Linear hyperfine tuning of donor spins in silicon using hydrostatic strain:

Supplementary Material

J. Mansir,1 P. Conti,1 Z. Zeng,2 J.J. Pla,3 P. Bertet,4 M.W. Swift,5 C. G. Van de Walle,5 M.L.W. Thewalt,6 B. Sklenard,7 Y.M. Niquet,2 and J.J.L. Morton1,8

1London Centre for Nanotechnology, UCL, 17-19 Gordon St, London, WC1H 0AH, United Kingdom
2Université Grenoble Alpes, CEA, INAC-MEM, L_Sim, F-38000 Grenoble, France
3School of Elec. Engineering & Telecomm., University of New South Wales, Sydney, NSW 2052, Australia
4Quantronics Group, SPEC, CEA, CNRS, Université Paris-Saclay, CEA-Saclay, 91191 Gif-sur-Yvette, France
5Materials Department, University of California, Santa Barbara, California 93106-5050, USA
6Department of Physics, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6
7Université Grenoble Alpes & CEA, LETI, MINATEC Campus, F-38000 Grenoble, France
8Dept of Electronic and Electrical Engineering, UCL, London WC1E 7JE, UK
I. EXPERIMENTAL METHODS

The sample ‘Bi’ is a $2 \times 2 \times 10$ mm single crystal of isotopically purified $^{28}\text{Si}$ doped with $4.4 \times 10^{14}$ Bi donors/cm$^3$, and ‘Buffet’ is a $2 \times 2 \times 7$ mm $^{28}\text{Si}$ single crystal doped with $1.5 \times 10^{14}$ $^{31}\text{P}$ donors/cm$^3$, $5 \times 10^{14}$ $^{75}\text{As}$ donors/cm$^3$, and $1.1 \times 10^{14}$ $^{121}\text{Sb}$ donors/cm$^3$.

A cylindrical aluminium plate rests on the floor of the cryostat (see Fig. S1). A rod made from the plastic PEEK screws into this plate and extends into the centre of the sapphire resonator, providing a bottom support for the sample. A second PEEK rod, which is supported radially by the resonator structure but is free to move along its axial direction, holds the sample in place from above.

Each echo is averaged 300 times and the spins are reset after each cycle with a 5 ms flash of 50 mW above-bandgap 1047 nm laser light through an optical window in the cryostat.

The time-domain Hahn echo signals can be expressed as a periodic oscillation at the frequency of the detuning between the fixed microwave drive frequency and the strain-shifted transition frequency, multiplied by an envelope function given by the inverse Fourier transform of the frequency-domain transition spectral lineshape. Then, by the convolution theorem, the Fourier transform of such a signal results in a frequency-domain spectral peak centred at the detuning frequency. The strain-induced detuning is then extracted by fitting a Voigt profile to this spectral peak. See Ref. [1] Chapter 5 for more details.
FIG. S1. a) Schematic of experimental setup showing Bruker ESR resonator mounted inside liquid Helium flow cryostat. The sample is held in place in the centre of the sapphire resonator by two engineered plastic rods. b) Magnified view of sample inside strain mount.
II. TIGHT-BINDING MODELING

A. Model

We consider a single bismuth impurity at the center of a large box of silicon with side $L = 48a \simeq 26$ nm ($a = 5.431\text{Å}$ being the lattice parameter of silicon).

The electronic structure of silicon is described by the $sp^3d^5s^*$ tight-binding (TB) model of Ref. 2. This model reproduces the effects of arbitrary strains on the band edges and effective masses of silicon. Note that this model includes two $s$ orbitals per atom (“$s$” and “$s^*$”).

The bismuth impurity is described by a Coulomb tail and an “on-site” chemical and Coulomb correction [3, 4]. The expression of the Coulomb tail is based on the dielectric function proposed by Nara [5]. The potential on atom $i$ reads:

$$V_i = V(R_i) \text{ with } V(r) = -\frac{e^2}{\kappa r} \left(1 + A\kappa e^{-\alpha r} + (1 - A)\kappa e^{-\beta r} - e^{-\gamma r}\right),$$  \hspace{1cm} (S1)

where $R_i$ is the position of the atom (the bismuth impurity being at $R_1 = 0$), $\kappa = 11.7$ is the dielectric constant of silicon, $A = 1.175$, $\alpha = 0.757$ Bohrs$^{-1}$, $\beta = 0.312$ Bohrs$^{-1}$, and $\gamma = 2.044$ Bohrs$^{-1}$ [5, 6]. This expression deviates from a simple $-e^2/\kappa R_i$ tail mostly on the first and second nearest neighbors of the bismuth atom.

