Rev. B* **72**, 024545 (2005).
72. Marques, M. A. L. et al. Ab initio theory of superconductivity. II. Application to elemental metals. *Phys. Rev. B* **72**, 024546 (2005).
73. Sanna, A., Pellegrini, C. & Gross, E. K. U. Combining Eliashberg theory with density functional theory for the accurate prediction of superconducting transition temperatures and gap functions. *Phys. Rev. Lett.* **125**, 057001 (2020).
74. Allen, P. B. & Dynes, R. C. Transition temperature of strong-coupled superconductors reanalyzed. *Phys. Rev. B* **12**, 905–922 (1975).
75. Carbotte, J. P. Properties of boson-exchange superconductors. *Rev. Mod. Phys.* **62**, 1027–1157 (1990).
76. Sanna, A. et al. Ab initio Eliashberg theory: making genuine predictions of superconducting features. *J. Phys. Soc. Jpn.* **87**, 041012 (2018).
77. Xie, S. R. et al. Machine learning of superconducting critical temperature from Eliashberg theory. *npj Comput. Mater.* **8**, 1–8 (2022).
78. Gor'kov, L. P. Microscopic derivation of the Ginzburg–Landau equations in the theory of superconductivity. *Sov. Phys.-JETP* **7**, 505 (1958).
79. Nambu, Y. Quasi-particles and gauge invariance in the theory of superconductivity. *Phys. Rev.* **117**, 648–663 (1960).
80. Scalapino, D. J., Schrieffer, J. R. & Wilkins, J. W. Strong-coupling superconductivity. I. *Phys. Rev.* **148**, 263–279 (1966).
81. Scalapino, D. J. *The Electron–phonon Interaction and Strong-coupling Superconductors*, Vol. 1 (Dekker, 1969).
82. Allen, P. B. & Mitrović, B. *Theory of Superconducting T_c*, Vol. 37 of *Solid State Physics* (Academic Press, 1983).
83. Choi, H. J., Cohen, M. L. & Louie, S. G. Anisotropic Eliashberg theory of MgB_2: T_c, isotope effects, superconducting energy gaps, quasiparticles, and specific heat. *Phys. C: Supercond.* **385**, 66–74 (2003).
84. Marsiglio, F. & Carbotte, J. P. *Electron–Phonon Superconductivity* (Springer, 2008).
85. Margine, E. R. & Giustino, F. Anisotropic Migdal–Eliashberg theory using Wannier functions. *Phys. Rev. B* **87**, 024505 (2013).
86. Marsiglio, F. Eliashberg theory: a short review. *Ann. Phys.* **417**, 168102 (2020).
87. Migdal, A. Interaction between electrons and lattice vibrations in a normal metal. *Sov. Phys.-JETP* **34**, 996 (1958).
88. Hedin, L. New method for calculating the one-particle Green's function with application to the electron-gas problem. *Phys. Rev.* **139**, A796 (1965).
89. Hybertsen, M. S. & Louie, S. G. Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies. *Phys. Rev. B* **34**, 5390 (1986).
90. Lee, K.-H., Chang, K. J. & Cohen, M. L. First-principles calculations of the coulomb pseudopotential μ^*: application to Al. *Phys. Rev. B* **52**, 1425–1428 (1995).
91. Margine, E. R., Lambert, H. & Giustino, F. Electron–phonon interaction and pairing mechanism in superconducting Ca-intercalated bilayer graphene. *Sci. Rep.* **6**, 1–6 (2016).
92. Aperis, A. & Oppeneer, P. M. Multiband full-bandwidth anisotropic Eliashberg theory of interfacial electron–phonon coupling and high T_c superconductivity in FeSe/$SrTiO_3$. *Phys. Rev. B* **97**, 060501 (2018).
93. Allen, P. B. Fermi-surface harmonics: a general method for nonspherical problems. application to Boltzmann and Eliashberg equations. *Phys. Rev. B* **13**, 1416–1427 (1976).
