

The calculation of the infrared and terahertz spectrum of sodium peroxodisulfate

John Kendrick* and Andrew Burnett*

Department of Chemistry, University of Leeds, Leeds, LS2 9JT, UK

E-mail: J.Kendrick@leeds.ac.uk; A.D.Burnett@leeds.ac.uk

Abstract

A number of DFT programs with various combinations of pseudo-potentials and van der Waals' dispersive corrections have been used to optimize the structure of sodium peroxodisulfate, $\text{Na}_2(\text{SO}_4)_2$, and to calculate the infra-red and terahertz absorption spectrum of the powdered crystal. Comparison of the results from the different methods highlights the problems of calculating the absorption spectrum reliably. In particular the low frequency phonon modes are especially sensitive to the choice of grids to represent the wavefunction or the charge distribution, k-point integration grid and the energy cutoff. A comparison is made between the Maxwell-Garnett (MG) and Bruggeman effective medium methods used to account for the effect of crystal shape on the predicted spectrum. Possible scattering of light by air inclusions in the sample and by larger particles of $\text{Na}_2(\text{SO}_4)_2$ is also considered using the Mie method. The results of the calculations are compared with experimental measurements of the transmission and attenuated total reflection spectra.

1 Introduction

Infrared and Terahertz (THz) spectroscopies are incredibly powerful analytical techniques with many applications across the physical and life sciences. While the origin of many of the spectral features in an infrared spectrum can be easily interpreted with a knowledge about characteristic vibrational frequencies of functional groups, to identify the origin of all peaks (particularly below 1000 cm^{-1}) and understand the subtleties in peak shape and position theoretical support is essential. There are now a number of density functional theory (DFT) based packages designed for both small molecules^{1,2} and periodic solids³⁻⁷ that are capable of calculating vibrational frequencies and infrared intensities that can be used to interpret complex experimental spectra. There are also post-processing tools such as PDielec⁸ which take into account effective medium approximations,⁹ the attenuated total reflection (ATR) effect¹⁰ and Mie scattering¹¹ to aid in the interpretation of complex experimental spectra.

While these packages are now readily accessible, the calculation of spectra of complex systems, that correlate well with experiment can still be tricky, particularly at frequencies below 200 cm^{-1} .¹²⁻¹⁴ Choice of basis set or pseudo-potential,^{15,16} density functional,¹⁷ DFT package¹⁶ and convergence criteria can all have a dramatic affect on any calculated spectral parameter. This accessibility has a downside, and can lead to unsuitable calculations being compared to experiment, for instance it is still commonplace to compare single molecule DFT calculations to THz measurements of crystalline material¹⁸⁻²³ which has often been shown to be unsuitable.²⁴

In this paper we will discuss the potential pitfalls in the calculation of the infrared spectra of crystalline materials using a range of solid state DFT packages. We will look at the reliability of such calculations with respect to method, basis set and pseudo-potential. We will also compare a number of van der Waals' dispersive corrections which have been shown to be

particularly important in the calculation of THz spectra.^{25,26}

The crystal structure of the material, sodium peroxodisulfate, is available from the Crystallographic Open Database²⁷ with code number 2208366 and was determined from single crystal X-ray measurements at 150 K.²⁸ Sodium peroxodisulfate forms an interesting monoclinic, $P\bar{1}$, crystal structure with half a formula unit in the asymmetric cell and one formula unit in the full cell. Each sodium ion is in a distorted octahedral environment of oxygen atoms. The two SO_4 groups are linked together by a bridging O-O bond with an experimental bond length of 1.479 Å, as can be seen in Figure 1. This is an unusual functional group which will test the suitability of any oxygen basis set or pseudo-potential. Sulphur can also be a problematic element in THz spectroscopic calculations²⁹ again providing a challenge for the choice of basis set or pseudo-potential. Finally, the sodium ions are bound in the lattice by ionic bonding, causing considerable charge separation which means it is likely that the infrared spectrum of the material is influenced by the crystal shape. Previous work^{8,10,30} has shown that for ionic materials it is necessary to take into account the interaction between the crystal phonon modes and the electric field of the infrared radiation as large shifts in the absorption peaks can occur.

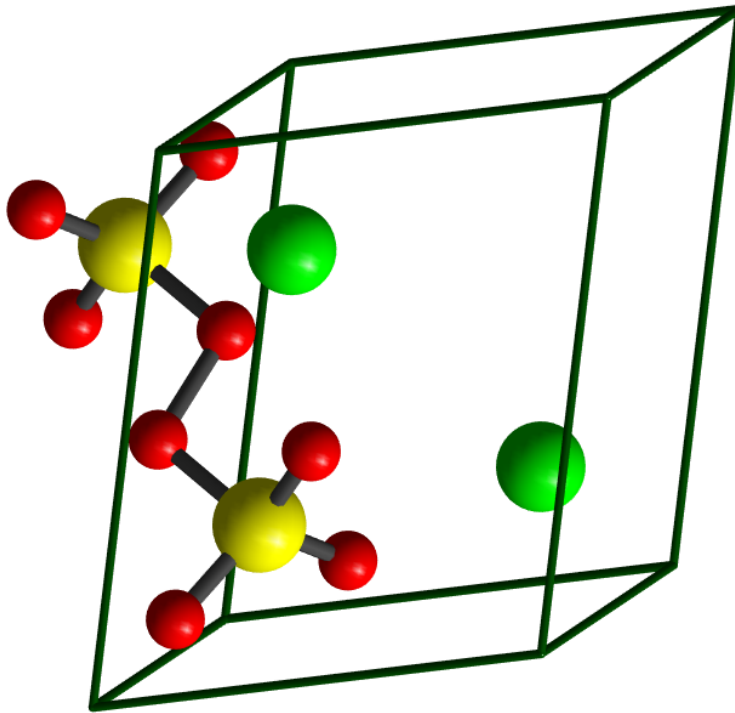


Figure 1: Unit cell of Sodium Peroxodisulfate - Na (green), S (yellow), O (red)

2 Methods

The following density functional theory (DFT) packages were used to calculate the optimized geometry and unit cell of $\text{Na}_2(\text{SO}_4)_2$; Abinit,⁵ CASTEP,³ Crystal14,⁶ VASP⁴ and Quantum Espresso⁷ (referred subsequently to as QE). For all, except Crystal14, pseudo-potentials were used to represent the core electrons. The Perdew-Burke-Ernzerhof (PBE) functional³¹ was used for all calculations. The pseudo-potentials and the associated packages are summarised in Table 1 along with the atomic configuration of the active electrons in the calculation.

Table 1: DFT packages, pseudo-potentials, basis sets and atomic electronic configurations

DFT Package	Pseudo-potential or basis set	Na	S	O
Abinit 8.2.3	FHI	$3s^1$	$3s^23p^4$	$2s^22p^4$
Abinit 8.2.3	ONCVSP	$2s^22p^63s^1$	$3s^23p^4$	$2s^22p^4$
CASTEP 17.2	NCP17	$2s^22p^63s^1$	$3s^23p^4$	$2s^22p^4$
QE 5.1	SSSP	$2s^22p^63s^1$	$3s^23p^4$	$2s^22p^4$
VASP 5.4.4	PAW	$2p^63s^1$	$3s^23p^4$	$2s^22p^4$
Crystal 14	TZVP	$1s^22s^22p^63s^1$	$1s^22s^22p^63s^23p^4$	$1s^22s^22p^4$
Crystal 14	DEF2	$1s^22s^22p^63s^1$	$1s^22s^22p^63s^23p^4$	$1s^22s^22p^4$

The ONCVSP pseudo-potentials³² and FHI pseudo-potentials³³ were obtained from the Abinit website.³⁴ The CASTEP norm-conserving pseudo-potentials were taken from the ‘on-the-fly’ pseudo potentials built using the NCP17 keyword as input to the ‘SPECIES.POT’ directive. The QE pseudo-potentials are Ultra Soft Pseudo-potentials (USPs), taken from the accurate set of Standard Solid State Pseudo-potentials (SSSP)^{16,35} and the VASP pseudo-potentials are Projector Augmented Wave (PAW) pseudo-potentials³⁶ distributed with VASP 5.4.4.

Crystal14 is an all electron code using atom centred contractions of Gaussian functions to represent the electronic wavefunction. Two basis sets were used; a triple zeta valence plus polarization (TZVP) basis³⁷ and a larger DEF2 basis set, based on the def2-TZVP molecular basis of Weigend and Ahlrichs.³⁸ For the DEF2 basis all f-functions were removed from the basis set and any Gaussian functions with an exponent less than 0.1 were removed from the sodium basis set in order to prevent linear dependencies in the calculation. Details of the number of uncontracted Gaussian functions and their contractions are given in the supplementary information (SI).

Alongside the choice of pseudo-potential there are also a number of possible choices for a suitable description of the dispersive interaction in solid state DFT codes. Table 2 shows

the dispersive correction used with each package;

Table 2: DFT packages and the dispersion correction

DFT Package	Dispersive correction	Designation
Abinit 8.2.3	Grimme D2	GD2
CASTEP 17.2	Grimme D3	GD3
CASTEP 17.2	Tkatchenko-Scheffler	TS
Crystal 14	Grimme D2	GD2
Quantum Espresso 5.1	Grimme D2	GD2
VASP 5.4.4	Grimme D3	GD3
VASP 5.4.4	Grimme D3 with Johnson and Becke damping	GD3-BJ
VASP 5.4.4	Tkatchenko-Scheffler	TS

The Grimme D2 method³⁹ adds a semiempirical correction to the energy of the form;

$$E_{disp} = -\frac{1}{2} \sum_{i \neq j} \frac{C_{6ij}}{r_{ij}^6} f(r_{ij}) \quad (1)$$

where the damping factor is;

$$f(r_{ij}) = \frac{S_6}{1 + e^{-d(r_{ij}/S_R R_{0ij} - 1)}} \quad (2)$$

The S_6 and S_R parameters are usually taken from the literature after fitting to reference calculations and are specific to the functional being used; for the PBE functional and the GD2 method values of 0.75 and 1.0 respectively are used. It has been common to vary their values so as to minimize the deviation between the calculated and experimental crystal structures. In some cases this is achieved by minimising the error in the calculated volume,^{40,41} in others the distortion of the cell is minimised.⁴²

Becke and Johnson⁴³ introduced an alternative damping function which leads to a constant contribution to the dispersion energy for spatially close pairs of atoms. The Grimme D3

method (GD3)⁴⁴ is a revision of the GD2 method which includes 3-body terms, allows for some geometry dependent information to be taken into account and includes higher multipole contributions. This method can also use the Becke and Johnson damping scheme (GD3-BJ). The method of Tkatchenko-Scheffler (TS)⁴⁵ is formally equivalent to the GD2 method but the parameters are functions of the charge density. For this dispersion correction the S_6 parameter is set to 1.0 and the parameter S_R is used for fitting (0.94 is the default value for the PBE functional).

2.1 k-point Integration

All calculations used a k-point integration scheme based on the Monkhorst-Pack method⁴⁶ with 7, 6 and 5 points in each reciprocal lattice direction respectively. Based on optimizations of the unit cell and the molecular geometry using VASP with an energy cutoff of 560 eV, increasing the k-point grid density by a factor of 2 showed that the energy of the optimized unit cell changed by less than 0.001 eV and the change in the calculated volumes of the optimized unit cells was less than 0.01 \AA^3 . Further details are given in the SI. Calculations of the phonon spectrum at the gamma point showed a difference of less than 0.05 cm^{-1} in the frequencies of the lowest nine non-zero modes. Although no explicit k-point convergence testing was performed with the other methods, it is expected that for this insulator, the choice of k-point sampling density suggested by VASP will be equally accurate for all methods.

2.2 Plane-wave Energy Cutoffs

Using the above k-point integration grid, optimizations of the atomic positions were performed at the, fixed, experimental unit-cell dimensions. From the atom-only optimized structures the internal pressure of the cells was calculated and the phonon modes calculated

at the gamma point. As reported in the SI, these calculations were carried out using a number of plane-wave energy cutoffs. Table 3 shows the chosen cutoff energy for each program and pseudo-potential along with the absolute difference between the calculated pressure, unit cell energy and frequencies for the chosen cutoff and the largest cutoff used. The root mean squared shifts in frequency was calculated by taking the frequencies of all the optical phonon modes of the calculation with the largest cutoff as a reference.

Table 3: Chosen energy cutoffs and associated absolute errors in energy, pressure and root mean squared shift (RMSS) in frequency

Program/Pseudo-potential	Cutoff Energy (eV)	Total Energy (eV)	Pressure (GPa)	Frequency RMSS (cm^{-1})
Abinit/FHI	1633	0.034	0.205	4.08
Abinit/ONCVSP	1633	0.009	0.012	1.84
CASTEP/NCP17	1300	0.015	0.024	1.75
QE/SSSP	680	0.045	0.253	0.23
VASP/PAW	600	0.050	0.369	1.14

The results for Abinit/FHI have the largest change in the mean squared frequencies. However, this was owing to the relatively poor translational invariance of the calculation; before the translational modes were projected out, the 3 lowest frequencies were far from zero and this caused some contamination of the 2 lowest optical modes. The wide range of cutoff energies reflects the type of pseudo-potentials being used by the packages. Those using USP or PAW pseudo-potentials are much softer than the norm-conserving pseudo-potentials and therefore need a smaller number of plane-waves in the basis set for a similar accuracy.

