Supplementary Information: Vibrational spectroscopy identifies the bond asymmetry of hexagonal diamond

Li Zhu
Department of Physics, Rutgers University, Newark, New Jersey 07102, USA

S1 Raman inversion: model, identifiability and uncertainty budget

Model. Linearized mode map about the relaxed structure ($O_A^0 = 1.5366$, $O_B^0 = 1.5607$ Å):

$$\omega_i(O_A, O_B) = \omega_i^0 + J_i \cdot (\Delta O_A, \Delta O_B) + \delta, \tag{1}$$

with reference frequencies $\omega^0 = (1211.0,\ 1301.2,\ 1329.2)$ cm^{-1} for (E_{2g}, A_{1g}, E_{1g}) and sensitivity matrix (cm^{-1} Å^{-1})

$$J_{E_{2g}} = (-3025,\ +300), \quad J_{A_{1g}} = (-300,\ -2156), \quad J_{E_{1g}} = (-2725,\ +300), \tag{2}$$

from the fixed-bond response surface (Sec. S2). δ is a common *additive* frequency offset (nuisance) bounded by the CD/graphite benchmarks computed with the identical workflow (CD T_{2g} -6.5 cm^{-1}; graphite G ≈ -13 cm^{-1} vs experiment): prior $|\delta| \leq 15$ cm^{-1}. A multiplicative scale s at the benchmarked 1% level alters a 28 cm^{-1} splitting by only $\sim$0.3 cm^{-1} and is subsumed in δ to leading order.

Identifiability. With two observed components ($y_A = 1310$, $y_E = 1338$ cm^{-1}, Lai) and three unknowns $(\Delta O_A, \Delta O_B, \delta)$ the system is underdetermined by one dimension. The one-parameter solution family is, however, nearly parallel to the $(\Delta O_A, \Delta O_B)$ diagonal: solving the two equations at fixed δ,

$$\frac{\partial(O_B - O_A)}{\partial\delta} = -0.005 \text{ mÅ per cm}^{-1}, \qquad \frac{\partial O_A}{\partial\delta} \approx \frac{\partial O_B}{\partial\delta} \approx +0.41 \text{ mÅ per cm}^{-1}. \tag{3}$$

The *asymmetry* is therefore identified nearly independently of the offset prior (total prior-induced spread < 0.15 mÅ over $|\delta| \leq 15$), while the *absolute* bond lengths inherit ± 6 mÅ from it. The asymmetry is controlled by the measured splitting: $\partial(O_B - O_A)/\partial(\text{splitting}) = 0.41$ mÅ per cm^{-1}.

Measured-cell consistency. The inversion is performed on the sensitivity map computed at the calculated lattice constants. Enforcing instead the measured cell of the phase-pure sample ($a = 2.5179$, $c = 4.1828$ Å), which reduces the structural freedom to the internal coordinate alone, and solving the two component equations for (z, δ) gives $z = 0.0625$, $O_A = 1.545$, $O_B = 1.569$ Å and $O_B - O_A = 24.0$ mÅ: the identical asymmetry, with both bonds shifted together by $+12$ mÅ. The required offset grows to $\delta = +28.7$ cm^{-1}, outside the benchmark-derived prior, for an identifiable reason: the measured cell is 1.6% larger in volume than the calculated one ($a = 2.5024$, $c = 4.1679$ Å), and mode-Grüneisen softening with $\gamma = 1.0$–1.5 predicts a uniform 21–31 cm^{-1} downshift of the computed frequencies at the measured cell, which the fitted δ absorbs; the benchmarks that set the ± 15 cm^{-1} prior were computed at DFT geometries and do not contain this cell term. The asymmetry is invariant along this family because it is fixed by the splitting (Eq. 3); only the absolute bond lengths trade off against δ. Exact finite-displacement phonons computed at this measured-cell structure confirm the linearized treatment: $E_{2g}/A_{1g}/E_{1g}$ = 1,189.0/1,280.6/1,308.8 cm^{-1}, a near-uniform -21 cm^{-1} shift from the calculated-cell equilibrium values (inside the Grüneisen estimate), with an $E_{1g} - A_{1g}$ splitting of 28.2 cm^{-1} against the measured 28.