94. Vidberg, H. & Serene, J. Solving the Eliashberg equations by means of N-point Padé approximants. *J. Low Temp. Phys.* **29**, 179–192 (1977).
95. Leavens, C. & Ritchie, D. Extension of the N-point Padé approximants solution of the Eliashberg equations to $T \sim T_c$. *Solid State Commun.* **53**, 137–142 (1985).
96. Marsiglio, F., Schossmann, M. & Carbotte, J. P. Iterative analytic continuation of the electron self-energy to the real axis. *Phys. Rev. B* **37**, 4965–4969 (1988).
97. Morel, P. & Anderson, P. W. Calculation of the superconducting state parameters with retarded electron–phonon interaction. *Phys. Rev.* **125**, 1263–1271 (1962).
98. Heil, C. et al. Origin of superconductivity and latent charge density wave in NbS_2. *Phys. Rev. Lett.* **119**, 087003 (2017).
99. Heil, C., di Cataldo, S., Bachelet, G. B. & Boeri, L. Superconductivity in sodalite-like yttrium hydride clathrates. *Phys. Rev. B* **99**, 220502 (2019).
100. Di Cataldo, S., Heil, C., von der Linden, W. & Boeri, L. $LaBH_8$: towards high-T_c low-pressure superconductivity in ternary superhydrides. *Phys. Rev. B* **104**, L020511 (2021).
101. Pickett, W. E. Generalization of the theory of the electron–phonon interaction: thermodynamic formulation of superconducting- and normal-state properties. *Phys. Rev. B* **26**, 1186–1207 (1982).
102. Ummarino, G. A. C. *Eliashberg Theory*, Vol. 3 of *Lecture Notes of the Autumn School Correlated Electrons* (Forschungszentrum Jülich GmbH Institute for Advanced Simulation, 2013).
103. Davydov, A. et al. Ab initio theory of plasmonic superconductivity within the Eliashberg and density-functional formalisms. *Phys. Rev. B* **102**, 214508 (2020).
104. Kafle, G. P. et al. Ab initio study of Li–Mg–B superconductors. *Phys. Rev. Mater.* **6**, 084801 (2022).
105. Sano, W., Koretsune, T., Tadano, T., Akashi, R. & Arita, R. Effect of van Hove singularities on high-T_c superconductivity in H_3S. *Phys. Rev. B* **93**, 094525 (2016).
106. Schrodi, F., Oppeneer, P. M. & Aperis, A. Full-bandwidth Eliashberg theory of superconductivity beyond Migdal's approximation. *Phys. Rev. B* **102**, 024503 (2020).
107. Valla, T. et al. Quasiparticle spectra, charge-density waves, superconductivity, and electron-phonon coupling in 2H-$NbSe_2$. *Phys. Rev. Lett.* **92**, 086401 (2004).
108. Weber, F. et al. Extended phonon collapse and the origin of the charge-density wave in 2H-$NbSe_2$. *Phys. Rev. Lett.* **107**, 107403 (2011).
109. Leroux, M. et al. Strong anharmonicity induces quantum melting of charge density wave in 2H-$NbSe_2$ under pressure. *Phys. Rev. B* **92**, 140303(R) (2015).
110. Yokoya, T. et al. Fermi surface sheet-dependent superconductivity in 2H-$NbSe_2$. *Science* **294**, 2518 (2001).
111. Anikin, A. et al. Ultrafast dynamics in the high-symmetry and in the charge density wave phase of 2H-$NbSe_2$. *Phys. Rev. B* **102**, 205139 (2020).
112. Sanna, A. et al. Real-space anisotropy of the superconducting gap in the charge-density wave material 2H-$NbSe_2$. *npj Quantum Mater.* **7**, 6 (2022).
113. Klimeš, J., Bowler, D. R. & Michaelides, A. van der Waals density functionals applied to solids. *Phys. Rev. B* **83**, 195131 (2011).
114. Klimeš, J., Bowler, D. R. & Michaelides, A. Chemical accuracy for the van der Waals density functional. *J. Phys.: Condens. Matter* **22**, 022201 (2010).