Using the plane-wave cutoffs shown in Table 3 and starting with the experimental crystal structure²⁷ the unit cell dimensions and the atomic positions were optimized, maintaining the space group symmetry.

2.3 Calculation of Infrared Absorption Spectra

The PDielec package⁸ was used to calculate the absorption spectrum from the Born charges, harmonic frequencies and phonon modes calculated by each of the packages. For the calculation of the effective permittivity PDielec assumes that the sodium peroxodisulfate is a powdered crystal dispersed in a supporting matrix and that the composite material has an effective complex permittivity, calculated using an effective medium theory such as Maxwell-Garnett or Bruggeman.⁹ The effective permittivity is calculated from the calculated permittivity of the sodium peroxodisulfate, the shape of the crystal and the permittivity of the supporting matrix. The effect of crystal shape on the absorption can be studied by comparing the effective medium theory spectra with that calculated at low concentrations using the Averaged Permittivity (AP) method,⁸ which shows absorption at the transverse optical (TO) phonon frequencies. In addition, using PDielec, it is possible to incorporate the effects of scattering during the absorption process using the Mie method for spherical particles.¹¹ This allows the program to describe the effect of air inclusions in the sample and to describe what effect larger particles of $\text{Na}_2(\text{SO}_4)_2$ may have on the spectrum.

The attenuated total reflection (ATR) technique is commonly used to record infrared spectra. An effective medium theory calculation of the effective permittivity for a high volume fraction of sodium peroxodisulfate embedded in air enables the calculation of the reflectance spectrum in an ATR configuration. The method used in PDielec to calculate the ATR reflectance is similar to that used by others.^{10,47} PDielec solves Fresnel's equations⁴⁸ for 45° incident radiation on a slab of non-absorbing, high permittivity material (such as diamond) supporting a layer of this effective medium.

2.4 Comparison of Calculated Infrared Absorption spectra

In order to compare the spectra calculated by the various packages, a spectrum was calculated using PDielec assuming a Maxwell-Garnett effective medium model of 10% by volume of small particles of sodium peroxodisulfate suspended in a PTFE matrix. Each spectrum was calculated with a frequency resolution of 0.1 cm^{-1} and the width of each absorption was taken to be 5 cm^{-1} . In order to calculate a normalized cross correlation between the calculated spectrum each spectrum, A , was normalized.

$$A_{norm}(i) = \frac{A(i) - \bar{A}}{\sqrt{n}\sigma(A)} \quad (3)$$

where n is the number of data points in the spectrum, \bar{A} is the mean value of the spectrum and $\sigma(A)$ is its standard deviation. The normalized cross-correlation coefficient can take values from -1 to +1, a value of zero indicates no correlation between the spectra.

The maximum cross-correlation between two spectra is calculated at a given frequency shift. The value of the maximum correlation coefficient and its 'lag' is used to calculate the similarity between the calculated spectra.

Using either the full cross-correlation matrix or the matrix of frequency shifts between all pairs of calculated spectra, a heat-map was calculated along with a clustering of the calculations according to their similarity. These calculations were performed using the gapmap package⁴⁹ in R.⁵⁰

3 Experimental

Sodium peroxodisulfate (99%) was bought from Sigma-Aldrich and ground using a Specamill stainless steel ball mill. The powder was mixed with a non-absorbing matrix material (for infrared measurements KBr was used and for THz measurements PTFE was used) and pressed using 7 t of force into pellets approximately 500 μm in thickness supported by a surrounding copper ring. A Nicolet iS5 FTIR was used for the transmission infrared measurements and 32 scans recorded for both background and sample at a 1 cm^{-1} frequency resolution.

For ATR infrared measurements a Bruker alpha platinum ATR instrument was used. A sample of sodium peroxodisulfate was placed onto the diamond ATR crystal, clamped in place, and an average of 32 scans with a spectral resolution of 2 cm^{-1} was recorded. THz spectra were recorded on a home-built THz time-domain spectrometer (THz-TDS) previously described elsewhere.⁵¹ In brief, spectral measurements were performed using a dry-air purged broadband THz-TDS using a mode-locked Ti:sapphire laser (Vitara, Coherent) which was used to produce a train of near-infrared pulses, each of duration 20 fs, centered at 800 nm at a repetition rate of 80 MHz. The beam was then focused onto a low-temperature-grown gallium arsenide (LT-GaAs) on quartz⁵¹ photoconductive switch with a large-area slot electrode design which was 200 μm wide and 4 mm long. The emitter was biased at 350 V using a 7 kHz modulation frequency with a 50% duty cycle to enable lock-in detection. The THz radiation emitted from the photoconductive switch was collected and collimated from the side of the emitter excited by the laser and focused onto the sample pellet by a set of off-axis parabolic mirrors. The THz radiation transmitted through the sample was then recollected and focused with a second pair of mirrors onto a second LT-GaAs-on-Quartz device used as a photoconductive detector. The current generated in the photoconductive switch was amplified using a transimpedance amplifier with a gain of $1 \times 10^8 \Omega$ with a time-delayed

probe beam (100 mW) split off from the original near-infrared laser pulse train used for detection. Spectra are an average of 60 scans recorded with a frequency resolution of 0.8 cm^{-1} . Low temperature measurements were performed by mounting the sample pellet, within a copper ring for good thermal contact, onto a coldfinger of a continuous-flow helium cryostat (MicrostatHe, Oxford Instruments) equipped with polymethylpentene (TPX) windows.

4 Results and Discussion

For comparison purposes, the experimental unit cell dimensions²⁸ are reported in Table 4 along with the length of the O-O bond in the crystal.

Table 4: Experimental unit cell dimensions of $\text{Na}_2(\text{SO}_4)_2$

a/Å	b/Å	c/Å	$\alpha/^\circ$	$\beta/^\circ$	$\gamma/^\circ$	Volume/Å ³	O-O/Å
4.780	5.575	6.091	101.871	103.337	97.418	151.88	1.479

4.1 Geometry optimization

Table 5 shows the percentage errors in the optimized unit cell parameters for the methods used which apply no dispersion correction. The calculated cell dimensions are provided in the SI. The calculated volumes from these optimizations are systematically larger than the experimental cell volume by more than 6%. The calculated O-O bond is also too large by more than 1.5%. This systematic error in volume is expected because the uncorrected DFT methods do not include any electron correlation, and this tends to increase the unit-cell volume. The experimental unit-cell dimensions were determined at 150K. For the purpose of comparison with calculation determination at a lower temperatures would be better, but based on the expansions coefficients of mirabilite ($\text{Na}_2\text{SO}_4(\text{D}_2\text{O})_{10}$)⁵² the unit-cell volume

of sodium peroxodisulfate would only be 0.36% smaller at 0K and each unit-cell dimension would be only 0.12% smaller.

Table 5: Calculated percentage errors in unit cell dimensions using no dispersion correction^a

Method	a	b	c	α	β	γ	Volume	RMFSD ^b	O-O
Abinit/FHI	3.2	4.7	-0.4	-1.3	0.0	2.7	7.3	2.6	1.5
Abinit/ONCVSP	3.0	3.9	1.1	-0.5	-0.6	1.2	8.3	2.1	1.6
CASTEP/NCP17	3.2	4.1	1.5	-0.6	-0.5	1.2	9.1	2.3	3.7
Crystal/TZVP	2.5	1.6	1.5	-0.5	0.0	-0.4	6.1	1.4	5.7
Crystal/DEF2	2.9	2.9	1.5	-0.9	-0.5	0.6	7.9	1.8	1.9
QE/SSSP	2.5	3.3	1.1	-0.6	-0.7	1.0	7.3	1.8	1.6
VASP/PAW	2.4	2.9	1.2	-0.5	-0.5	0.8	6.9	1.7	2.0

^a Negative numbers indicate that the calculated value is smaller than experiment

^b RMSFD is the root means squared fractional deviation of the optimised cell from the experimental and is given as a percentage

Table 6: Calculated percentage error in unit cell dimensions using dispersion corrections^a

Method	a	b	c	α	β	γ	Volume	RMSFD ^b	O-O
Abinit/FHI/GD2	0.1	0.4	-0.5	-1.0	-0.1	0.7	0.2	0.6	1.4
Abinit/ONCVSP/GD2	0.4	0.9	0.5	-0.6	-0.4	0.2	2.2	0.5	1.5
CASTEP/NCP17/GD3	0.1	0.7	1.7	-0.9	-0.8	-0.8	3.6	1.0	3.7
CASTEP/NCP17/TS	-0.5	-0.2	2.7	0.9	1.1	-1.1	1.2	1.3	3.3
CASTEP/NCP17/TS-v	-0.5	-0.9	2.0	0.8	1.2	-1.1	-0.1	1.2	3.3
Crystal/TZVP/GD2	-1.1	-3.0	3.5	-1.1	0.7	-3.1	0.3	2.4	5.8
Crystal/DEF2/GD2	-0.6	-0.5	2.1	-1.2	-0.6	-1.6	2.4	1.3	1.9
Crystal/DEF2/GD2-v	-1.7	-3.2	4.4	-0.6	1.1	-3.5	0.0	2.8	2.0
Crystal/DEF2/GD2-r	0.6	0.5	1.8	-1.1	-0.7	-1.0	4.1	1.0	1.9
QE/SSSP/GD2	-0.4	0.2	1.4	-0.9	-0.8	-0.9	2.2	0.9	1.6
VASP/PAW/GD2	-0.4	0.5	1.0	-0.8	-0.8	-0.4	2.0	0.7	2.1
VASP/PAW/GD3	-0.1	1.0	1.0	-0.9	-0.1	0.0	2.3	0.7	2.0
VASP/PAW/GD3-BJ	0.2	0.7	0.2	-0.7	-0.3	0.3	1.5	0.5	1.9
VASP/PAW/TS	-1.0	-0.1	1.2	0.4	0.4	-0.5	-0.2	0.7	1.8

^a Negative numbers indicate that the calculated value is smaller than experiment

^b RMSFD is the root means squared fractional deviation of the optimised cell from the experimental and is given as a percentage

The effects of the dispersion corrections on the calculated unit cell parameters are shown

in Table 6. In this Table the method designator includes the dispersion correction method used.

A few calculations were performed, using Crystal/DEF2/GD2, CASTEP/NCP17/GD3, and CASTEP/NCP17/TS where the S_6 parameter of the Grimme dispersion correction or the S_R parameter of the TS correction was optimized to improve the agreement between the experimental and calculated unit-cell dimensions. Details of the calculation of the optimized parameters can be found in the SI. In these cases the fact that optimized dispersion correction parameters have been used rather than the default values is indicated by '-v' or '-r' in the method label. A '-v' indicates the parameter was determined so as to reproduce the experimental volume, a '-r' indicates that the root mean squared fractional deviations (RMSFD) of the calculated unit-cell dimensions and angles from the experimental values was minimized.

In the case of Crystal/DEF2/GD2-v, the resulting value of S_6 (0.92) was larger than the default value of 0.75 and resulted in a distortion of the unit-cell, as evidenced by an RMSFD of 2.8%, which is greater than that of the non-dispersion corrected unit-cell (1.8%). Optimizing the parameter to minimize RMSFD, Crystal/DEF2/GD2-r, gave an S_6 value of 0.5 (RMSFD 1.0%), but the calculated cell volume is 4% larger than experiment. However, the distortion is lower than that of the non-dispersion corrected unit-cell.

Attempts were made to reproduce the experimental volume by optimizing the S_6 parameter for CASTEP/NCP17/GD3. The results are reported in the SI and indicate a sudden change in the packing of the cell around the value needed to give the experimental volume. The value of S_6 which minimized the RMSFD was the same as the default value (0.75), so only CASTEP/NCP17/GD3 results are reported.

For CASTEP/NCP17/TS the value of S_R which gave the experimental volume was 0.916, which is close to the value which minimizes RMSFD (0.90), so only the default

CASTEP/NCP17/TS ($S_R=0.94$) and CASTEP/NCP17/TS-v results are reported.

All methods which include a dispersion correction lead to a reduction in the calculated volume of the cell. Of those methods where no parameter optimization was performed, VASP/PAW/TS is closest to the experimental volume. VASP/PAW/GD2 and Abinit/ONCVSP/GD2 have the lowest RMSFD (0.5%), which is lower than that achieved by Crystal/DEF/GD2-r, even with minimization of RMSFD.