Uncertainty budget for $O_B - O_A$ (95% half-widths, treated as independent effective uncertainties and combined in quadrature; no formal probability model beyond that is implied): component-fit uncertainty of the splitting ($\sigma \approx 3$ cm^{-1}) $\to \pm 2.4$ mÅ; functional dependence of the reference asymmetry (24.1 r^{2}SCAN+rVV10 vs 23.4 PBE) $\to \pm 0.8$; closure spread of J (7%) acting on the inferred correction $\to \pm 0.6$; offset prior $\to \pm 0.15$. Combined:

$$\boxed{O_B - O_A = 24 \pm 3 \text{ mÅ (95\%), phase-pure sample}} \tag{4}$$

excluding zero, Yang's -55 mÅ and Lai's $+183$ mÅ by wide margins (zero lies eight half-widths of the 95% interval from the central value). Lai *et al.* do not publish component-fit uncertainties; $\sigma \approx 3$ cm^{-1} is our estimate for the splitting of two narrow, separately fitted components whose positions are quoted to 1 cm^{-1}. The result scales linearly with this choice: doubling it to $\sigma = 6$ cm^{-1} widens the interval to ± 5 mÅ (95%) and changes no conclusion. The interval is conditional on the A_{1g}/E_{1g} assignment, which is fixed independently (assignment matrix below).

Scope of this interval. The budget above propagates the uncertainties of *one* measurement (the two narrow, separately fitted components of the phase-pure sample), and the interval should be quoted as such. It is not a between-sample interval. The independent lonsdaleite spectrum of Goryainov *et al.* provides the only available external test: their A_{1g} = 1,305 cm^{-1} lies 4 cm^{-1} from the relaxed prediction, but their A_{1g}–E_{1g} splitting of 51 cm^{-1} (against 28 predicted) maps through $\partial(O_B - O_A)/\partial(\text{splitting}) = 0.41$ mÅ per cm^{-1} to $O_B - O_A \approx 33$ mÅ. Those authors report band widths of 50–60 cm^{-1} and a $\sim$10 cm^{-1} low-energy shift from lonsdaleite imperfection, on a spectrum obtained by subtracting two spectra of a mixture containing 0–42% lonsdaleite. Taking the component-position uncertainty as a quarter to a third of the band width gives a splitting uncertain to 18–28 cm^{-1} (we adopt 20), hence $O_B - O_A = 33 \pm 8$ mÅ, consistent with the phase-pure value at 1.1σ. The three independent experimental determinations therefore read $+24 \pm 3$, $+33 \pm 8$ and $+60 \pm 45$ mÅ: mutually consistent, all positive, and none compatible with either recent refinement.

Assignment matrix. All six ordered assignments of the two measured components to the three modes were inverted:

Table S1: Assignment matrix for Lai's two measured components (1,310/1,338 cm^{-1}).

assignment	O_A (Å)	O_B (Å)	$O_B - O_A$ (mÅ)	implied state	third-mode prediction
A_{1g}/E_{1g}	1.533	1.557	+24	low stress	E_{2g} 1,221 (in envelope tail)
A_{1g}/E_{2g}	1.495	1.563	+67	$\sim$42 GPa in-plane	E_{1g} 1,443 (unobserved)
E_{2g}/A_{1g}	1.503	1.549	+46	$\sim$30 GPa in-plane	E_{1g} 1,418 (unobserved)
E_{1g}/A_{1g}	1.542	1.543	+1.2	near-equilibrium, low stress	E_{2g} 1,191 (exact; unobserved)
E_{2g}/E_{1g}	1.237	–	–	unphysical ($\Delta O_A = -300$ mÅ)	–
E_{1g}/E_{2g}	1.049	–	–	unphysical ($\Delta O_A = -487$ mÅ)	–

Two assignments are unphysical outright. Two (A_{1g}/E_{2g} and E_{2g}/A_{1g}) require tens-of-GPa in-plane stress states excluded by the ambient lattice constants and predict a strong E_{1g} band where nothing is observed. The remaining alternative, E_{1g}/A_{1g} (the mode ordering inverted), is mechanically innocuous: it inverts to a near-symmetric structure ($O_B - O_A = 1.2$ mÅ; 1.16–1.31 over $|\delta| \leq 15$ cm^{-1}) close to equilibrium, and neither energy nor stress excludes it. It is excluded by the spectra, on three counts. (i) Its third mode falls at 1,191 cm^{-1} (E_{2g}; the exact finite-displacement value at the swap structure, which also reproduces the two components at 1,312/1,338 cm^{-1} and so confirms the linearized inversion to ≤ 2 cm^{-1}), in-window and unobserved — and exact finite-field Raman tensors at the same structure (method of Sec. S5) give degeneracy-summed activities E_{2g}:A_{1g}:E_{1g} = 98:100:103: as the bonds approach equality the three bands equalize (the approach to the equal-bond cubic limit), so the unobserved 1,191 cm^{-1} band is predicted to be as strong as the two observed components. (The single-variable activity interpolation of Sec. S5, valid along the scan, fails in this bond-crossover region; the exact tensors supersede it here.) (ii) The same near-equal activities give no account of the strongly asymmetric measured envelope: its dominant 1,310 cm^{-1} component is reproduced by the preferred assignment as the dominant A_{1g} (100:41:28 at the relaxed structure), while the inverted reading predicts two components of equal strength. (iii) On the one sample where all three modes are resolved (Goryainov *et al.*), the E-splitting bound of Sec. S4 admits only the pairing $E_{2g} = 1{,}244/E_{1g} = 1{,}356$ (splitting 112 cm^{-1}, inside 107–131; the alternative pairings give 61 and 51), which forces $A_{1g} = 1{,}305$ as the middle, most intense band, i.e. the normal ordering; the same ordering appears in the simulated spectrum of Lai *et al.* Under the inverted reading the three-determination concordance of this section would also collapse (a +1 mÅ value against $+33 \pm 8$). The A_{1g}/E_{1g} assignment is thus fixed by intensity ordering, the third-mode constraint and the resolved third-sample spectrum, and in turn implies a scattering geometry (or texture distribution) with E_{1g} visibility; the rejections that invoke an unobserved E_{1g} band apply under that same visibility condition, so the analysis is self-consistent.