115. Thonhauser, T. et al. Spin signature of nonlocal correlation binding in metal-organic frameworks. *Phys. Rev. Lett.* **115**, 136402 (2015).
116. Thonhauser, T. et al. van der Waals density functional: self-consistent potential and the nature of the van der Waals bond. *Phys. Rev. B* **76**, 125112 (2007).
117. Methfessel, M. & Paxton, A. T. High-precision sampling for Brillouin-zone integration in metals. *Phys. Rev. B* **40**, 3616 (1989).
118. Zheng, F. & Feng, J. Electron-phonon coupling and the coexistence of superconductivity and charge-density wave in monolayer $NbSe_2$. *Phys. Rev. B* **99**, 161119 (2019).
119. Wickramaratne, D., Khmelevskyi, S., Agterberg, D. F. & Mazin, I. I. Ising superconductivity and magnetism in $NbSe_2$. *Phys. Rev. X* **10**, 041003 (2020).
120. Das, S., Paudyal, H., Margine, E. R., Agterberg, D. F. & Mazin, I. I. Electron–phonon coupling and spin fluctuations in the Ising superconductor $NbSe_2$. *npj Comput. Mater.* **9**, 66 (2023).
121. Alexandrov, A. S. & Devreese, J. T.*Advances in Polaron Physics*, Vol. 159 of *Springer Series in Solid-State Sciences* (Springer, 2010).
122. Emin, D. *Polarons* (Cambridge University Press, 2012).
123. Franchini, C., Reticcioli, M., Setvin, M. & Diebold, U. Polarons in materials. *Nat. Rev. Mater.* **6**, 560 (2021).

124. Deskins, N. A. & Dupuis, M. Electron transport via polaron hopping in bulk TiO_2: a density functional theory characterization. *Phys. Rev. B* **75**, 195212 (2007).
125. Franchini, C., Kresse, G. & Podloucky, R. Polaronic hole trapping in doped $BaBiO_3$. *Phys. Rev. Lett.* **102**, 256402 (2009).
126. Lany, S. & Zunger, A. Polaronic hole localization and multiple hole binding of acceptors in oxide wide-gap semiconductors. *Phys. Rev. B* **80**, 085202 (2009).
127. Varley, J. B., Janotti, A., Franchini, C. & Van de Walle, C. G. Role of self-trapping in luminescence and p-type conductivity of wide-band-gap oxides. *Phys. Rev. B* **85**, 081109(R) (2012).
128. Setvin, M. et al. Direct view at excess electrons in TiO_2 rutile and anatase. *Phys. Rev. Lett.* **113**, 086402 (2014).
129. Kokott, S., Levchenko, S. V., Rinke, P. & Scheffler, M. First-principles supercell calculations of small polarons with proper account for long-range polarization effects. *New J. Phys.* **20**, 033023 (2018).
130. Falletta, S. & Pasquarello, A. Polarons free from many-body self-interaction in density functional theory. *Phys. Rev. B* **106**, 125119 (2022).
131. Sio, W. H., Verdi, C., Poncé, S. & Giustino, F. Polarons from first principles, without supercells. *Phys. Rev. Lett.* **122**, 246403 (2019).
132. Sio, W. H., Verdi, C., Poncé, S. & Giustino, F. Ab initio theory of polarons: Formalism and applications. *Phys. Rev. B* **99**, 235139 (2019).
133. Lafuente-Bartolomé, J. et al. Unified approach to polarons and phonon-induced band structure renormalization. *Phys. Rev. Lett.* **129**, 076402 (2022).
134. Lafuente-Bartolomé, J. et al. Ab initio self-consistent many-body theory of polarons at all couplings. *Phys. Rev. B* **106**, 075119 (2022).
135. Maradudin, A. A. & Vosko, S. H. Symmetry properties of the normal vibrations of a crystal. *Rev. Mod. Phys.* **40**, 1–37 (1968).