The process of optimizing the S_6 parameter by minimizing the error in the volume of the cell can lead to some unintended consequences. This is most clearly shown by the Crystal/DEF/GD2-v results. Although the volume calculated by this calculation agrees with the experiment, the percentage errors in the cell dimensions are much larger than the results of non-optimized case.

The O-O bond length is calculated to be larger than the experimental value for all calculations. The addition of dispersion corrections does not significantly change the value. The Crystal/TZVP and its dispersion corrected methods calculate the longest O-O bonds, indicating that there may be a problem with the oxygen basis set used for this calculation.

4.2 Translational Invariance

Accurate calculations of the phonon modes at the gamma point of the crystal should give three acoustic modes which have zero frequency reflecting the translational invariance of the crystal. If the calculation has not converged sufficiently in sampling the Brillouin zone or in basis set, or there is some underlying grid (e.g. FFT grid) being used within the calculation that is insufficiently fine, then these modes may have non-zero values. In some cases an imaginary value indicates that the cell is unstable with respect to atomic motion. No such instabilities were encountered for the calculations reported here. Any (small) imaginary

frequencies were as a result of losing translational invariance. In the case of CASTEP a FINE_GRID_SCALE parameter of 6 was chosen to minimise the deviations of the acoustic mode frequencies from zero as described in the SI. Table 7 shows the root mean squared error (RMSE) in frequencies of the acoustic modes at the gamma point, as calculated by each package, without imposing translational invariance. The squared error is averaged over all the methods used by each package. The error in the acoustic mode frequencies depended principally on the package used and was independent of the method. In addition the Table shows the root mean squared shift (RMSS) of the optical frequencies for each calculation as a result of imposing translational invariance on the dynamical matrix, by projecting out the translational degrees of freedom from the dynamical matrix using PDielec and recalculating the phonon frequencies.

Table 7: The root mean squared error (RMSE) in the acoustic mode frequencies and the root mean squared shift (RMSS) in the optical frequencies on projection of the translational degrees of freedom

Package	RMSE (cm^{-1})	RMSS (cm^{-1})
Abinit	24.0	0.1345
CASTEP	6.4	0.0003
Crystal	2.7	0.0045
QE	10.8	0.0010
VASP	1.1	0.0015

Abinit shows large deviations from zero in the unprojected acoustic mode frequencies and also shows a significant effect on the optical frequencies as a result of projecting out the translational modes from the dynamical matrix. The shift in frequency as a result of projection is largest for the lower frequency modes. CASTEP, Crystal and VASP have only small deviations from zero in the unprojected acoustic mode frequencies and there is only a small shift in the optical mode frequencies after projection. The QE results show a significant deviation from zero in the unprojected acoustic mode frequencies, but this does not lead to

significant shifts in the optical mode frequencies upon projection.

4.3 Calculated Frequencies and Intensities

Figure 2 shows a comparison of the non-dispersion corrected calculations of frequencies and intensities over the full frequency range. The spectrum of $\text{Na}_2(\text{SO}_4)_2$ falls into three distinct ranges; low (below 300 cm^{-1}), intermediate ($300\text{-}750\text{ cm}^{-1}$) and high (above 750 cm^{-1}) The low frequency region up to 300 cm^{-1} is complex as can be seen in Figure 3 but all calculations predict six absorptions of varying intensity and position in this region. The calculated frequencies and intensities for all calculations are given in the SI.

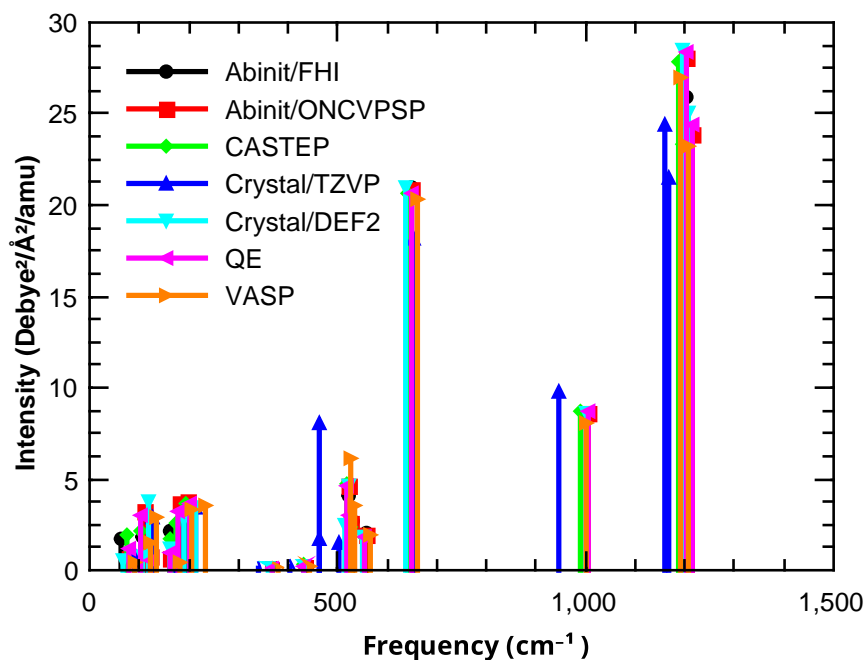


Figure 2: Non-Dispersion Corrected Frequencies and Intensities

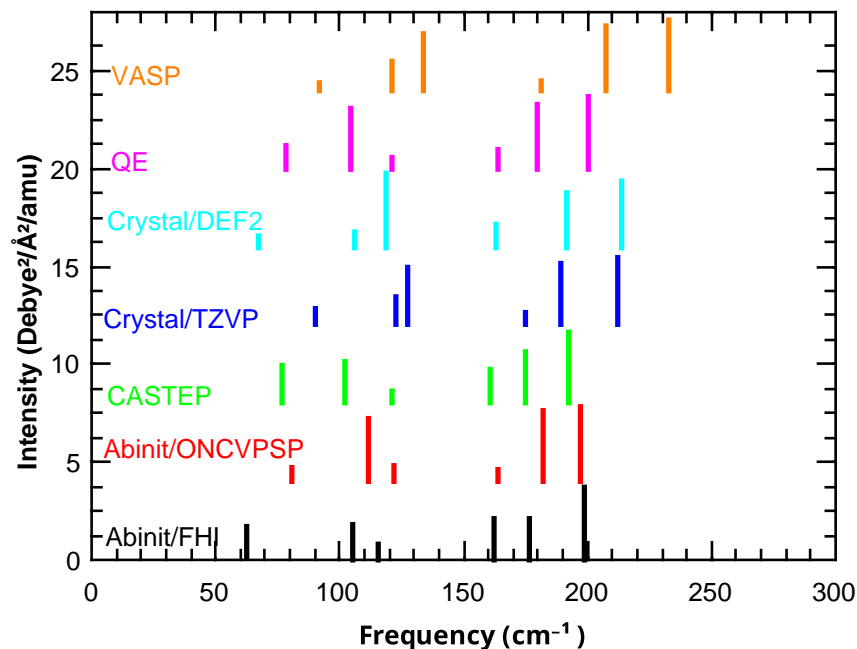


Figure 3: Non-Dispersion Corrected Frequencies and Intensities - Low Frequencies

4.4 Phonon Mode Analysis

The make-up of each phonon mode in terms of either the internal/external contributions or in terms of the contributions from particular groups of atoms can be determined from their percentage kinetic energy contribution to that mode. The approach adopted in PDielec for this analysis is described in previous work.¹³

The internal contributions can be regarded as molecular vibrations and the external contributions as whole molecule translatory or rotatory motion. There are two obvious ‘molecular’ groupings for this crystal. In one the S_2O_8 moiety can be treated as a single unit or it can be treated as two independent SO_4 units. The sodium atoms are treated as being able to contribute to external translatory modes. Figures 4 and 5 show the results of the analysis of the Crystal/DEF2 calculation using SO_4 and Na molecular groups. Analysis of the results

from all the other packages, pseudo-potentials and dispersion corrections gives very similar results to those presented here for Crystal/DEF2. Figure 4 shows the contribution of external (translatory and rotatory) and internal (vibrational) contributions. Figure 5 shows the break-down into 'molecular' components. The frequencies of the individual normal modes for this particular calculation are given in the supporting information, as are the related results where the S_2O_8 unit is treated as the molecular unit.

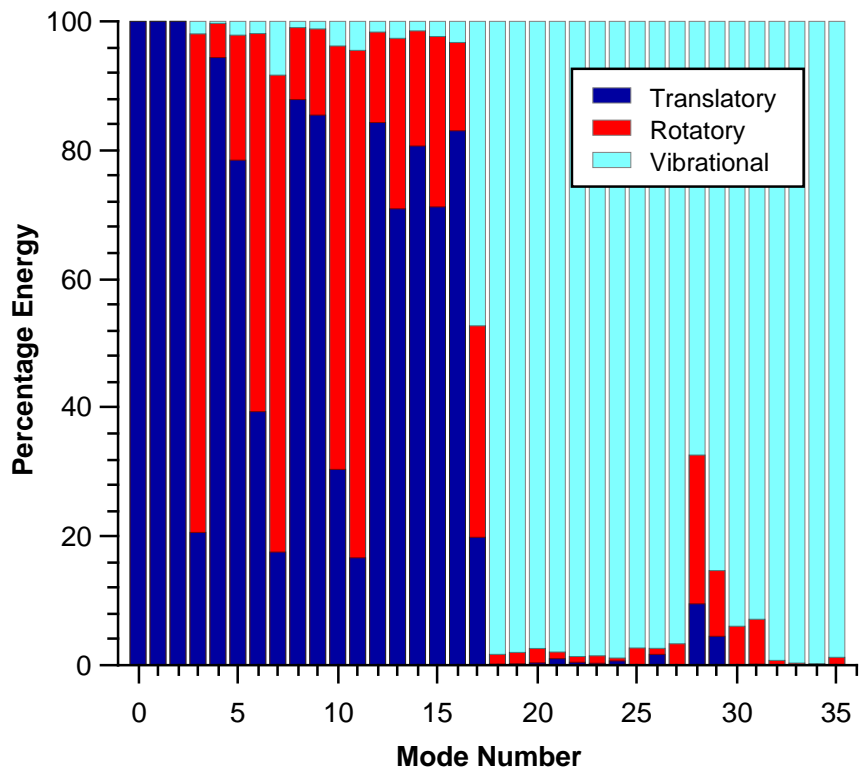


Figure 4: Mode Analysis of Crystal/DEF2 Phonon Modes, Internal and External Contributions

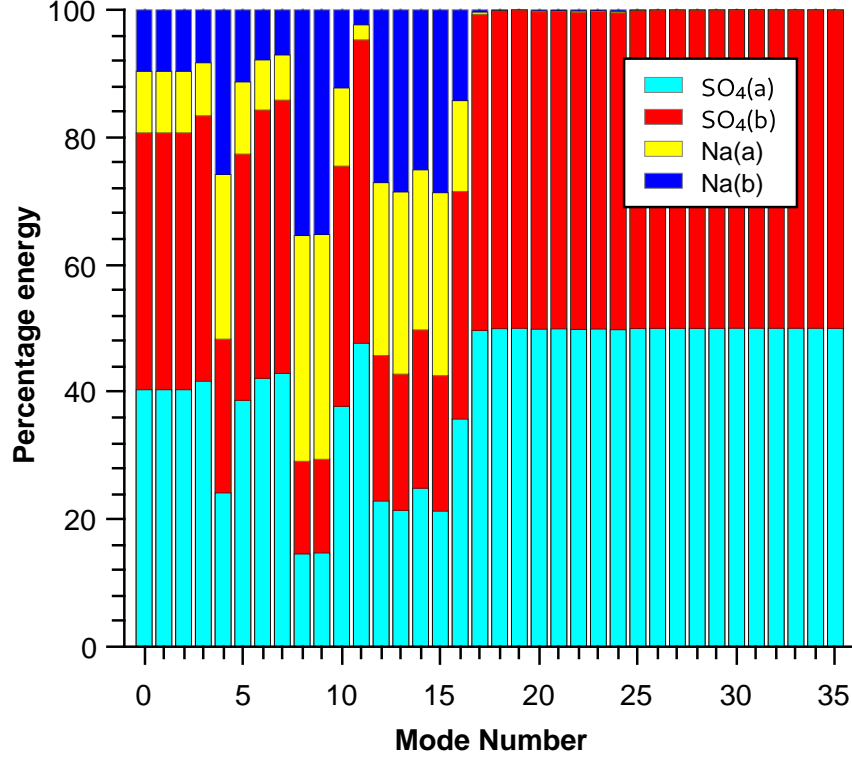


Figure 5: Analysis of Crystal/DEF2 Phonon Modes, Molecular Contributions

The Figures show the 3 acoustic modes (modes 0 to 2) to be purely translational modes, as expected. All modes up to and including mode 17 are external modes dominated by rotatory and translatory motion of the molecular groups. All 18 modes include contributions from the sodium atom translatory motion, as can be seen from Figure 5. Above 311.5 cm^{-1} (the frequency of mode 17) modes are all internal vibrations of the SO_4 group. Inspection of the Figures in the SI calculated assuming an S_2O_8 group show that the lowest optical mode (mode 3) has a large contribution from the vibration of the S_2O_8 group originating from movement of the O-O bond. Mode 28 which shows significant rotatory motion of the SO_4 groups is associated with stretching of the O-O bond.