The “on-site” correction is a shift of the energies of the bismuth orbitals that accounts for the different chemical nature of the impurity and for the short-range part of the Coulomb tail. This shift $\Delta E = -U$ reads for each orbital:

$$U_s = 5.862 \text{ eV}$$
$$U_p = 3.690 \text{ eV}$$
$$U_d = 0.000 \text{ eV}$$
$$U_{s^*} = 5.862 \text{ eV}.$$  \hspace{1cm} (S2)

These values were adjusted on the experimental binding energies of the $1s(A_1)$, $1s(E)$ and $1s(T_2)$ states of bismuth in silicon (see Fig. S2) [7–9]. We have set $U_s = U_{s^*}$ on purpose since it is practically difficult to adjust $U_s$ and $U_{s^*}$ separately. We have designed an other model with $U_{s^*} = 0$ that gives very similar results.

We include spin-orbit coupling (SOC) in the calculations. SOC is described by an intra-atomic Hamiltonian acting on the $p$ orbitals of each atom, $H_{SO} = \lambda \mathbf{L}_i \cdot \mathbf{S}$, where $\mathbf{S}$ is the
FIG. S2. Experimental (Exp.) and TB bound states of a bismuth impurity in silicon. The horizontal dash-dotted line is the bulk conduction band edge.

spin, \( \mathbf{L}_i \) the angular momentum on atom \( i \), \( \lambda = 0.0185 \) eV for silicon, and \( \lambda = 0.350 \) eV for bismuth (adjusted on the experimental spin splittings of bismuth in silicon [9]).

The hyperfine coupling constant \( A \) is proportional to the probability of presence \( |\Psi(0)|^2 \) of the electron on the bismuth nucleus [Eq. (2) of main text]. In the TB framework,

\[
|\Psi(0)|^2 = |c_s s(0) + c_{s^*} s^*(0)|^2 = |c_s|^2 |s(0)|^2 \left| 1 + \frac{c_{s^*} s^*(0)}{c_s s(0)} \right|^2,
\]

(S3)

where \( c_s \) and \( c_{s^*} \) are the coefficients of the \( s \) and \( s^* \) orbitals of the bismuth atom in the TB wavefunctions (we discard the spin index here for the sake of simplicity). This expression is, in principle, ambiguous because the radial parts of the \( s \) and \( s^* \) orbitals of the TB model are not explicitly known. We have tentatively set \( R_{s^*s} = s^*(0)/s(0) = 0.058 \) [4]. Yet the choice for \( R_{s^*s} \) is practically little relevant, as \( c_{s^*}/c_s \) is almost independent on the strains \( \hat{\varepsilon} = \{\varepsilon_{xx}, \varepsilon_{yy}, \varepsilon_{zz}, \varepsilon_{yz}, \varepsilon_{xz}, \varepsilon_{xy}\} \). Therefore, the quantity

\[
A(\hat{\varepsilon}) = \frac{A(\hat{\varepsilon})}{A_0} = \frac{|\Psi(0, \hat{\varepsilon})|^2}{|\Psi(0, \hat{\varepsilon} = 0)|^2},
\]

(S4)

which describes the relative change of the hyperfine coupling constant under strains, is well defined within TB, irrespective of the assumptions made for the radial parts of the \( s \) and \( s^* \) orbitals.
B. Strains

We consider uniaxial stress along [001] and [110].

For uniaxial stress $\sigma_{zz} = \sigma_{\parallel}$ along [001], the infinitesimal strains in the cubic axis set can be found from Hooke’s law $\sigma_{xx} = \sigma_{yy} = 0$:

$$\varepsilon_{zz} = \varepsilon_{\parallel} = \frac{c_{11} + c_{12}}{(c_{11} - c_{12})(c_{11} + 2c_{12})}\sigma_{\parallel}$$

(S5a)

$$\varepsilon_{xx} = \varepsilon_{yy} = \varepsilon_{\perp} = -\frac{c_{12}}{c_{11} + c_{12}}\varepsilon_{\parallel},$$

(S5b)

where $c_{11} = 166$ GPa, $c_{12} = 64$ GPa and $c_{44} = 79.6$ GPa are the elastic constants of bulk silicon.