136. Davidson, E. R. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J. Comput. Phys.* **17**, 87–94 (1975).
137. Makov, G. & Payne, M. C. Periodic boundary conditions in ab initio calculations. *Phys. Rev. B* **51**, 4014–4022 (1995).
138. Castner, T. G. & Känzig, W. The electronic structure of V-centers. *J. Phys. Chem. Solids* **3**, 178–195 (1957).
139. Jette, A. N., Gilbert, T. L. & Das, T. P. Theory of the self-trapped hole in the alkali halides. *Phys. Rev.* **184**, 884–894 (1969).
140. Fröhlich, H., Pelzer, H. & Zienau, S. XX. Properties of slow electrons in polar materials. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **41**, 221–242 (1950).
141. Holstein, T. Studies of polaron motion: Part I. The molecular-crystal model. *Ann. Phys.* **8**, 325–342 (1959).
142. Bassani, F. & Parravicini, G. P. *Electronic States and Optical Transitions in Solids* (Pergamon Press, 1975).
143. Noffsinger, J., Kioupakis, E., Van de Walle, C. G., Louie, S. G. & Cohen, M. L. Phonon-assisted optical absorption in silicon from first principles. *Phys. Rev. Lett.* **108**, 167402 (2012).
144. Giustino, F. *Materials Modelling Using Density Functional Theory: Properties and Predictions* (Oxford University Press, 2014).
145. Tiwari, S. & Giustino, F. Ab-initio many-body calculations of optical absorption spectra including both direct and indirect phonon-assisted processes (2023) (submitted).
146. Levine, Z. H. & Allan, D. C. Quasiparticle calculation of the dielectric response of silicon and germanium. *Phys. Rev. B* **43**, 4187–4207 (1991).
147. Rohlfing, M. & Louie, S. G. Electron-hole excitations and optical spectra from first principles. *Phys. Rev. B* **62**, 4927–4944 (2000).
148. Song, B. et al. Optical properties of cubic boron arsenide. *Appl. Phys. Lett.* **116**, 141903 (2020).
149. Zhang, X. & Kioupakis, E. Phonon-assisted optical absorption of SiC polytypes from first principles. *Phys. Rev. B* **107**, 115207 (2023).
150. Zhang, X., Shi, G., Leveillee, J. A., Giustino, F. & Kioupakis, E. Ab initio theory of free-carrier absorption in semiconductors. *Phys. Rev. B* **106**, 205203 (2022).
151. Mo, Y., Tang, H., Bansil, A. & Tao, J. Accurate lattice geometrical parameters and bulk moduli from a semilocal density functional. *AIP Adv.* **8**, 095209 (2018).
152. Haas, P., Tran, F. & Blaha, P. Calculation of the lattice constant of solids with semilocal functionals. *Phys. Rev. B* **79**, 085104 (2009).
153. Favot, F. & Dal Corso, A. Phonon dispersions: Performance of the generalized gradient approximation. *Phys. Rev. B* **60**, 11427–11431 (1999).
154. Deslippe, J. et al. BerkeleyGW: a massively parallel computer package for the calculation of the quasiparticle and optical properties of materials and nanostructures. *Comput. Phys. Commun.* **183**, 1269–1289 (2012).
155. Sze, S. M., Li, Y. & Ng, K. K. *Physics and Properties of Semiconductors—A Review*, Ch. 1, 5–75 (Wiley, 2006).
156. Wachs, A. L., Miller, T., Hsieh, T. C., Shapiro, A. P. & Chiang, T. C. Angle-resolved photoemission studies of Ge(111)-c(2 × 8), Ge(111)-(1 × 1)H, Si(111)-(7 × 87), and Si(100)-(2 × 1). *Phys. Rev. B* **32**, 2326–2333 (1985).
157. Straub, D., Ley, L. & Himpsel, F. J. Conduction-band and surface-state critical points in Si: an inverse-photoemission study. *Phys. Rev. Lett.* **54**, 142–145 (1985).