4.5 Infrared Spectra determined using a Maxwell-Garnett Effective Medium Theory

Unless otherwise stated the infrared spectra were calculated using PDielec⁸ from the normal modes and Born charges using a Maxwell-Garnett effective medium theory model for 10% by volume of small crystallites embedded in a PTFE matrix support. The Lorentzian line width for each absorption peak in the calculation was taken to be 5 cm^{-1} .

4.5.1 No Dispersion Correction

A comparison of the full frequency range for calculations involving no dispersion correction can be seen in Figure 6. Because the spectrum is dominated by the high frequency range, Figures 7 to 9 show the same spectrum but over the high, intermediate and low frequency ranges respectively. If the calculations were fully converged in all aspects including basis set, k-point integration and grid representation of the charge and wavefunction, then it should be expected that the spectra should agree with one another. In the high frequency region (Figure 7) this is clearly not the case. Although there is some qualitative agreement there are clear differences between the calculations. The absorptions in the region from 900 to 1050 cm^{-1} are the bending modes of the SO_4 units. The Crystal/TZVP results are more than 40 cm^{-1} lower in frequency than any other calculation, including Crystal/DEF2. The Crystal/DEF2, QE/SSSP, Abinit/FHI and Abinit/ONCVSP are higher in frequency than other calculations and the VASP/PAW and CASTEP/NCP17 calculations predict absorption at around 1000 cm^{-1} .

The S-O stretch region near 1200 cm^{-1} is more complex but shows a similar pattern, with the Crystal/TZVP calculation showing absorption at significantly lower frequencies and the QE/SSSP, Abinit/FHI and Abinit/ONCVSP calculations showing absorption at slightly

higher frequencies. Crystal/TZVP was previously shown to predict a long O-O bond, indicating a problem with this particular basis set. The deficiency in the basis set is manifesting itself with a poor prediction of the O-S-O bending and S-O stretching frequencies.

In the intermediate frequency region (Figure 8) the pattern is not quite the same. All calculations except CASTEP/NCP17 and Crystal/DEF2 agree that there is an absorption around 680 cm^{-1} . The CASTEP and Crystal/DEF2 results are shifted to slightly lower frequency. In the region from 400 to 600 cm^{-1} the Crystal/TZVP calculations shows absorption peaks at 480 and 510 cm^{-1} which are lower in frequency than all the other calculations by about 50 cm^{-1} . In this frequency region all the plane-wave calculations are in agreement with each other. The Crystal/DEF2 results are similar to the plane-wave calculations but there are two distinct peaks at 520 and 530 cm^{-1} , instead of a single absorption around 530 cm^{-1} .

The low frequency regime shown in Figure 9 is harder to unravel. There appears to be general agreement that there is significant absorption just above 200 cm^{-1} which comes from two strong absorptions. There is little agreement as to where the lowest frequency absorption occurs, although all methods predict some absorption below 100 cm^{-1} . Crystal/DEF2 predicts the lowest frequency absorption just above 70 cm^{-1} . Between 100 and 200 cm^{-1} there are 3 peaks, the middle peak of which is the most intense and whose frequencies can shift by up to 20 cm^{-1} depending on the package being used.

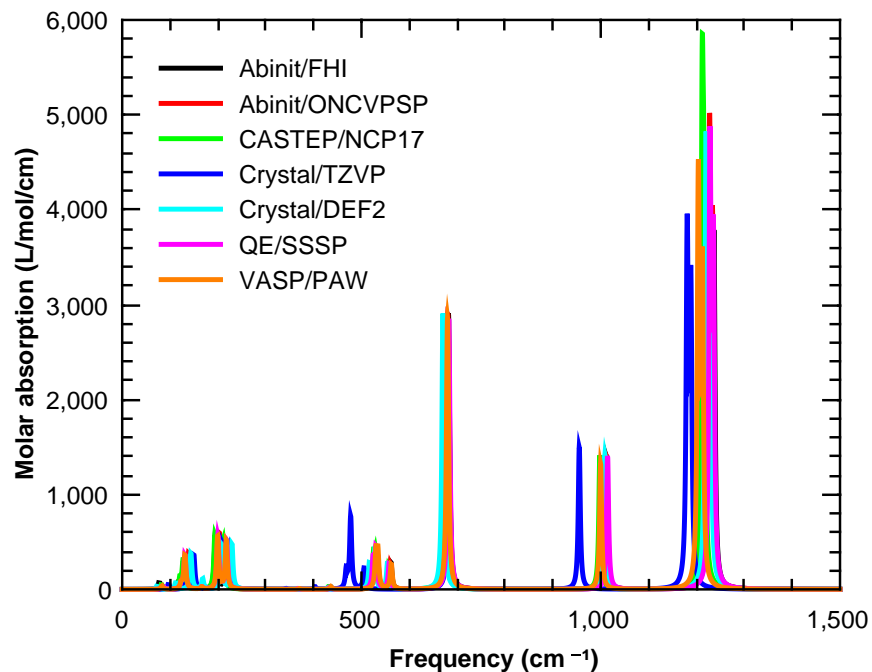


Figure 6: Non-Dispersion Corrected Spectra using Maxwell-Garnett Effective Medium Theory

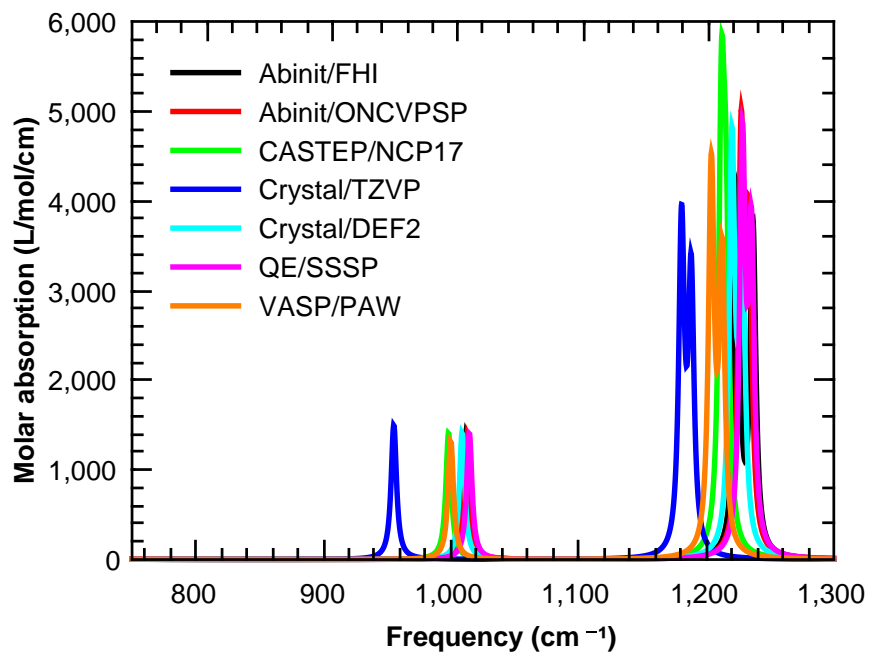


Figure 7: Non-Dispersion Corrected Spectra using Maxwell-Garnett Effective Medium Theory - High Frequency Range

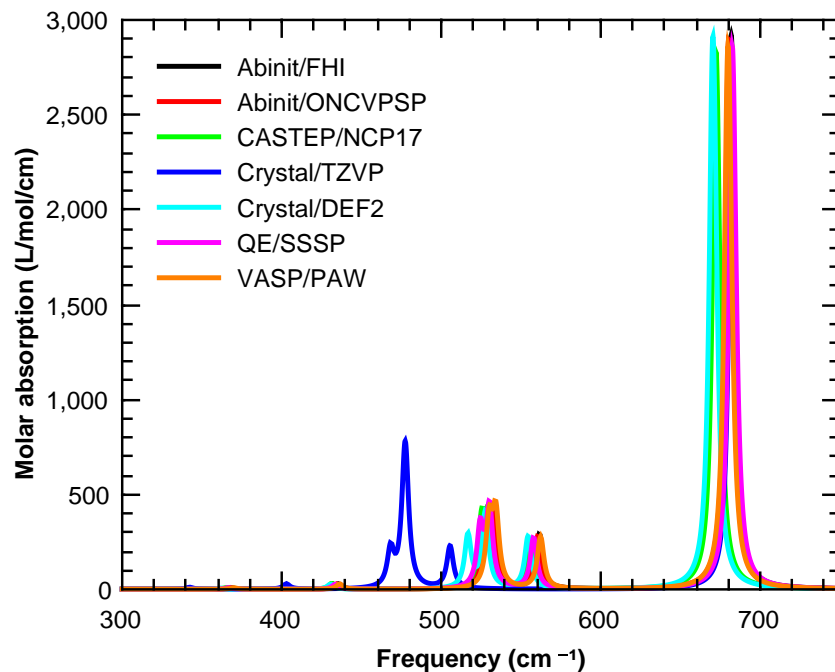


Figure 8: Non-Dispersion Corrected Spectra using Maxwell-Garnett Effective Medium Theory - Intermediate Frequency Range

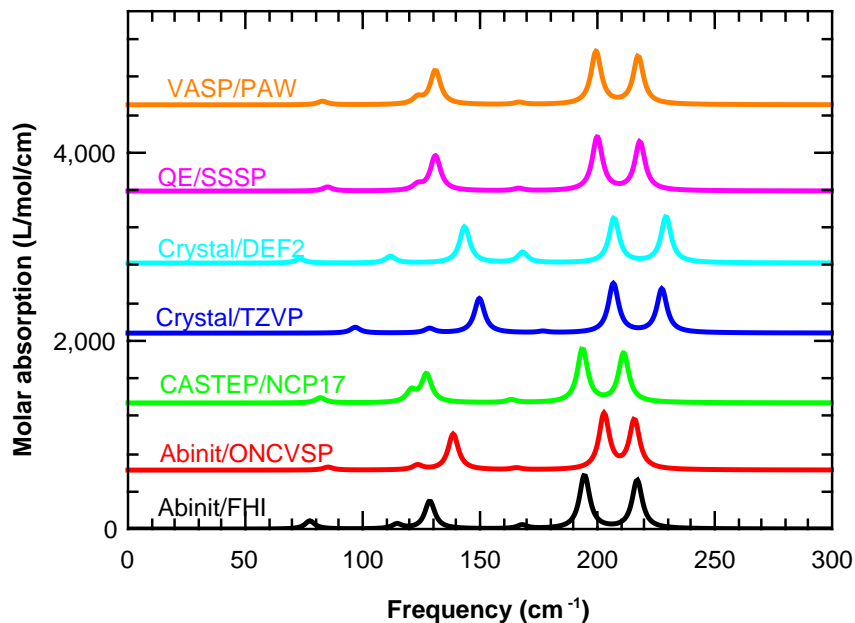


Figure 9: Non-Dispersion Corrected Spectra using Maxwell-Garnett Effective Medium Theory - Low Frequency Range

4.5.2 Dispersion Corrected Spectra

The effect of including a dispersion correction in the calculation of the unit cell and the phonon modes can be best seen by comparing the results of the VASP calculations with different dispersion corrections. Figure 10 shows the predicted spectra over the full frequency range. The intermediate and high frequency show small changes on including a dispersion correction and can be found in the SI. The uncorrected spectrum shows a single, slightly more intense absorption at about 1200 cm^{-1} , which comes from two transitions with similar frequencies. The dispersion corrected methods show two peaks at slightly higher frequencies. There are significant changes in the low frequency spectra, that can be more clearly seen in Figure 11. All VASP results show two low intensity modes at very low frequency with three more intense absorption peaks above 130 cm^{-1} . Whilst this pattern is the same for all calculations, the actual positions of the peaks vary for the differing methods. The VASP/PAW/GD2 results seem to show the highest shift in frequency from the non-dispersion corrected results with up to 40 cm^{-1} shift to higher frequencies in absorption. The GD3, BJD3 and TS dispersion correction methods predict absorption spectra in the low frequency regime which are very similar.