Scattering geometries. D_{6h} Raman tensors (calculated, arbitrary units): A_{1g}: $\alpha_{xx} = \alpha_{yy} = -1.31$, $\alpha_{zz} = +6.53$; E_{2g}: $(\alpha_{xx} - \alpha_{yy}, \alpha_{xy})$ components 2.15; E_{1g}: $(\alpha_{xz}, \alpha_{yz})$ components 1.78/0.48. Visibility by geometry: backscattering along c: A_{1g} (via α_{xx}), E_{2g} allowed; E_{1g} forbidden. Edge-on with in-plane analyzer: A_{1g}, E_{2g}. Edge-on with a c-axis polarization component: all three, A_{1g} strongly enhanced via α_{zz}. Powder average: all three with degeneracy-summed activities 100:41:28 (A_{1g}:E_{2g}:E_{1g}) at the relaxed geometry. Activities are static, non-resonant (PBE); both experiments used UV excitation where resonance may reweight intensities; frequencies are the primary evidence throughout.

S2 Fixed-bond response surface and constructions

Structures with (O_A, O_B) fixed were generated by the exact geometric closure $z = (c/2 - O_B)/(2c)$, $a = \sqrt{3(O_A^2 - (c/2 - O_B)^2)}$, with the remaining degree of freedom c resolved either by energy minimization over a 5-point scan (E-min closure) or fixed to the experimental value (fixed-c closure). Sensitivities were obtained from the fixed-O_A path ($O_A = 1.58$, $O_B = 1.44$–1.60 Å), the fixed-O_B pair ($O_B = 1.5607$, $O_A = 1.50/1.58$), and the fixed-cell z path; the A_{1g} O_B-sensitivity is -2000 ± 150 cm^{-1} Å^{-1} across all three constructions (the constructions differ in which cell parameters relax; the spread bounds the closure dependence).

S3 Physical state of all non-equilibrium structures

ΔE values for the Yang/Lai refined structures are static energies at the published coordinates relative to relaxed. Residual-force subtraction (phonopy `--fz`) was validated at the equilibrium structure (<4 cm^{-1} against the unconstrained calculation); frequencies of strongly non-stationary

Table S2: Physical-state data for every constrained structure. Stresses are positive-compressive (the confining stress required to hold the structure); residual forces are per-atom maxima in the undisplaced supercell before subtraction. No structure has imaginary Γ modes. Stress sign: positive = compressive (confining); negative = tensile.

structure	a (Å)	c (Å)	z	O_A/O_B (Å)	ΔE (meV/at)	σ_{xx}	σ_{zz}	$F_{\max}$ (eV/Å)
relaxed	2.5024	4.1679	0.0628	1.537/1.561	0	0	0	0.00
z-scan z=0.063	2.5024	4.1679	0.0630	1.537/1.559	∼0	–	–	0.07
Yang Rietveld	2.5182	4.1780	0.0693	1.565/1.510	31	−8.5	−2.0	2.18
Lai Rietveld	2.5179	4.1828	0.0479	1.508/1.691	122	−5.8	+8.2	3.54
Lai mixed	2.5150	4.1720	0.0504	1.512/1.666	87	−4.5	+7.5	3.13
z-scan O_B=1.442	2.5024	4.1679	0.0770	1.581/1.442	149	+0.3	+10.3	5.30
fixed-O_A=1.58, E-min	2.5423	4.0494	0.0722	1.580/1.440	104	−20.6	+54.1	3.76
fixed-O_A=1.58, exp-c	2.4951	4.1780	0.0777	1.580/1.440	164	+4.2	+6.9	5.55
fixed-O_A=1.56, E-min	2.5136	4.0245	0.0711	1.560/1.440	90	−7.0	+64.7	3.42
fixed-O_A=1.56, exp-c	2.4571	4.1780	0.0777	1.560/1.440	179	+24.9	+6.1	5.61
compressed LSQ	2.4640	3.9904	0.0674	1.521/1.457	81	22.1	77.3	1.95

structures are curvatures of constrained structural models, not predictions for free-standing phases, which is precisely their role in this analysis.