158. Madelung, O. *Semiconductors: Group IV Elements and III–V Compounds* (Springer, 1991).
159. Petretto, G. et al. High-throughput density-functional perturbation theory phonons for inorganic materials. *Sci. data* **5**, 1–12 (2018).
160. Dolling, G. Inelastic scattering of neutrons in solids and liquids. *IAEA* **2**, 37 (1963).
161. Nilsson, G. & Nelin, G. Study of the homology between silicon and germanium by thermal-neutron spectrometry. *Phys. Rev. B* **6**, 3777–3786 (1972).
162. Schinke, C. et al. Uncertainty analysis for the coefficient of band-to-band absorption of crystalline silicon. *AIP Adv.* **5**, 067168 (2015).
163. Antonius, G., Poncé, S., Boulanger, P., Côté, M. & Gonze, X. Many-body effects on the zero-point renormalization of the band structure. *Phys. Rev. Lett.* **112**, 215501 (2014).
164. Zacharias, M. & Giustino, F. One-shot calculation of temperature-dependent optical spectra and phonon-induced band-gap renormalization. *Phys. Rev. B* **94**, 075125 (2016).
165. Zacharias, M. & Giustino, F. Theory of the special displacement method for electronic structure calculations at finite temperature. *Phys. Rev. Res.* **2**, 013357 (2020).
166. Zacharias, M. et al. Multiphonon diffuse scattering in solids from first principles: Application to layered crystals and two-dimensional materials. *Phys. Rev. B* **104**, 205109 (2021).
167. Brüesch, P. *Phonons, Theory and Experiments*, Vol. 1 (Springer, 1982).
168. Zacharias, M. & Kelires, P. C. Temperature dependence of the optical properties of silicon nanocrystals. *Phys. Rev. B* **101**, 245122 (2020).
169. Zacharias, M. & Kelires, P. C. Quantum confinement of electron–phonon coupling in graphene quantum dots. *J. Phys. Chem. Lett.* **12**, 9940–9946 (2021).
170. Ramírez, R., Herrero, C. P. & Hernández, E. R. Path-integral molecular dynamics simulation of diamond. *Phys. Rev. B* **73**, 245202 (2006).
171. Allen, P. B. & Heine, V. Theory of the temperature dependence of electronic band structures. *J. Phys. C: Solid State Phys.* **9**, 2305 (1976).
172. Popescu, V. & Zunger, A. Extracting E versus $\vec{k}$ effective band structure from supercell calculations on alloys and impurities. *Phys. Rev. B* **85**, 085201 (2012).
173. Zacharias, M. et al. Efficient first-principles methodology for the calculation of the all-phonon inelastic scattering in solids. *Phys. Rev. Lett.* **127**, 207401 (2021).
174. Perdew, J. P. & Zunger, A. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B* **23**, 5048–5079 (1981).
175. Ceperley, D. M. & Alder, B. J. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.* **45**, 566–569 (1980).
176. Giustino, F., Louie, S. G. & Cohen, M. L. Electron–phonon renormalization of the direct band gap of diamond. *Phys. Rev. Lett.* **105**, 265501 (2010).
177. Varshni, Y. P. Temperature dependence of the energy gap in semiconductors. *Physica* **34**, 149–154 (1967).
178. Karsai, F., Engel, M., Kresse, G. & Flage-Larsen, E. Electron-phonon coupling in semiconductors within the GW approximation. *N. J. Phys.* **20**, 123008 (2018).
179. Poncé, S. et al. Temperature dependence of the electronic structure of semiconductors and insulators. *J. Chem. Phys.* **143**, 102813 (2015).
180. Nery, J. P. & Allen, P. B. Influence of Fröhlich polaron coupling on renormalized electron bands in polar semiconductors: results for zinc-blende GaN. *Phys. Rev. B* **94**, 115135 (2016).
181. Kang, Y., Peelaers, H., Krishnaswamy, K. & Van de Walle, C. G. First-principles study of direct and indirect optical absorption in $BaSnO_3$. *Appl. Phys. Lett* **112**, 062106 (2018).