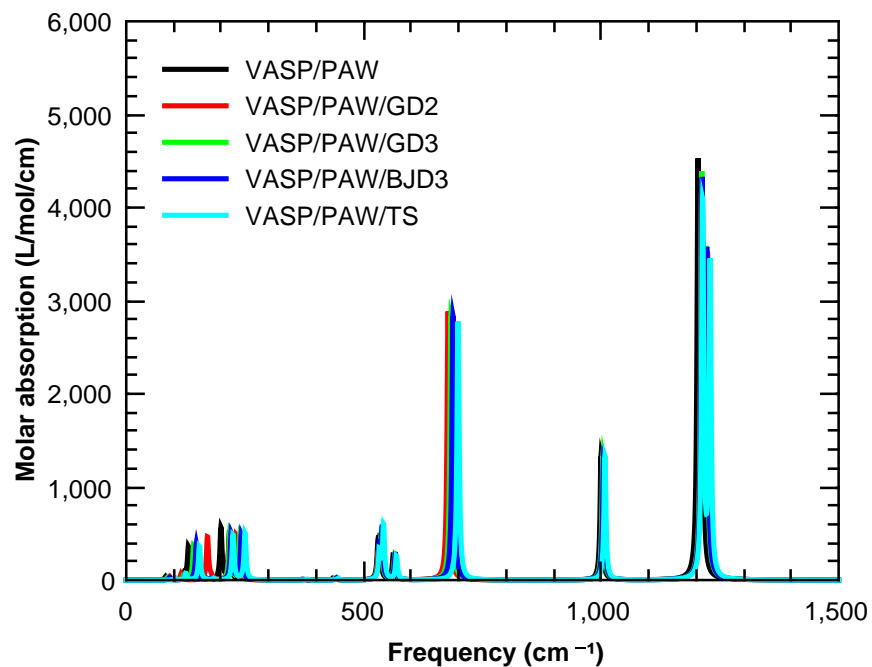


Figure 10: IR Spectra from VASP calculations - Full Frequency Range

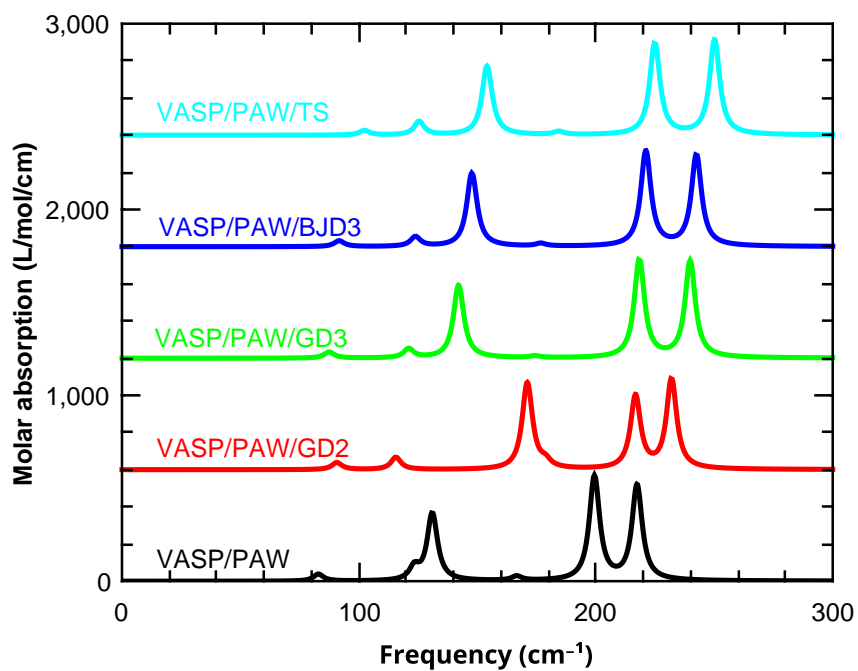


Figure 11: IR Spectra from VASP calculations - Low Frequency Range

The SI gives the full, high, intermediate and low frequency range calculated spectra for all of the methods used. In many respects the observations drawn from the VASP example shown above can be seen in the other methods. There tends to be a small shift to higher frequencies when dispersion corrections are included in the intermediate and high frequency ranges. The intensities are not affected. However in the low frequency range, although the qualitative pattern of absorption is similar, there are significant shifts in the frequency of absorption owing to the inclusion of a dispersion correction. The shift of absorption to higher frequency on the inclusion of dispersion is consistent with the decreased volume of the unit-cell, relative to the non-corrected volume. In the cases of CASTEP/NCP17/TS, Crystal/TZVP/GD2 and Crystal/DEF2/GD2 optimization of the S_6 parameter resulted in a smaller unit-cell and at least in the low frequency regime a shift to higher frequency (see SI).

4.6 Comparison of Calculated Spectra

The calculated spectra were compared with each other by calculating the normalized cross-correlation coefficient between each pair of spectra. This calculation also provides a 'lag' or frequency shift which maximizes the cross-correlation for each pair of spectra.

4.6.1 Full Frequency Range Comparison

Figure 12 shows the calculated cross-correlation coefficients for the complete frequency range. The cross-correlation matrix is symmetric and the results presented using a gap-map, where the methods have been clustered and reordered according to their similarity. The clustering is made clear by the dendrogram at the top of the heat-map. The heat-map itself shows the values of the cross-correlation coefficient as a colour map. Yellow is

used to describe the highest cross-correlation coefficient (1.0) and blue the lowest (0.5). VASP/PAW, VASP/PAW/GD3 and VASP/PAW/GD3-BJ are shown to be very similar. The none dispersion corrected methods tend to cluster at the top/left of the gap-map. Although the Abinit/FHI/GD2, Abinit/ONCVSP/GD2 and CASTEP/NCP17/GD3 methods also cluster in this region. For CASTEP/NCP17/TS-v the optimization of the S_6 coefficient seems to give results which are very similar to the unoptimized result. However, Crystal/DEF2/GD2-v stands out as being quite different to all the other calculations, which is probably a result of the large errors in the unit-cell dimensions (despite the agreement with the experimental volume).

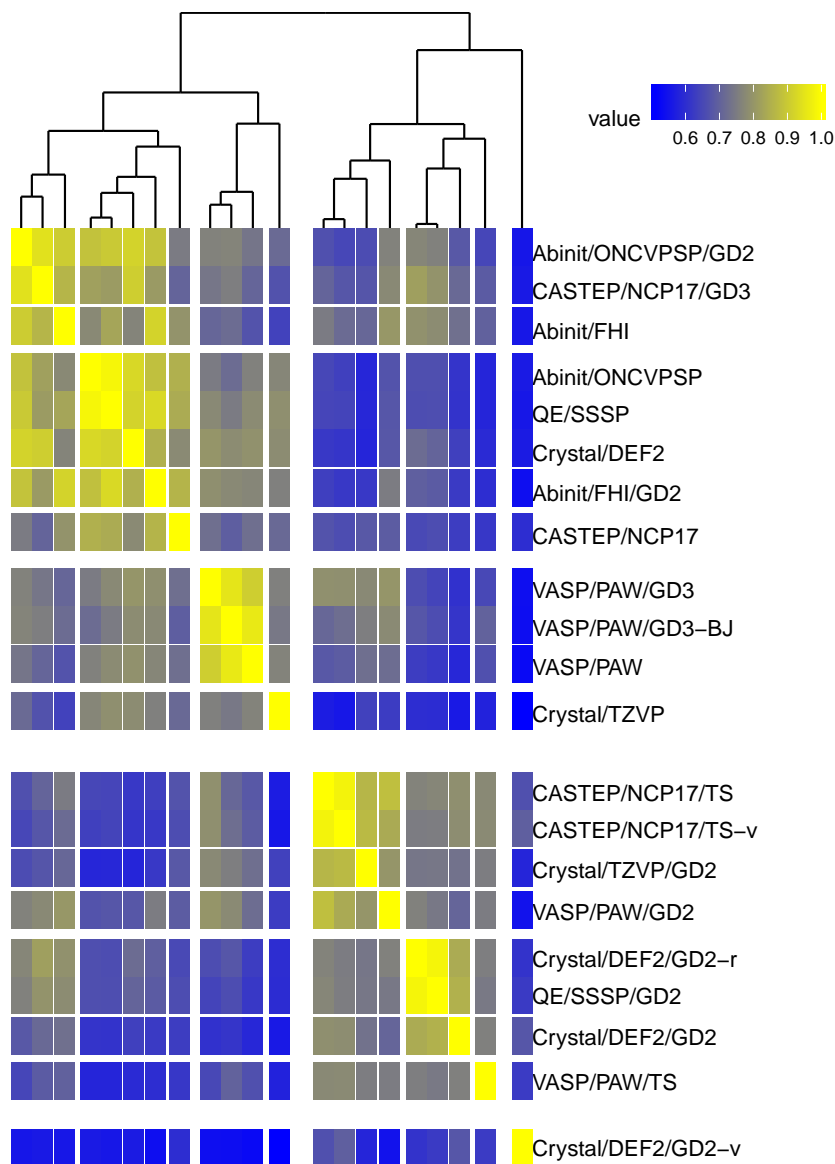


Figure 12: Cross-correlation heat-map of full frequency spectra after clustering

Figure 13 shows a gap-map created by using the lag frequency to calculate the similarity of each method. The lag frequencies in this plot have been calculated using the full frequency range of the spectra and they vary between -50 (blue) and $+50 \text{ cm}^{-1}$ (yellow). Surprisingly this method of clustering shows that the frequency shift which maximizes the correlation between spectra is strongly related to the program used to perform the calculation. It

does not seem to be related to the dispersion correction used. All the VASP calculations are clustered at the top-left of the diagram. Crystal/DEF2 and QE/SSSP calculations are clustered together, as are Abinit and CASTEP. For instance all of the Crystal/TZVP calculations are clustered together, probably reflecting the observed trend for Crystal/TZVP calculations to predict lower frequency absorption in the SO_4 bending region of the spectrum.

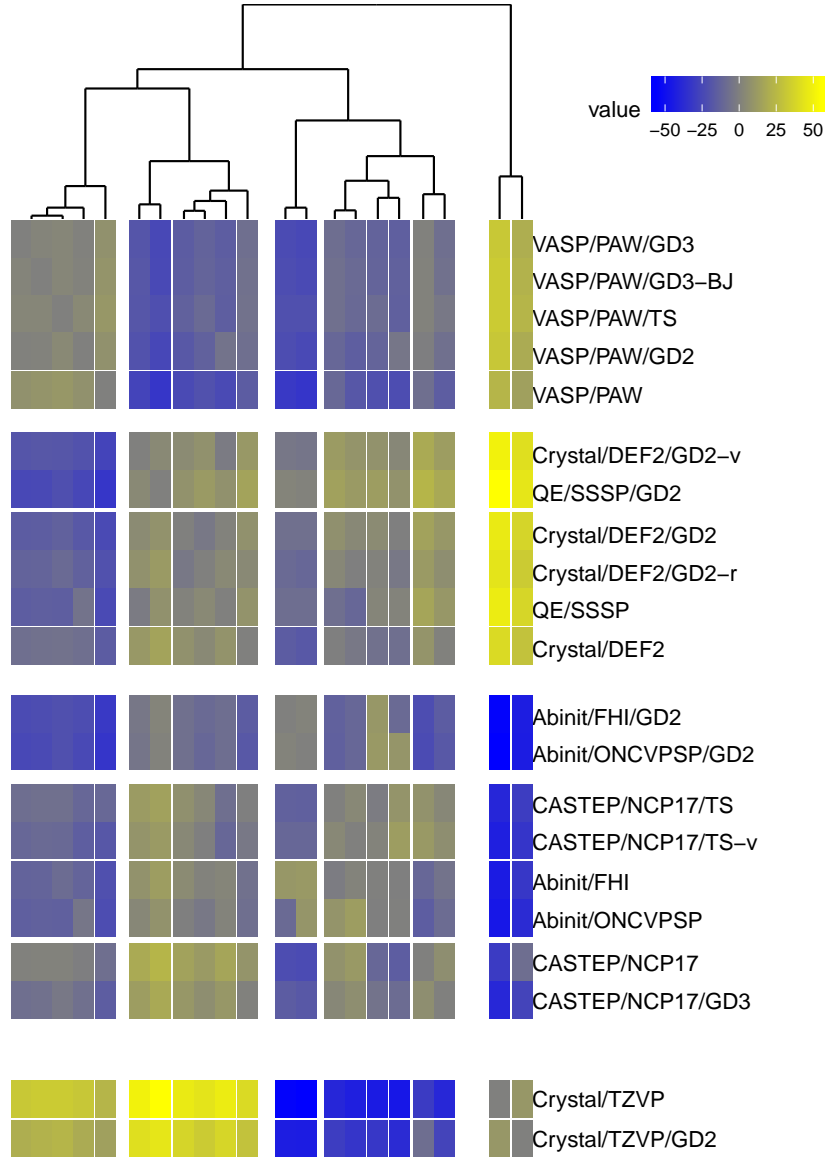


Figure 13: Frequency lag heat-map of full frequency spectra after clustering

4.6.2 Low Frequency Range Comparison

The absorption spectrum is dominated by the high frequency region, so it is interesting to see if calculating the cross-correlation function of the low frequency absorption gives similar results. The gap-maps of the correlation function and the lag are provided in the supporting information. The correlation coefficient shows a wider range of values, ranging from 0.4 to 1.0. Methods with the same dispersion correction seem to be clustered together, as are the non-corrected results. The TS and TS-v results are clustered together at the bottom/right of the table. The least similar groups are the GD2 and TS results. Inspection of the lag heat-map obtained from the low frequency spectrum shows no obvious pattern.