S4 Two-bond least-squares search: construction, residuals and uniqueness

Construction. The search asks whether *any* homogeneous $P6_3/mmc$ structure reproduces all three bands reported by Yang *et al.* (1,249, 1,321, 1,529 cm^{-1}). Over the (O_A, O_B) manifold we evaluate the linearized mode map of Sec. 1 (Eqs. 1–2) on a grid, and score each point by the offset-minimized root-mean-square misfit

$$R(O_A, O_B) = \min_{|\delta|\le 15} \sqrt{\tfrac{1}{3}\sum_i \left[\omega_i(O_A, O_B) + \delta - y_i\right]^2}, \tag{5}$$

with $y = (1249, 1321, 1529)$ cm^{-1} assigned to (E_{2g}, E_{1g}, A_{1g}) and δ the same common additive nuisance offset used in Sec. 1, bounded by the CD/graphite benchmarks. The resulting misfit surface is Fig. 5a of the main text. It has a single connected low-misfit island, doubly compressed relative to equilibrium; the least-squares point was then recomputed with exact finite-displacement DFT rather than the linearization.

Exact check at the least-squares point. At $(O_A, O_B) = (1.521, 1.457)$ Å with the energy-minimized cell ($a = 2.4640$, $c = 3.9904$ Å, $z = 0.06744$), the exact Γ phonons are $E_{2g} = 1{,}233$, $E_{1g} = 1{,}345$, $A_{1g} = 1{,}552$ cm^{-1}, against the linearized prediction (1,230/1,343/1,529). Per-band residuals against the measured values are +16, −24 and −23 cm^{-1} at $\delta = 0$, i.e. every band matched to within the 15–30 cm^{-1} scale error. The offset-minimized RMS misfit is 18.6 cm^{-1} at $\delta = -10.3$ cm^{-1}.

Residual structure. The residual is not random but dominated almost entirely by the E-mode splitting. From the sensitivity matrix of Sec. 1, the splitting is controlled by O_A alone to linear order, the O_B cross-terms cancelling exactly:

$$\omega(E_{1g}) - \omega(E_{2g}) = 118.2 + 300\,\Delta O_A \text{ cm}^{-1} \ (\Delta O_A \text{ in Å}). \tag{6}$$

The least-squares structure therefore predicts an E-splitting of 112 cm^{-1} where 72 cm^{-1} is observed; the +16 and −24 cm^{-1} per-band residuals are that single 40 cm^{-1} discrepancy distributed over two bands by the fit. This is worth stating plainly because it bounds how good any match of this kind can be.

Uniqueness. Eq. (6) also settles which assignments are admissible, without any further calculation. Across the entire range of O_A sampled in this work (1.50–1.58 Å) the predicted E-mode splitting is confined to 107–131 cm^{-1}. The three ways of assigning two of the measured bands to the two E modes give observed splittings of 72 cm^{-1} (1,249/1,321), 208 cm^{-1} (1,321/1,529) and 280 cm^{-1} (1,249/1,529). Only the first lies within 40 cm^{-1} of the accessible band; the other two are excluded by 77 and 149 cm^{-1} respectively, several times the scale error, for *every* structure on the manifold. The assignment used above is thus the only one worth searching under, and the low-misfit island it selects is unique within it. The complementary four-way assignment analysis for the two-component spectrum of Lai *et al.* is given in Sec. 1.

Disposition. The existence of this solution refutes the narrower claim that no HD structure reproduces all three bands, and the main text states the stronger claim instead: the matching structure lies 81 meV atom^{-1} above equilibrium, requires $\sigma_{xx} = 22.1$ and $\sigma_{zz} = 77.3$ GPa of confining stress (Table S2), and has lattice constants 2.2% and 4.5% smaller than the ambient values measured on the same sample. It describes hexagonal diamond under tens of gigapascals, not a recovered ambient specimen.