For uniaxial stress along [110], the infinitesimal strains in the $\{1 \equiv [110], 2 \equiv [\overline{1}10], 3 \equiv [001]\}$ axis set read:

$$\varepsilon_{11} = \varepsilon_{\parallel} = \frac{(c_{11} - c_{12})(c_{11} + 2c_{12}) + 2c_{11}c_{44}}{4(c_{11} - c_{12})(c_{11} + 2c_{12})c_{44}}\sigma_{\parallel}$$

(S6a)

$$\varepsilon_{22} = -\frac{(c_{11} - c_{12})(c_{11} + 2c_{12}) - 2c_{11}c_{44}}{4c_{12}c_{44}}\varepsilon_{\parallel}$$

(S6b)

$$\varepsilon_{33} = -\frac{(c_{11} - c_{12})(c_{11} + 2c_{12}) + 2c_{11}c_{44}}{4c_{12}c_{44}}\varepsilon_{\parallel}.$$  

(S6c)

In the original cubic axis set, the strains are therefore:

$$\varepsilon_{xx} = \varepsilon_{yy} = \frac{2c_{11}c_{44}}{(c_{11} - c_{12})(c_{11} + 2c_{12}) + 2c_{11}c_{44}}\varepsilon_{\parallel}$$

(S7a)

$$\varepsilon_{zz} = -\frac{4c_{12}c_{44}}{(c_{11} - c_{12})(c_{11} + 2c_{12}) + 2c_{11}c_{44}}\varepsilon_{\parallel}$$

(S7b)

$$\varepsilon_{xy} = \frac{(c_{11} - c_{12})(c_{11} + 2c_{12})}{(c_{11} - c_{12})(c_{11} + 2c_{12}) + 2c_{11}c_{44}}\varepsilon_{\parallel}.$$  

(S7c)

Note that there is an additional shear component with respect to uniaxial [001] stress.

C. Results

$\mathcal{A}(\varepsilon_{\parallel})$ is plotted in Fig. S3 for uniaxial [001] and [110] stress.

In the valley repopulation model (VRM) [10], $\mathcal{A}(\varepsilon_{\parallel})$ is expected to be quadratic with small $\varepsilon_{\parallel}$ (the changes in $\mathcal{A}$ being exclusively driven by the loss of symmetries). In the TB approximation, $\mathcal{A}(\varepsilon_{\parallel})$ indeed describes a parabola for weak stress, but centered on some $\varepsilon_{\parallel} > 0$. Therefore, $\mathcal{A}(\varepsilon_{\parallel})$ appears to behave almost linearly with small compressive $\varepsilon_{\parallel}$.
To understand this trend, it is very instructive to split the strain into a hydrostatic component, an uniaxial component, and a shear component. The hydrostatic component $\dot{\varepsilon}_{hs}$ is defined as $\varepsilon_{xx} = \varepsilon_{yy} = \varepsilon_{zz} = \dot{\varepsilon}_{hs}$. It accounts for the changes in the total volume $\Omega$ ($\Delta\Omega/\Omega = 3\dot{\varepsilon}_{hs}$). The uniaxial and shear components account for the changes in symmetries (at constant volume). The uniaxial component $\dot{\varepsilon}_{uni}$ is defined as $\varepsilon_{zz} = \dot{\varepsilon}_{uni}$, $\varepsilon_{xx} = \varepsilon_{yy} = -\dot{\varepsilon}_{uni}/2$, and the shear component $\dot{\varepsilon}_{shear}$ as $\varepsilon_{xy} = \dot{\varepsilon}_{shear}$.

For uniaxial stress along [001] [Eqs. (S5)],

$$\dot{\varepsilon}_{hs}^{001} = \frac{k}{3} \varepsilon_{\parallel},$$  \hspace{1cm} (S8a)

$$\dot{\varepsilon}_{uni}^{001} = \left(1 - \frac{k}{3}\right) \varepsilon_{\parallel},$$  \hspace{1cm} (S8b)

where:

$$k = \frac{c_{11} - c_{12}}{c_{11} + c_{12}}.$$  \hspace{1cm} (S9)
For uniaxial stress along [110] [Eqs. (S6)],

\[ \varepsilon_{110}^{hs} = \frac{2k_1 - k_2}{3} \varepsilon_{\parallel} \]  
\[ (S10a) \]

\[ \varepsilon_{110}^{uni} = -\frac{2(k_1 + k_2)}{3} \varepsilon_{\parallel} \]  
\[ (S10b) \]

\[ \varepsilon_{110}^{shear} = (1 - k_1) \varepsilon_{\parallel} \]  
\[ (S10c) \]

where:

\[ k_1 = \frac{2c_{11}c_{44}}{(c_{11} - c_{12})(c_{11} + 2c_{12}) + 2c_{11}c_{44}} \]  
\[ (S11a) \]

\[ k_2 = \frac{4c_{12}c_{44}}{(c_{11} - c_{12})(c_{11} + 2c_{12}) + 2c_{11}c_{44}} \]  
\[ (S11b) \]