4.7 Comparison of Effective Medium Theories

The results presented so far calculate the effective permittivity of the composite material using a Maxwell-Garnett homogenization formula.⁹ Maxwell-Garnett is commonly used in a wide variety of circumstances, but it is not symmetrical with respect to the two components of the composite material. As a result the Bruggeman method⁹ is often preferred when the volume fractions of the two components are similar.

For comparison purposes an Averaged Permittivity (AP) effective medium theory⁸ with a low volume fraction of sodium peroxodisulfate is used to indicate the position and intensity of absorption from transverse optical phonons with no interaction with the field within the crystal.

Finally, there are occasions where scattering from the particles is important and to understand this calculations have been performed using a Mie methodology, which is relevant for low concentrations of spherical particles embedded in a non-absorbing medium.¹¹

The calculations presented here use the same parameters as above with only the effective medium method varied. For the purposes of comparing the MG, AP and Bruggeman effective medium methods, Figures 14 to 16 show the calculated molar absorption spectra from VASP/PAW/GD3-BJ calculations.

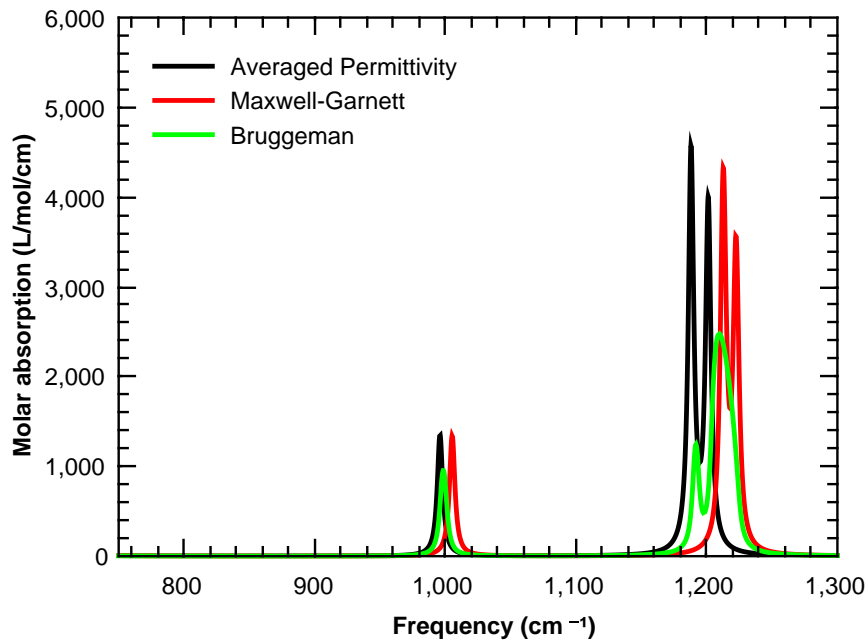


Figure 14: Comparison of Effective Medium Theories - VASP/PAW/GD3-BJ high frequency range

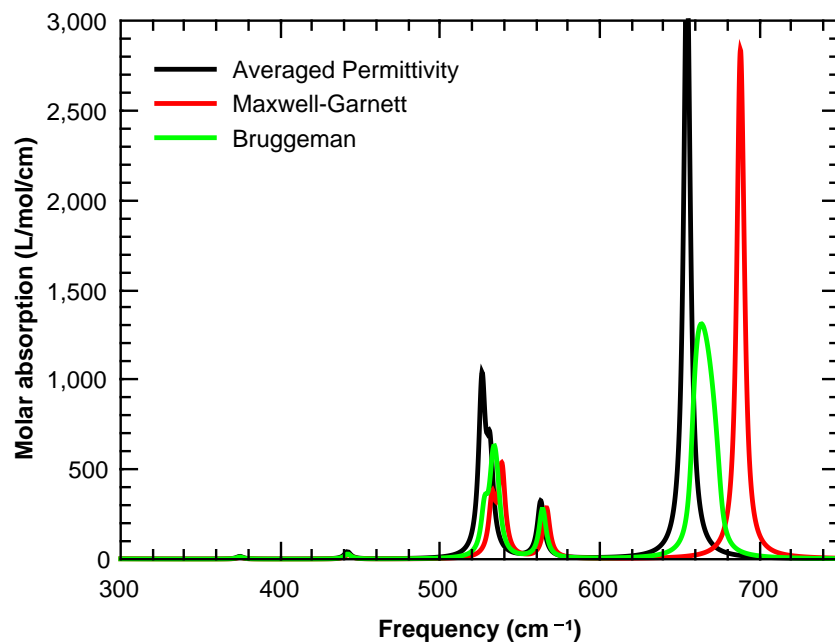


Figure 15: Comparison of Effective Medium Theories - VASP/PAW/GD3-BJ intermediate frequency range

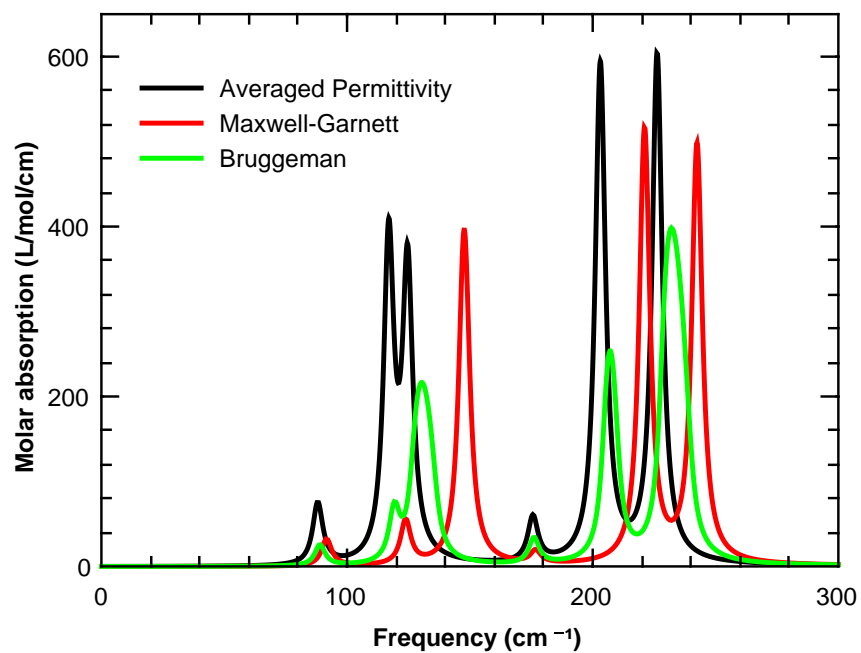


Figure 16: Comparison of Effective Medium Theories - VASP/PAW/GD3-BJ low frequency range

Figure 14 shows that in the high frequency range the MG method shifts the absorption to higher frequencies compared with the TO frequencies (shown by AP results), whilst the Bruggeman method produces much broader absorption peaks, with the peak maxima positioned between the AP and MG methods. The intermediate frequency range (Figure 15) is similar, though the single peak at 550 cm^{-1} is independent of the method of calculation. Just above 500 cm^{-1} the twin AP peaks are shifted to a slightly higher absorption maximum by the Bruggeman method, whilst the MG method show two peaks both at higher frequency than the AP calculation.

The low frequency spectrum shows the onset of absorption at between 80 and 90 cm^{-1} . All methods show six absorption peaks. Similarly to the higher frequency ranges, the MG method results in a shift of the absorption maxima to higher frequencies relative to the TO frequencies (shifts of up to 30 cm^{-1} are seen), whilst the Bruggeman method tends to show similar, but less marked trends, and much broader absorption peaks. As an example, similar trends can be seen for the Crystal/DEF2 calculations in the SI.

4.8 Spectra from Mie Scattering Calculations

When the wavelength of light is similar to or smaller than the particles being studied, scattering of light by the particles has to be considered. For spherical particles this can be described well using Mie scattering theory, as long as no multiple scattering events take place. In other words the particles must be very dilute. The spectra shown in Figures 17 and 18 were obtained from the VASP/PAW/GD3-BJ phonon calculations using the PDielec package. A 10% volume fraction of spheres in PTFE was used with a Lorentzian line width of 5 cm^{-1} . The smallest particle size ($0.1\text{ }\mu\text{m}$) results for Mie scattering coincide with those of the Maxwell-Garnett method. As the particle size increases the higher frequencies are more affected and the absorption broadens with very little change seen at lower frequencies.

Figures showing the full and low frequency ranges can be seen in the SI.

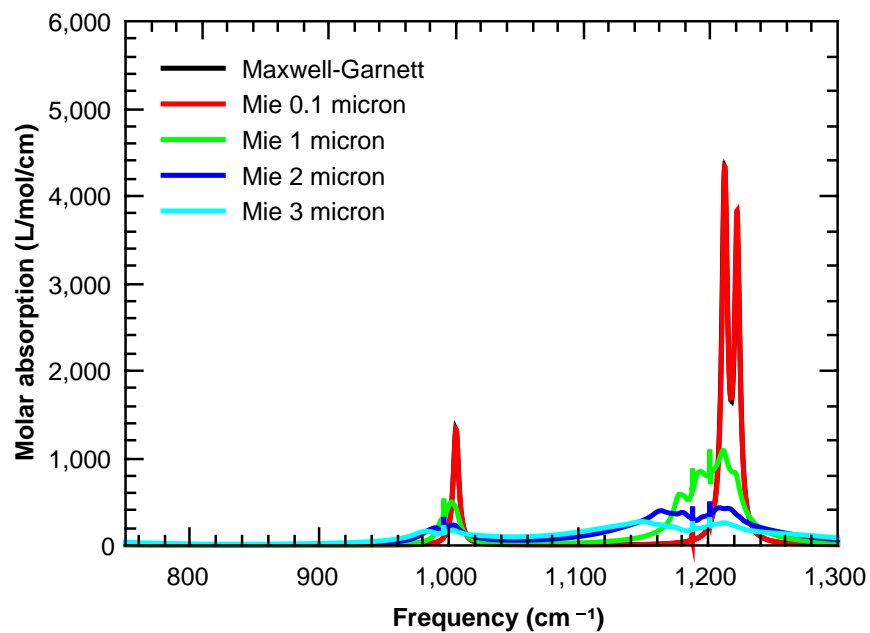


Figure 17: Mie scattering calculations - high frequency range

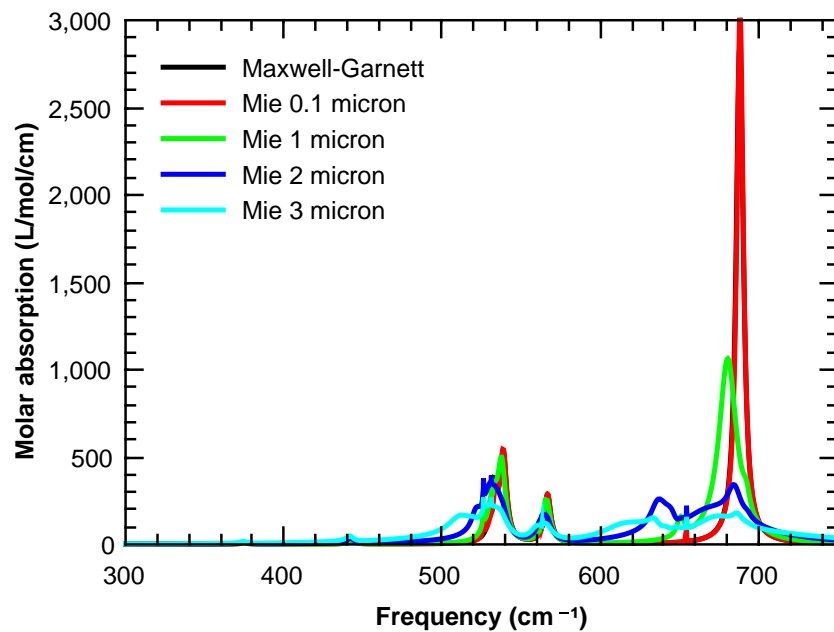


Figure 18: Mie scattering calculations - intermediate frequency range

This method can also be used to understand the effect of scattering by air voids or bubbles that are unavoidable in pelletised samples and have shown to contribute to the background spectral response at low frequencies.^{53,54}

4.9 Comparison with experiment

In this section we compare the calculated spectra with experimentally measured IR and THz spectra of $\text{Na}_2(\text{SO}_4)_2$. In order to improve the correlation between calculation and experiment and identify any systematic error in the DFT calculations we have also explored re-scaling of the calculated spectra. Such re-scaling is common in molecular calculations where the systematic errors in the calculated frequency of a particular method are corrected by a scale factor.⁵⁵ The frequency re-scaling can be expressed as;

$$f_{new} = lag + scale f_{calc} \quad (4)$$

4.9.1 ATR Spectra

Figure 19 shows a comparison of the experimental ATR spectra with that calculated by VASP/PAW/GD3-BJ using Maxwell-Garnett and Bruggeman effective medium theories with an 80% volume fraction of $\text{Na}_2(\text{SO}_4)_2$ in air. For all calculated ATR spectra in this section the effective medium is assumed to be on a slab of diamond with a refractive index of 2.4 with the angle of incidence of the incoming radiation was 45° and the radiation assumed to have equal S and P polarization.