Surrogate documentation. The misfit surface of main-text Fig. 5a uses the linearized map of Eqs. 1–2; its accuracy at the far point is bounded by the exact check above (largest deviation 23 cm^{-1}, for A_{1g} at the least-squares point, where the linearization under-predicts the frequency). The energy surface of Fig. 5b is the quadratic form $E = k_{AA}\,\Delta O_A^2 + 2k_{AB}\,\Delta O_A\,\Delta O_B + k_{BB}\,\Delta O_B^2$ least-squares fitted to the 16 computed energies of the scan and Table-S2 structures, with $(k_{AA}, k_{AB}, k_{BB}) = (22{,}900,\ 1{,}220,\ 6{,}380)$ meV atom^{-1} Å^{-2}, rms residual 8 meV atom^{-1} and maximum residual 20 meV atom^{-1} (fit and points in the archived plotting script). Energies of structures with different cell closures differ by the closure spread discussed in Sec. S2; the $\gtrsim$80 meV atom^{-1} contrast used in the argument is an order of magnitude above both residual scales.

S5 Intensity-weighted comparison with the measured spectra

Figure 4 of the main text plots predicted mode positions as equal-height sticks, so that the frequency comparison is not conditioned on calculated intensities. This section supplies the intensity-weighted reading. Degeneracy-summed band activities were computed at three geometries spanning the scan (Methods; PBE, static and non-resonant):

Table S3: Degeneracy-summed relative band activities, normalized to $A_{1g} = 100$.

geometry	O_B (Å)	A_{1g}	E_{2g}	E_{1g}
z-scan z=0.055	1.626	100	3	0.4
relaxed	1.561	100	41	28
z-scan z=0.073	1.475	100	27	107

The trend is monotonic in O_B and brackets every structure considered, which is what licenses the qualitative readings below; activities were not computed separately at the refined coordinates, so these are interpolations of the computed trend and are used only as orderings, never as fitted quantities.

Relaxed structure (O_B = 1.561 Å). A dominant A_{1g} at 1,301 cm^{-1} overlapped by E_{1g} at 1,329 cm^{-1} (28%), with a weaker E_{2g} at 1,211 cm^{-1} (41%). The predicted spectrum is one asymmetric envelope with its maximum near 1,305 cm^{-1} and a low-frequency shoulder, the observed form in both samples and the reason the two fitted components of Lai *et al.* are reproduced to 9 cm^{-1} each.

Yang Rietveld ($O_B = 1.510$ Å). Interpolating between the relaxed and $O_B = 1.475$ rows, A_{1g} remains the dominant band; the model therefore predicts the *strongest* feature of the spectrum at 1,399 cm^{-1}. The strongest measured band is at 1,321 cm^{-1}, and nothing is reported near 1,399 cm^{-1}. The intensity ordering fails independently of the frequency comparison.

Lai Rietveld ($O_B = 1.691$ Å). Beyond the long-bond end of the sampled range, where A_{1g} is overwhelmingly dominant (100:3:0.4 already at 1.626 Å). The model thus places almost all of the predicted scattering strength in the A_{1g} band at 959 cm^{-1}, below the published measurement window, and predicts the in-window bands to be weak. Its failure is that the one band it does place in-window with appreciable strength (E_{1g}, 1,425 cm^{-1}) is unobserved, while the observed envelope is unaccounted for.

Short-O_B structures ($O_B = 1.442$–1.457 Å). Past the $O_B = 1.475$ row, A_{1g} and E_{1g} are comparable (100:27:107), so these structures predict two strong bands: one near 1,529–1,552 cm^{-1} and one near 1,161–1,345 cm^{-1}. This is the basis of the intensity argument in the main text against a locally compressed HD origin for the 1,529 cm^{-1} feature.

Two caveats apply throughout. Activities are static and non-resonant, whereas both experiments used ultraviolet excitation, where resonance can reweight bands; and they are computed at the PBE level. Frequencies therefore remain the primary evidence and these orderings are supporting. The A_{1g}-dominant pattern near equilibrium is independently corroborated by the simulated spectrum published by Lai *et al.*, computed with the same functional class.

S6 Diffraction sensitivity analysis of the published Lai pattern

Data provenance. Observed, calculated, background and difference curves for Extended Data Fig. 2a of Lai *et al.* were taken from the publisher-hosted source-data spreadsheet (623 rows, $2\theta = 14.96$–$40.02°$, $\lambda = 0.6888$ Å); no digitization from images was used. A uniform $+0.08°$ offset of the source data relative to the nominal peak positions was detected by comparison against the authors' own calculated curve and absorbed into peak positions.