182. Lee, H. et al. *EPW Project* (2023) (last retrieved 02-06-2023) https://epw-code.org.
183. Engel, M. et al. Zero-point renormalization of the band gap of semiconductors and insulators using the projector augmented wave method. *Phys. Rev. B* **106**, 094316 (2022).
184. Bushick, K. & Kioupakis, E. Phonon-assisted Auger-Meitner recombination in silicon from first principles. *Phys. Rev. Lett.* **131**, 076902 (2023).
185. Zacharias, M., Volonakis, G., Giustino, F. & Even, J. Anharmonic lattice dynamics via the special displacement method. *Phys. Rev. B* **108**, 035155 (2023).
186. Sio, W. H. & Giustino, F. Unified ab initio description of fröhlich electron-phonon interactions in two-dimensional and three-dimensional materials. *Phys. Rev. B* **105**, 115414 (2022).
187. Poncé, S., Royo, M., Stengel, M., Marzari, N. & Gibertini, M. Long-range electrostatic contribution to electron–phonon couplings and mobilities of two-dimensional and bulk materials. *Phys. Rev. B* **107**, 155424 (2023).
188. Li, Z., Antonius, G., Wu, M., da Jornada, F. H. & Louie, S. G. Electron–phonon coupling from ab initio linear-response theory within the *GW* method: correlation-enhanced interactions and superconductivity in $Ba_{1-x}K_xBiO_3$. *Phys. Rev. Lett.* **122**, 186402 (2019).

189. Gonze, X. et al. The Abinit project: Impact, environment and recent developments. *Comput. Phys. Commun.* **248**, 107042 (2020).
190. Kresse, G. & Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558–561 (1993).
191. Soler, J. M. et al. The SIESTA method for ab initio order-N materials simulation. *J. Phys.: Condens. Matter* **14**, 2745 (2002).
192. Enkovaara, J. et al. Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method. *J. Phys.: Condens. Matter* **22**, 253202 (2010).
193. Gamma, E., Helm, R., Johnson, R. & Vlissides, J. *Design Patterns, Elements of Reusable Object-Oriented Software* (Addison-Wesley, 1995).
194. Ismail-Beigi, S. & Arias, T. A. New algebraic formulation of density functional calculation. *Comput. Phys. Commun.* **128**, 1–45 (2000).
195. Pizzi, G. et al. Wannier90 as a community code: new features and applications. *J. Phys.: Condens. Matter.* **32**, 165902 (2019).
196. Vitale, V. et al. Automated high-throughput wannierisation. *npj Comput. Mater.* **6**, 66 (2020).
197. Qiao, J., Pizzi, G. & Marzari, N. Projectability disentanglement for accurate and automated electronic-structure Hamiltonians. Preprint athttps://doi.org/10.48550/arXiv.2303.07877 (2023).
198. Agapito, L. A., Ferretti, A., Calzolari, A., Curtarolo, S. & Buongiorno Nardelli, M. Effective and accurate representation of extended Bloch states on finite Hilbert spaces. *Phys. Rev. B* **88**, 165127 (2013).
199. Damle, A., Lin, L. & Ying, L. Compressed representation of Kohn–Sham orbitals via selected columns of the density matrix. *J. Chem. Theory Comput.* **11**, 1463–1469 (2015).
200. Damle, A. & Lin, L. Disentanglement via entanglement: a unified method for wannier localization. *Multiscale Model. Simul.* **16**, 1392–1410 (2018).
201. Yates, J. R., Wang, X., Vanderbilt, D. & Souza, I. Spectral and Fermi surface properties from Wannier interpolation. *Phys. Rev. B* **75**, 195121 (2007).
202. Giannozzi, P., de Gironcoli, S., Pavone, P. & Baroni, S. Ab initio calculation of phonon dispersions in semiconductors. *Phys. Rev. B* **43**, 7231–7242 (1991).
203. Vogl, P. Microscopic theory of electron–phonon interaction in insulators or semiconductors. *Phys. Rev. B* **13**, 694–704 (1976).