The *lag* and *scale* factors used in Figure 19 are respectively 0.0 cm^{-1} and 1.04 for Maxwell-Garnett and 8.6 cm^{-1} and 1.04 for Bruggeman.

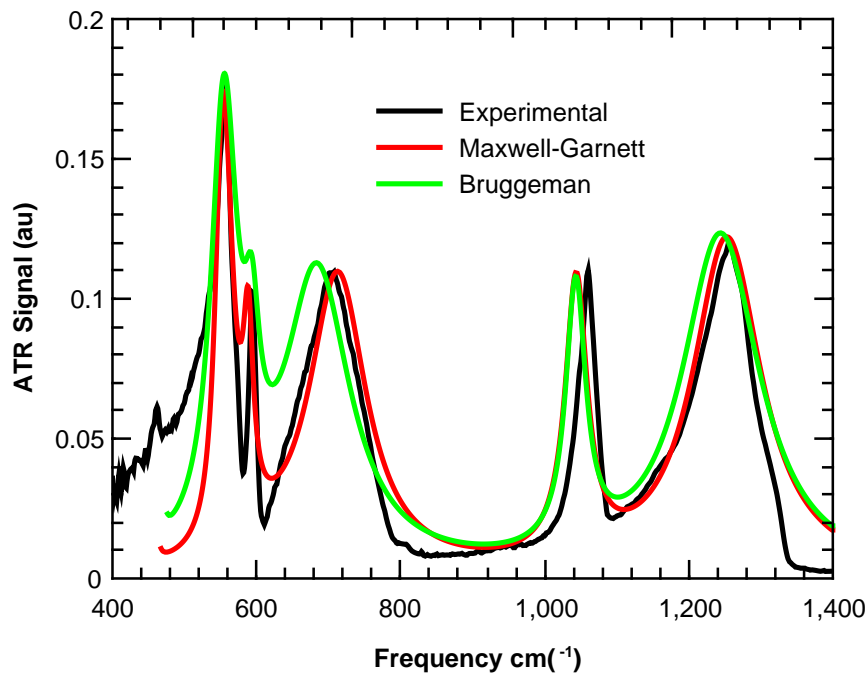


Figure 19: Comparison of Experimental ATR signal with Bruggeman and Maxwell-Garnett calculated simulations with scaled frequencies

Both effective medium theories show good agreement with experiment, when re-scaling of the frequency scale is employed. Examination of Table 8 shows that if frequency scaling is used in the comparison of calculated and experimental spectra, all of the methods show a cross-correlation over 0.81 and there is little to choose between the methods. The calculated ATR spectra for all calculations, along with both effective medium approximations for the Crystal/DEF2 calculation can be seen in the SI.

To compare all the calculated spectra against the experimental spectra a normalized cross-correlation coefficient (as previously discussed in Section 4.6) was calculated between the experimental spectrum in the range 450 to 1400 cm^{-1} . The calculations were performed with a Maxwell-Garnett effective medium representation of 80% volume fraction of spherical particles $\text{Na}_2(\text{SO}_4)_2$ in air. The Lorentzian widths of each transition were chosen so that the peak height of the calculated spectrum agreed with that of the experimental spectrum. The

reported cross-correlation coefficients Table 8 are the maximum coefficients at a constant frequency shift. There are therefore two parameters which are optimised to improve the fit with experiment, a frequency *lag* and a frequency *scale* factor. The first three columns in Table 8 show the results for the case that no frequency scaling is employed. The last three columns show the results after optimizing the frequency scaling factor to improve the cross-correlation coefficient.

Table 8: *scale* factors, *lag* shifts and cross-correlation coefficients between calculated and experimental ATR spectra

Method	Scale factor	Lag (cm ⁻¹)	Cross-correlation	Scale factor	Lag (cm ⁻¹)	Cross-correlation
Abinit/FHI	1	28.2	0.808	1.018	16.6	0.825
Abinit/FHI/GD2	1	21.2	0.835	1.012	13.0	0.844
Abinit/ONCVSP	1	28.8	0.818	1.020	15.8	0.833
Abinit/ONCVSP/GD2	1	22.0	0.845	1.010	15.2	0.851
CASTEP/NCP17	1	33.0	0.764	1.038	9.4	0.821
CASTEP/NCP17/GD3	1	28.8	0.798	1.032	8.2	0.834
CASTEP/NCP17/TS	1	24.8	0.807	1.035	1.4	0.839
CASTEP/NCP17/TS-v	1	22.6	0.812	1.030	2.4	0.840
Crystal/TZVP	1	76.0	0.642	1.011	68.2	0.646
Crystal/TZVP/GD2	1	68.4	0.784	1.009	61.8	0.786
Crystal/DEF2	1	36.0	0.830	1.020	22.8	0.843
Crystal/DEF2/GD2	1	32.0	0.848	1.008	26.4	0.849
Crystal/DEF2/GD2-v	1	33.4	0.846	1.009	26.8	0.851
Crystal/DEF2/GD2-r	1	31.0	0.834	1.002	30.0	0.834
QE/SSSP	1	30.0	0.832	1.014	20.4	0.839
QE/SSSP/GD2	1	24.6	0.853	1.008	18.8	0.854
VASP/PAW	1	28.4	0.738	1.049	-1.0	0.838
VASP/PAW/GD2	1	25.4	0.760	1.043	-1.0	0.851
VASP/PAW/GD3	1	25.2	0.758	1.042	0.0	0.845
VASP/PAW/GD3-BJ	1	23.6	0.766	1.039	0.0	0.841
VASP/PAW/TS	1	20.6	0.764	1.022	7.6	0.810

With no re-scaling of the frequencies the optimum cross-correlation coefficients are found by using a *lag* shift for the calculated spectra by between 20 and 36 cm⁻¹ to higher frequency. The combined use of re-scaling and shifting the frequencies results in almost all

methods having a cross-correlation with experiment of over 0.8. Only Crystal/TZVP and Crystal/TZVP/GD2 calculations have a cross-correlation below 0.8 and in addition they require a *lag* shift of over 60 cm^{-1} . The VASP calculations without re-scaling the frequencies have poor cross-correlations with experiment (below 0.8). However, with re-scaling the cross-correlation is as good as any of the others and the required *lag* shift in frequency to achieve the best cross-correlation coefficient is small. This behaviour is also shown by CASTEP calculations, where after re-scaling the frequency *lag* shift required to get the optimum cross-correlation is relatively small. The SI includes a similar comparison for the Bruggeman method. Without re-scaling the frequencies, on average the Bruggeman method requires an additional 9.2 cm^{-1} to find the maximum cross-correlation coefficient and the average cross-correlation coefficient is lower by -0.01. However with re-scaling the average cross-correlation coefficient increases by 0.017.

4.10 Transmission Infrared

The experimental transmission infrared spectrum is shown in Figure 20 and compared with Bruggeman and Maxwell-Garnett effective medium theory calculations based on VASP/PAW/GD3-BJ phonon calculations. The experimental absorption has been re-scaled to show similar peak heights to those calculated and the calculated frequencies re-scaled to improve the position of the calculated peaks. Additional experimental repeat measurements of differing sample concentrations are shown in the SI.

In Figure 20 the values of *lag* and *scale* are 4.02 cm^{-1} and 1.04 respectively which are similar to those needed for the ATR calculated spectra above with the differences likely owing to the wider and asymmetric peak shapes seen in the experimental spectra. This result indicates a systematic underestimation of the calculated absorption frequencies. Both Bruggeman and the Maxwell-Garnett effective medium theories predict very similar absorption in this region

and the same frequency scaling has been applied to both methods.

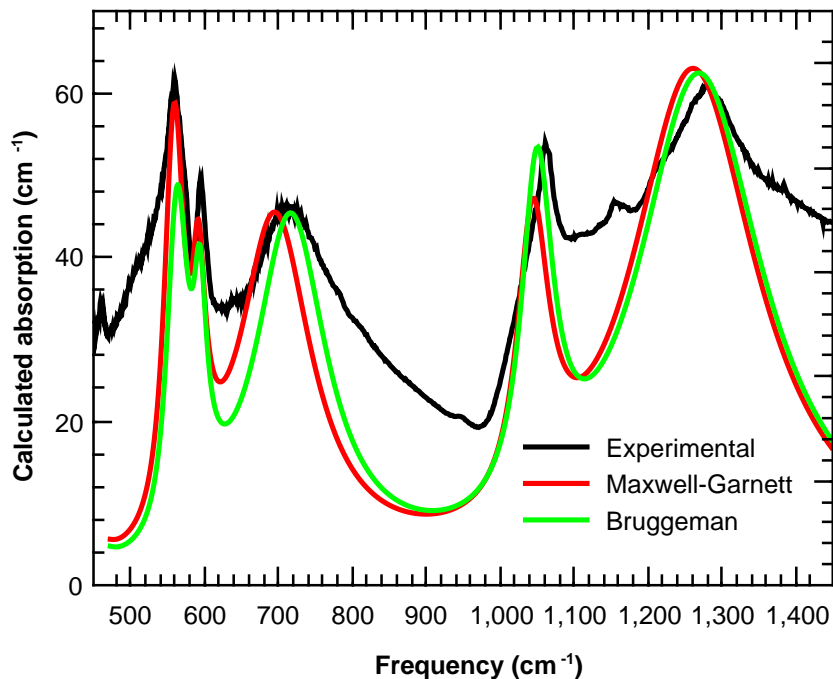


Figure 20: Experimental Infrared Spectrum for 2.67% mass fraction compared with Bruggeman and Maxwell-Garnett simulations with scaled frequencies

4.10.1 Terahertz Spectra

Figure 21 shows a comparison of the experimental room temperature terahertz spectrum and the calculated Maxwell-Garnett and Bruggeman effective medium theories using VASP/PAW/GD3-BJ phonon calculations. The experimental spectrum shows a strong background signal which is assumed to arise from scattering of air bubbles trapped in the PTFE supporting matrix.^{53,54} Both simulations reported in the figure account for this scattering through consideration of Mie scattering off a 15% volume fraction of 50 μm air bubbles. The calculated absolute absorption agrees well with experimental measurements. The peak positions and shapes predicted by the Bruggeman effective medium theory are in excellent agreement with experiment. The Maxwell-Garnett method predicts strong absorption at

too high a frequency compared with experiment. This shows the choice of effective medium approximation is often crucial at low frequencies to aid in spectral interpretation. Similar results can be seen for a number of the other calculations (not shown) although the best correlation with experiment at these low frequencies is with the VASP/PAW/GD3 phonon calculations using the Bruggeman effective medium approximation.

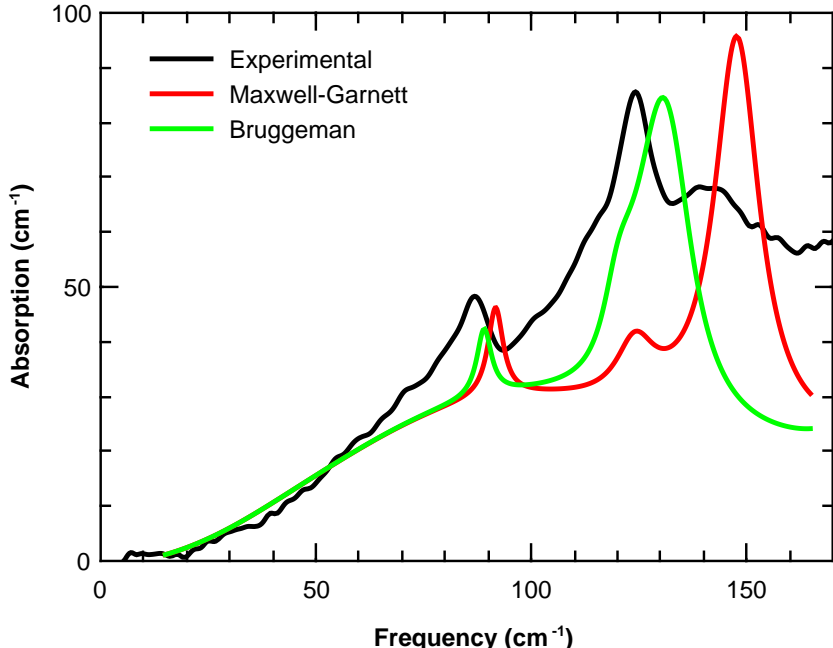


Figure 21: Comparison of Experimental room temperature terahertz measurements with Bruggeman and Maxwell-Garnett effective medium theory simulations

5 Conclusions

In this paper we have compared a number of infra-red and terahertz spectra of the powdered crystal $\text{Na}_2(\text{SO}_4)_2$ to the calculated spectra from a range of DFT programs with various combinations of pseudo-potentials and van der Waals' dispersive corrections. The inclusion of a van der Waals' dispersion correction has a significant effect on the calculated absorption spectrum, and are crucial for good correlation at low frequencies, where there is a systematic

shift of absorption to higher frequencies; shifts of over 40 cm^{-1} were seen. The default values of the dispersion correction parameters S_6 and S_R have been determined for a wide range of molecules and optimising these parameters to improve the predictions for a single molecule can lead to poor results, especially if only a single parameter such as the volume is chosen for improvement. Determining an optimum parameter does have an impact on the predicted spectrum, but generally speaking it is smaller than other factors in the calculation.