Method. Ten reflections in range were modelled: (100), (002), (101), (102), (110), (103), (200), (112), (201), (004). The $hk0$ intensities are z-independent (verified analytically) and anchor scale and B_{iso}. Background-subtracted intensities were extracted by pseudo-Voigt fits (shared width scaling within the overlapping low-angle triplet); uncertainties and the full 10×10 covariance by block bootstrap over the residual noise (300 replicates). The intensity model used carbon Cromer–Mann form factors [S1], multiplicity, Lorentz-polarization, B_{iso}, and a one-parameter March–Dollase texture proxy along [001]. At each z on a grid 0.040–0.080, (scale, B_{iso}, r_{MD}) were re-optimized; χ^2_{eff} is a *sensitivity metric* (generalized least squares against the bootstrap covariance, inflated by the goodness-of-fit), not a formal likelihood: the original refinement used an 8th-order spherical-harmonic texture model (texture index 5.5) that a one-parameter proxy cannot reproduce, and no conventional $\Delta\chi^2$ confidence interpretation is claimed.

Result. The profile is flat: all $z \in [0.040, 0.080]$ lie within $\Delta\chi^2_{\mathrm{eff}} \leq 1$ of the minimum ($z = 0.0705$; minimum position shifts to 0.0645 under a diagonal-covariance treatment, itself a degeneracy signature). B_{iso} pins at the zero bound across the grid and r_{MD} tracks z, the classic ill-conditioning signatures; the most z-sensitive reflections ($00l$) are exactly those most affected by the strong texture. In every treatment examined, the fit at $z \approx 0.063$–0.070 is at least as good as at the refined $z = 0.0479$. Representative fits at $z = 0.0479$, 0.0628 and the minimum, the extracted intensity table, covariances and parameter traces are archived with the analysis script (`zrefit_lai.py`).

The sensitivity profile itself is shown in Fig. S1, and the peak-extraction quality-control overlay in Fig. S2.

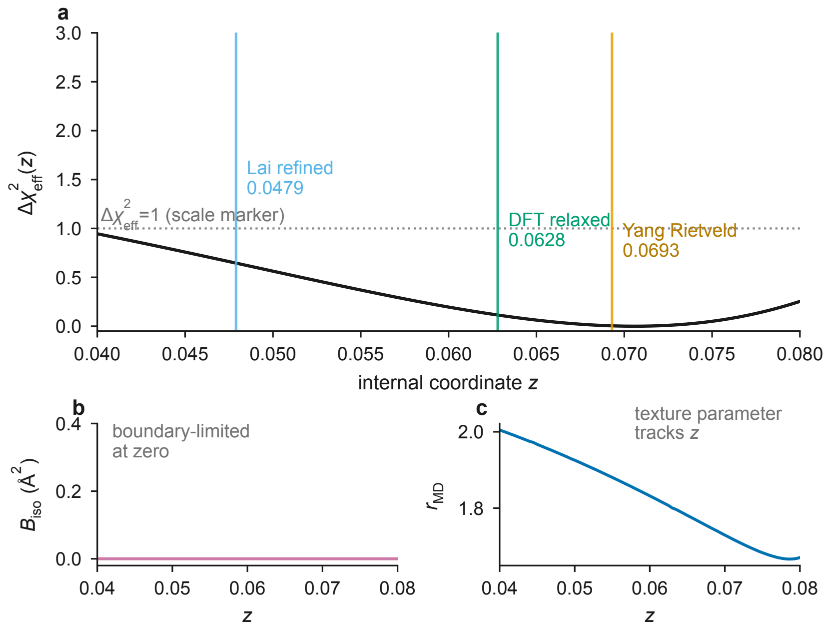

Figure S1: Sensitivity profile $\Delta\chi^2_{\mathrm{eff}}(z)$ for the published Lai pattern with the three coordinates marked, and the accompanying $B_{\mathrm{iso}}(z)$ and $r_{\mathrm{MD}}(z)$ traces showing the parameter degeneracy.