204. Fröhlich, H. Electrons in lattice fields. *Adv. Phys.* **3**, 325–361 (1954).
205. Verdi, C. & Giustino, F. Fröhlich electron–phonon vertex from first principles. *Phys. Rev. Lett.* **115**, 176401 (2015).
206. Sjakste, J., Vast, N., Calandra, M. & Mauri, F. Wannier interpolation of the electron-phonon matrix elements in polar semiconductors: polar-optical coupling in GaAs. *Phys. Rev. B* **92**, 054307 (2015).
207. Marzari, N. & Vanderbilt, D. Maximally localized generalized Wannier functions for composite energy bands. *Phys. Rev. B* **56**, 12847–12865 (1997).
208. Souza, I., Marzari, N. & Vanderbilt, D. Maximally localized Wannier functions for entangled energy bands. *Phys. Rev. B* **65**, 035109 (2001).
209. Jhalani, V. A., Zhou, J.-J., Park, J., Dreyer, C. E. & Bernardi, M. Piezoelectric electron–phonon interaction from ab initio dynamical quadrupoles: Impact on charge transport in wurtzite GaN. *Phys. Rev. Lett.* **125**, 136602 (2020).
210. Park, J., Zhou, J.-J., Jhalani, V. A., Dreyer, C. E. & Bernardi, M. Long-range quadrupole electron–phonon interaction from first principles. *Phys. Rev. B* **102**, 125203 (2020).
211. Royo, M. & Stengel, M. First-principles theory of spatial dispersion: dynamical quadrupoles and flexoelectricity. *Phys. Rev. X* **9**, 021050 (2019).
212. Gonze, X. et al. Recent developments in the ABINIT software package. *Comput. Phys. Commun.* **205**, 106–131 (2016).
213. Poncé, S., Royo, M., Gibertini, M., Marzari, N. & Stengel, M. Accurate prediction of Hall mobilities in two-dimensional materials through gauge-covariant quadrupolar contributions. *Phys. Rev. Lett.* **130**, 166301 (2023).
214. Sohier, T., Calandra, M. & Mauri, F. Two-dimensional Fröhlich interaction in transition-metal dichalcogenide monolayers: theoretical modeling and first-principles calculations. *Phys. Rev. B* **94**, 085415 (2016).
215. Deng, T. et al. Ab initio dipolar electron–phonon interactions in two-dimensional materials. *Phys. Rev. B* **103**, 075410 (2021).
216. Sio, W. H. & Giustino, F. Polarons in two-dimensional atomic crystals. *Nat. Phys.* **19**, 629 (2023).
217. Verdi, C. & Giustino, F. Fröhlich electron–phonon vertex from first principles. *Phys. Rev. Lett.* **115**, 176401 (2015).
218. Towns, J. et al. XSEDE: accelerating scientific discovery. *Comput. Sci. Eng.* **16**, 62–74 (2014).
219. Ottaviani, G., Reggiani, L., Canali, C., Nava, F. & Alberigi-Quaranta, A. Hole drift velocity in silicon. *Phys. Rev. B* **12**, 3318–3329 (1975).
220. Kawamura, M. Fermisurfer: Fermi-surface viewer providing multiple representation schemes. *Comput. Phys. Commun.* **239**, 197–203 (2019).
221. Spicer, W. E. & Eden, R. C. Photoemission investigation of the band structure of semiconductors. In *Proc. Ninth International Conference on the Physics of Semiconductors*, Vol. 1 (ed. Ryvkin, S. M.) p. 65 (Nauka Publishers, Moscow, 1968).
222. Zucca, R. R. L. & Shen, Y. R. Wavelength-modulation spectra of some semiconductors. *Phys. Rev. B* **1**, 2668–2676 (1970).
223. Aspnes, D. E. & Studna, A. A. Dielectric functions and optical parameters of Si, Ge, GaP, GaAs, GaSb, InP, InAs, and InSb from 1.5 to 6.0 eV. *Phys. Rev. B* **27**, 985–1009 (1983).