Low frequencies were particularly influenced by aliasing issues associated with the grids used to store the wavefunction and charge. But in all cases projection of the crystal translation from the dynamical matrix provided sensible results.

For the plane-wave based calculations convergence of the calculation was relatively straightforward to achieve. For Crystal it was more difficult to be sure that the atom centred basis set was adequate. The TZVP basis set was not adequate and the DEF2 basis which was slightly larger gave results very close to the plane-wave calculations.

The use of the cross-correlation of the predicted spectra to generate gap-maps of the similarities between the calculation was a useful tool and highlighted how the TZVP basis stood out from the other calculations. It also showed how the use of cell volume to determining S_6 in Crystal/DEF2/GD2-v resulted in an absorption spectrum which was different to other calculations.

In the calculations of the absorption spectrum, the use of effective medium methods to calculate the changes in absorption frequency and intensity owing to the interaction of the electromagnetic radiation field with the internal field generated by the vibrating crystal, was important over the whole frequency range. The Maxwell-Garnett method predicted the largest changes with shifts to higher frequencies of up to 40 cm^{-1} . The Bruggeman method tends to show broader absorption at frequencies intermediate between the TO frequencies and the absorption maxima predicted by Maxwell-Garnett. For small particle sizes relative

to the wavelength of the radiation, the Mie method for incorporating scattering effects agrees well with the Maxwell-Garnett effective medium theory. For particles smaller than $1\ \mu\text{m}$ the Maxwell-Garnett and Mie methods agree well up to $600\ \text{cm}^{-1}$, above this frequency the transitions get broader and show additional scattering artifacts in the calculated absorption. The incorporation of Mie scattering from air bubbles trapped in the support matrix greatly improves the agreement of the calculated THz spectrum with the experimental spectrum.

All the post analysis methods described in this paper including the effective medium approximations, Mie scattering, cross-correlation with experiment and the optimisation of *lag* shift and *scale* factors are available in the latest release of PDielec.⁵⁶

Acknowledgement

The authors thank Dr Paul Kendrick for helpful discussions on how to compare the infrared spectra using the cross-correlation between the absorption signals. ADB thanks the Engineering and Physical Sciences Research Council (EP/P007449/1) for funding. This work was undertaken on ARC3, part of the High Performance Computing facilities at the University of Leeds, UK. The data associated with this paper are openly available from the University of Leeds data repository.⁵⁷

References

- (1) Frisch, M. J. et al. Gaussian16 Revision B.01. 2016; Gaussian Inc. Wallingford CT.
- (2) Neese, F. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 73–78.

- (3) Clark, S. J.; Segall, M. D.; Pickard, C. J.; Hasnip, P. J.; Probert, M. I. J.; Refson, K.; Payne, M. C. *Zeitschrift für Kristallographie* **2005**, *220*, 567–570.
- (4) Hafner, J. *Journal of Computational Chemistry* **2008**, *29*, 2044–2078.
- (5) Gonze, X. et al. *Computer Physics Communications* **2016**, *205*, 106–131.
- (6) Dovesi, R.; Orlando, R.; Erba, A.; Zicovich-Wilson, C. M.; Civalleri, B.; Casassa, S.; Maschio, L.; Ferrabone, M.; De La Pierre, M.; D’Arco, P.; Noël, Y.; Causà, M.; Rérat, M.; Kirtman, B. *International Journal of Quantum Chemistry* **2014**, *114*, 1287–1317.
- (7) Giannozzi, P. et al. *Journal of Physics Condensed Matter* **2009**, *21*, 395502.
- (8) Kendrick, J.; Burnett, A. D. *Journal of Computational Chemistry* **2016**, *37*, 1491–1504.
- (9) Sihvola, A. In *Electromagnetic Mixing Formulas and Applications*; Clarricoats, P., Jull, E., Eds.; IET: The Institution of Engineering and Technology, Michael Faraday House, Six Hills Way, Stevenage SG1 2AY, UK, 1999.
- (10) Aufort, J.; Segalen, L.; Gervais, C.; Brouder, C.; Balan, E. *Physics and Chemistry of Minerals* **2016**, *43*, 615–626.
- (11) Van De Hulst, H. C. *Light Scattering by Small Particles*; Dover: New York, 1981; Vol. 1; p 470.
- (12) Ruggiero, M. T.; Kölbel, J.; Li, Q.; Zeitler, J. A. *Faraday discussions* **2018**, *211*, 425–439.
- (13) Burnett, A. D.; Kendrick, J.; Cunningham, J. E.; Hargreaves, M. D.; Munshi, T.; Edwards, H. G. M.; Linfield, E. H.; Davies, A. G. *ChemPhysChem* **2010**, *11*, 368–378.

- (14) Druzicki, K.; Mielcarek, J.; Kiwilsza, A.; Toupet, L.; Collet, E.; Pajzderska, A.; Wasicki, J. *Crystal Growth & Design* **2015**, *15*, 2817–2830.
- (15) Ruggiero, M. T.; Bardon, T.; Strlic, M.; Taday, P. F.; Korter, T. M. *The Journal of Physical Chemistry A* **2014**, *118*, 10101–10108.
- (16) Lejaeghere, K. et al. *Science* **2016**, *351*, aad3000–1–aad3000–7.
- (17) Hakey, P. M.; Allis, D. G.; Hudson, M. R.; Korter, T. M. *IEEE Sensors Journal* **2010**, *10*, 478–484.
- (18) Qu, F.; Pan, Y.; Lin, L.; Cai, C.; Dong, T.; He, Y.; Nie, P. *Journal of Infrared, Millimeter, and Terahertz Waves* **2018**, *39*, 1015–1027.
- (19) Qu, F.; Lin, L.; He, Y.; Nie, P.; Cai, C.; Dong, T.; Pan, Y.; Tang, Y.; Luo, S. *Molecules* **2018**, *23*, 1607.
- (20) Ullah, R.; Khan, S. U.-D.; Aamir, M.; Ullah, R. *Journal of Spectroscopy* **2013**, *2013*, 148903.
- (21) Cai, Q.; Xue, J.; Wang, Q.; Du, Y. *Journal of Molecular Structure* **2018**, *1153*, 170–178.
- (22) Qu, F.; Lin, L.; Cai, C.; Dong, T.; He, Y.; Nie, P. *Applied Sciences* **2018**, *8*, 420.
- (23) Hasegawa, R.; Kimura, T.; Tanabe, T.; Nishihara, K.; Taniyama, A.; Oyama, Y. *Journal of Biomedical Graphics and Computing* **2018**, *8*, 29–34.
- (24) Allis, D. G.; Hakey, P. M.; Korter, T. M. *Chemical Physics Letters* **2008**, *463*, 353–356.
- (25) King, M. D.; Buchanan, W. D.; Korter, T. M. *Physical Chemistry Chemical Physics* **2011**, *13*, 4250–4259.
- (26) Juliano, T. R.; King, M. D.; Korter, T. M. *IEEE Transactions on Terahertz Science and Technology* **2013**, *3*, 281–287.

- (27) Grazulis, S.; Daskevicius, A.; Merkys, A.; Chateigner, D.; Lutterotti, L.; Quiros, M.; Serebryanaya, N. R.; Moeck, P.; Downs, R. T.; Le Bail, A. *Nucleic Acids Research* **2012**, *40*, D420–D427.
- (28) Allan, D. R. *Acta Crystallographica Section E Structure Reports Online* **2006**, *62*, i44–i46.
- (29) Korter, T.; Balu, R.; Campbell, M.; Beard, M.; Gregurick, S.; Heilweil, E. *Chemical Physics Letters* **2006**, *418*, 65–70.
- (30) Balan, E.; Aufort, J.; Saldi, G. D.; Brouder, C.; Lazzeri, M. *European Journal of Mineralogy* **2018**, *31*, 73–81.
- (31) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Physical Review Letters* **1996**, *77*, 3865–3868.
- (32) Hamann, D. *Physical Review B* **2013**, *88*, 085117.
- (33) Fuchs, M.; Scheffler, M. *Computer Physics Communications* **1999**, *119*, 67–98.
- (34) Abinit FHI Pseudopotentials. https://www.abinit.org/sites/default/files/PrevAtomicData/psp-links/gga_fhi, Accessed on 17.05.2019.
- (35) G. Prandini, I. C. N. M., A. Marrazzo; Marzari, N. Standard Solid State Pseudopotentials (SSSP). <http://materialscloud.org/sssp>, or <https://doi.org/10.24435/materialscloud:2018.0001/v3>, Accessed on 17.05.2019.
- (36) Blöchl, P. E. *Physical Review B* **1994**, *50*, 17953.
- (37) Peintinger, M. F.; Oliveira, D. V.; Bredow, T. *Journal of Computational Chemistry* **2013**, *34*, 451–459.
- (38) Weigend, F.; Ahlrichs, R. *Physical Chemistry Chemical Physics* **2005**, *7*, 3297–3305.

- (39) Grimme, S. *Journal of Computational Chemistry* **2006**, *27*, 1787–1799.
- (40) Zhang, F.; Hayashi, M.; Wang, H.; Tominaga, K.; Kambara, O.; Nishizawa, J. N.; Sasaki, T. *The Journal of Chemical Physics* **2014**, *140*, 174509.
- (41) Zhang, H., Feng Wang; Tominaga, K.; Hayashi, M. *The Journal of Physical Chemistry B* **2016**, *120*, 1698–1710.
- (42) Nickel, D. V.; Delaney, S. P.; Bian, H.; Zheng, J.; Korter, T. M.; Mittleman, D. M. *The Journal of Physical Chemistry A* **2014**, *118*, 2442–2446.
- (43) Becke, A. D.; Johnson, E. R. *The Journal of Chemical Physics* **2005**, *123*, 154101.
- (44) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *The Journal of Chemical Physics* **2010**, *132*, 154104.
- (45) Tkatchenko, A.; Scheffler, M. *Physical Review Letters* **2009**, *102*, 073005.
- (46) Monkhorst, H. J.; Pack, J. D. *Physical Review B* **1976**, *13*, 5188–5192.
- (47) Balan, E.; Mauri, F.; Lemaire, C.; Brouder, C.; Guyot, F.; Saitta, A. M.; Devouard, B. *Physical Review Letters* **2002**, *89*, 177401.
- (48) Hecht, E. *Optics*, 5th ed.; Pearson Education Ltd: Boston, 2017.
- (49) Engle, S.; Whalen, S.; Joshi, A.; Pollard, K. S. *BMC Bioinformatics* **2017**, *18*, 63.
- (50) Crawley, M. J. *The R Book*; John Wiley and Sons, Ltd: Chichester, UK, 2007.
- (51) Bacon, D. R.; Burnett, A. D.; Swithenbank, M.; Russell, C.; Li, L.; Wood, C. D.; Cunningham, J.; Linfield, E. H.; Davies, A. G.; Dean, P.; Freeman, J. R. *Optics Express* **2016**, *24*, 26986–26997.

- (52) Brand, H. E.; Fortes, A. D.; Wood, I. G.; Knight, K. S.; Vočadlo, L. *Physics and Chemistry of Minerals* **2009**, *36*, 29–46.
- (53) Parrott, E. P.; Zeitler, J. A.; Gladden, L. F. *Optics Letters* **2009**, *34*, 3722–3724.
- (54) Zurk, L. M.; Orłowski, B.; Winebrenner, D. P.; Thorsos, E. I.; Leahy-Hoppa, M. R.; Hayden, L. M. *J. Opt. Soc. Am. B* **2007**, *24*, 2238–2243.
- (55) Merrick, J. P.; Moran, D.; Radom, L. *The Journal of Physical Chemistry A* **2007**, *111*, 11683–11700.
- (56) Kendrick, J.; Burnett, A. D. <https://doi.org/10.5281/zenodo.2529047> for recent updates please see <https://github.com/JohnKendrick/PDielec>.
- (57) Kendrick, J.; Burnett, A. D. Data associated with the calculation of the infrared and terahertz spectrum of sodium peroxodisulfate. <http://archive.researchdata.leeds.ac.uk/538/>.