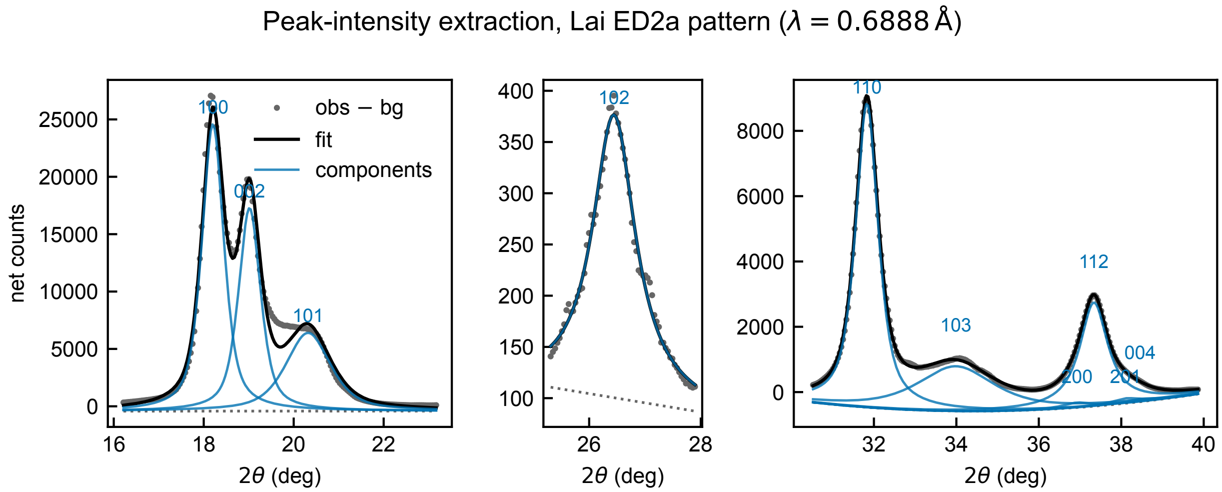

Figure S2: Quality control on the intensity extraction. Windowed pseudo-Voigt fits to the background-subtracted source data of Extended Data Fig. 2a of Lai *et al.*, with shared width scaling inside the overlapping low-angle triplet and free widths for the fault-broadened (101) and (103) reflections. The extracted integrated intensities and their block-bootstrap covariance are the input to the $\chi^2(z)$ profile of Fig. S1.

S7 Strained-sp^2 candidate for the high-frequency band

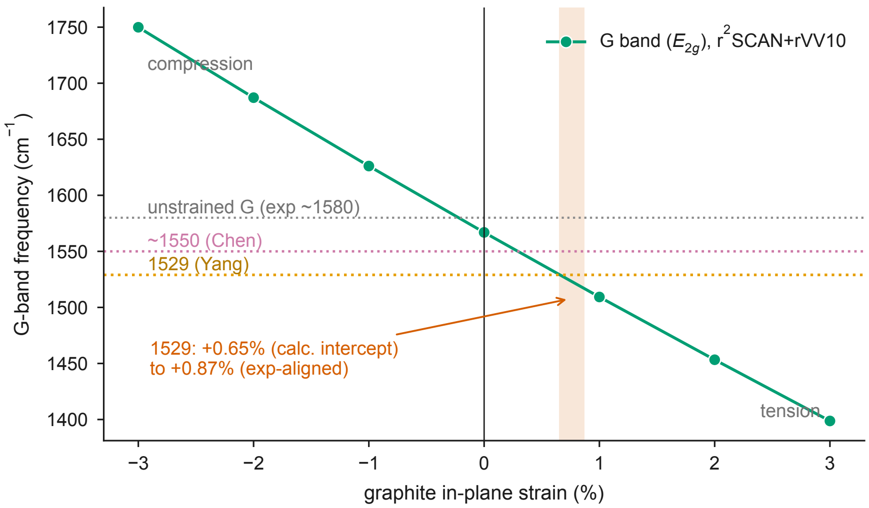

Figure S3: Graphite G band versus in-plane strain (-58.5 cm^{-1} %$^{-1}$; unstrained calculated 1,567 vs experimental ≈1,580 cm^{-1}). The 1,529 cm^{-1} band of Yang *et al.* corresponds to +0.65–0.87% tension depending on intercept calibration; the ≈1,550 cm^{-1} band of Chen *et al.* to +0.3–0.5%. This establishes plausibility of a strained minority-sp^2 origin; it does not identify the carrier (no interface model, resonance profile or linewidth calculation is implied).

S8 Numerical convergence and validation

Computed with the identical workflow: CD T_{2g} = 1,325.5 vs measured 1,332 cm^{-1} (−0.5%; $4\times4\times4$ supercell) and graphite G = 1,567 vs ≈1,580 cm^{-1} (−0.8%; $3\times3\times2$) bound the absolute frequency scale near equilibrium at ∼1%, consistent with the PBE-vs-r^{2}SCAN+rVV10 spread (12–26 cm^{-1}) on identical structures. Residual-force subtraction reproduces the unconstrained equilibrium frequencies to <4 cm^{-1}. The z-scan and refined-structure phonons use the same $3\times3\times2$ supercell, $7\times7\times5$ k-mesh, 800 eV cutoff and 10^{-8} eV SCF threshold as the validated equilibrium calculation; phonopy displacement amplitude 0.01 Å (default), Raman-tensor finite-field displacements ±0.02 Å. Symmetry assignments track irreducible representations computed at each geometry (no nearest-frequency relabeling across the scan). At strongly non-stationary structures (residual forces up to 5.6 eV Å^{-1}, Table S2) the leading finite-displacement error of a *one-sided* difference grows with the residual force through cubic anharmonicity; the calculations here are protected against it by construction, since all force constants are built from symmetric plus–minus displacement pairs (verified in the archived displacement metadata of the equilibrium, scan and refined-structure runs), which cancel the cubic term and leave a quartic-order residual. A direct test at the two highest-residual-force structures (2.2 and 5.6 eV Å^{-1}: the Yang-Rietveld and fixed-O_A=1.56/experimental-c rows of Table S2) confirms this: halving the displacement amplitude to 0.005 Å with forced plus–minus pairs reproduces every Raman-active frequency to 0.1 cm^{-1}. The discriminating margins in the main text (78–350 cm^{-1}) in any case exceed the equilibrium-validated scale error by an order of magnitude. All inputs are archived, so the $2\times2\times2$ supercell cross-check and the displacement-amplitude sensitivity test can be reproduced directly from the deposited data.