224. Tiedje, T., Yablonovitch, E., Cody, G. & Brooks, B. Limiting efficiency of silicon solar cells. *IEEE Trans. Electron Devices* **31**, 711–716 (1984).
225. Green, M. A. & Keevers, M. J. Optical properties of intrinsic silicon at 300 K. *Prog. Photovolt.* **3**, 189–192 (1995).
226. Kim, H. J. et al. Physical properties of transparent perovskite oxides (Ba,La)SnO_3 with high electrical mobility at room temperature. *Phys. Rev. B* **86**, 165205 (2012).
227. Lee, H. et al. EPW v6-alpha (last accessed 5 November 2023). https://docs.epw-code.org/doc/DownloadAndInstall.html (2023).

ACKNOWLEDGEMENTS

This research is supported by: the Computational Materials Sciences Program funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, under Award No. DE-SC0020129 (project coordination, scale-up, polaron module, transport module, optics module, special displacement module); the National Science Foundation, Office of Advanced Cyberinfrastructure and Division of Materials Research under Grants Nos. 2103991 and 2035518 (superconductivity module, interoperability); the NSF Characteristic Science Applications for the Leadership Class Computing Facility program under Grant No. 2139536 (preparation for LCCF); the Fond National de la Recherche Scientifique of Belgium (F.R.S.-FNRS) and the European Union's Horizon 2020 research and innovation program under grant agreements No. 881603-Graphene Core3 (transport module); the NSF DMREF award 2119555 (quasi-degenerate perturbation theory module). This research used resources of the National Energy Research Scientific Computing Center and the Argonne Leadership Computing Facility, which are DOE Office of Science User Facilities supported by the Office of Science of the U.S. Department of Energy, under Contracts No. DE-AC02-05CH11231 and DE-AC02-06CH11357, respectively. The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing access to Frontera, Lonestar6, and Texascale Days, which have contributed to the research results reported within this paper (http://www.tacc.utexas.edu); the Extreme Science and Engineering Discovery Environment (XSEDE)[218] which is supported by National Science Foundation grant number ACI-1548562, and in particular Expanse at the San Diego Supercomputer Center through allocation TG-DMR180071. S.P. acknowledges computational resources provided by the PRACE award granting access to Discoverer in SofiaTech, Bulgaria (OptoSpin project id. 2020225411), and by the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the FRS-FNRS under Grant No. 2.5020.11 and by the Walloon Region, as well as computational resources awarded on the Belgian share of the EuroHPC LUMI supercomputer. K.B. acknowledges the support of the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Department of Energy Computational Science Graduate Fellowship under Award Number DE-SC0020347. The authors wish to thank Zhenbang Dai, Nikolaus Kandolf, Viet-Anh Ha, and Amanda Wang for their contributions to the EPW project that are not discussed in this manuscript; John Cazes and Hang Liu at TACC for their support with the Characteristic Science Applications project, Paolo Giannozzi for his support with `Quantum ESPRESSO`; and Stefano Baroni for fruitful discussions.

AUTHOR CONTRIBUTIONS

H.L. coordinated the development of the code, developed hybrid parallelism, and performed scaling tests. S.P. co-developed the transport module and maintained the test farm. K.B. co-developed the indirect absorption module. S.H. developed the full-band anisotropic Eliashberg module. J.L.B. tested and expanded the polaron module. J.L. developed the ionized-impurity scattering module. C.L. co-developed the polaron module. J.M.L. developed the Wannier-function perturbation theory module. F.M. co-developed the transport module. H.M. developed the intermediate representation for the superconducting module. H.P. tested and expanded the anisotropic Eliashberg module. W.S. developed the initial version of the polaron module. S.T. developed the quasi-degenerate perturbation theory module. M.Z. developed the special displacement module. X.Z. co-developed the indirect absorption module. N.B. supervised the development of the transport module. E.K. supervised the development of the indirect absorption module. E.R.M. supervised the development of the superconductivity module. F.G. supervised the project. All authors contributed to the writing of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Feliciano Giustino.

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

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