One-phonon spectral bound. The full phonon density of states of relaxed HD, computed from the archived equilibrium force constants on a $24\times24\times16$ wavevector mesh, has support up to 1,329 cm^{-1} only; the top of the one-phonon spectrum is the zone-centre E_{1g} mode. No one-phonon state of the relaxed structure, at any wavevector, lies within 200 cm^{-1} of the 1,529 cm^{-1} band discussed in the main text. The density of states is regenerable from the deposited `FORCE_SETS` with a single phonopy mesh calculation.

S9 Scope of spectral comparisons

All comparisons in the main text are made at the level of *reported band and fitted component positions*; no digitized experimental traces were used. Statements about envelope asymmetry and unresolved components refer to the published figures qualitatively and are worded accordingly. The underlying Raman traces of Lai *et al.* (their Supplementary Fig. 4) are not among that article's deposited source-data files, so component-fit line shapes and covariances are not independently accessible; the $\sigma \approx 3$ cm^{-1} adopted in Sec. S1, and its factor-of-two sensitivity bound, parametrize that dependence explicitly.

The archived data package (2,611 files, with a full manifest; Zenodo DOI 10.5281/zenodo.21712712) contains the structures (`POSCAR` and `CONTCAR` files) and VASP inputs for every calculation reported here; the phonopy force sets and displacement metadata; the converged `OUTCAR`s carrying the total energies, stress tensors, dielectric tensors and elastic constants quoted in the text and in Table S2; the LOBSTER COHP/ICOHP outputs behind the bonding analysis; the Bader and ELF outputs; the validation runs of this Supplementary Information (the assignment-check, measured-cell, displacement-amplitude and swap-structure Raman-tensor calculations); and every analysis and plotting script, including those that generate the published figures. Two classes of file are deliberately excluded: the VASP PAW potentials (`PAW_PBE C 08Apr2002`, VASP 6.4.2), whose licence forbids redistribution and which are identified by name instead; and charge densities and wavefunctions, which are regenerable from the included inputs. Everything required to reproduce the numbers and figures of this work is present, subject to supplying the licensed potential.

Supplementary references

[S1] Cromer, D. T. & Mann, J. B. X-ray scattering factors computed from numerical Hartree–Fock wave functions. *Acta Crystallographica Section A* **24**, 321–324 (1968).

[S2] Yang, L. *et al.* Synthesis of bulk hexagonal diamond. *Nature* **644**, 370–375 (2025).

[S3] Lai, S. *et al.* Bulk hexagonal diamond. *Nature* **651**, 621–625 (2026).

[S4] Chen, D. *et al.* General approach for synthesizing hexagonal diamond by heating post-graphite phases. *Nature Materials* **24**, 513–518 (2025).

[S5] Goryainov, S. V., Likhacheva, A. Yu. & Ovsyuk, N. N. Raman scattering in lonsdaleite. *Journal of Experimental and Theoretical Physics* **127**, 20–24 (2018).

[S6] Yoshiasa, A., Murai, Y., Ohtaka, O. & Katsura, T. Detailed structures of hexagonal diamond (lonsdaleite) and wurtzite-type BN. *Japanese Journal of Applied Physics* **42**, 1694–1704 (2003).

[S7] Dollase, W. A. Correction of intensities for preferred orientation in powder diffractometry: application of the March model. *Journal of Applied Crystallography* **19**, 267–272 (1986).

Experimental values quoted in this Supplementary Information are from refs. [S2] (Yang *et al.*), [S3] (Lai *et al.*), [S4] (Chen *et al.*), [S5] (Goryainov *et al.*) and [S6] (Yoshiasa *et al.*); the texture proxy follows ref. [S7]. The density-functional and phonon methodology references are given in the main-text reference list.