\( \mathcal{A}(\hat{\varepsilon}) \) is plotted as a function of \( \hat{\varepsilon}_{hs} \), \( \hat{\varepsilon}_{uni} \), and \( \hat{\varepsilon}_{shear} \) in Fig. S4. \( \mathcal{A}(\hat{\varepsilon}) \) shows a quadratic behavior as a function of \( \varepsilon_{uni} \) and \( \varepsilon_{shear} \). It does, however, behave linearly as a function of \( \varepsilon_{hs} \). The trends at small \( \varepsilon_{\parallel} < 0 \) evidenced in Fig. S3 can, therefore, be ascribed to the effects of hydrostatic strains on the hyperfine coupling constant. As a matter of fact, \( \mathcal{A}(\varepsilon_{\parallel}) = \mathcal{A}(\hat{\varepsilon}_{hs}) + \mathcal{A}(\hat{\varepsilon}^{001}_{uni}) \) for uniaxial [001] stress, and \( \mathcal{A}(\varepsilon_{\parallel}) = \mathcal{A}(\hat{\varepsilon}_{hs}) + \mathcal{A}(\hat{\varepsilon}^{110}_{uni}) + \mathcal{A}(\hat{\varepsilon}^{110}_{shear}) \) for uniaxial [110] stress, showing the relevance of this decomposition.

**D. Discussion**

Although not predicted by the valley repopulation model, the existence of a \( \propto \varepsilon_{hs} \) term in \( \mathcal{A} \) is allowed by symmetries [11]. Indeed, a symmetry analysis suggests that, to second order in \( \varepsilon_{ij} \):

\[ \mathcal{A} - 1 = \frac{K}{3}(\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}) \]
\[ + L(\varepsilon_{xx}^2 + \varepsilon_{yy}^2 + \varepsilon_{zz}^2) + M(\varepsilon_{yy}\varepsilon_{zz} + \varepsilon_{xx}\varepsilon_{zz} + \varepsilon_{xx}\varepsilon_{yy}) + N(\varepsilon_{yz}^2 + \varepsilon_{xz}^2 + \varepsilon_{xy}^2), \]  
\[ (S12) \]

where \( K = \partial \mathcal{A}/\partial \varepsilon_{hs} \), \( L, M \) and \( N \) are constants. A fit to the tight-binding data yields:

\[ K = 29.3 \]
\[ L \approx -M = -9064 \]
\[ N = -225 \]  
\[ (S13) \]
FIG. S4. $A(\varepsilon)$ for (a) hydrostatic strain ($\varepsilon = \varepsilon_{hs}$), (b) uniaxial ($\varepsilon = \varepsilon_{uni}$) and shear strain ($\varepsilon = \varepsilon_{shear}$).

Eq. (S12) then simplifies into:

$$A - 1 = \frac{K}{3} (\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz})$$

$$+ \frac{L}{2} \left[ (\varepsilon_{yy} - \varepsilon_{zz})^2 + (\varepsilon_{xx} - \varepsilon_{zz})^2 + (\varepsilon_{xx} - \varepsilon_{yy})^2 \right] + N (\varepsilon_{yz}^2 + \varepsilon_{xz}^2 + \varepsilon_{xy}^2).$$  \hspace{1cm} (S14)

We find $L \simeq -M$ because there is no sizable non-linearity in the dependence of $A$ on hydrostatic strain in the investigated range $|\varepsilon_{hs}| < 10^{-4}$ (no $\propto \varepsilon_{hs}^2$ term above when $\varepsilon_{xx} = \varepsilon_{yy} = \varepsilon_{zz} = 0$).
\( \varepsilon_{yy} = \varepsilon_{zz} \). Also note that the effects of the quadratic shear terms are usually negligible with respect to the effects of the quadratic uniaxial terms \( N \ll L \).

The quadratic \( L \) term is mostly due to the coupling of the \( 1s(A_1) \) with the \( 1s(E) \) states of the impurity under uniaxial strain. The VRM of Ref. 10 actually suggests \( L = -2\Xi_u^2/(9\Delta^2) \), where \( \Xi_u = 8.6 \) eV is the uniaxial deformation potential of the conduction band of silicon and \( \Delta \) is the splitting between the \( 1s(A_1) \) and the \( 1s(E) \) state. For Bi \( (\Delta = 41 \) meV), the VRM predicts \( L = -9720 \), in close agreement with the TB data. The linear hydrostatic term is due, on the other hand, to the coupling of the \( 1s(A_1) \) with \( 2s(A_1) \) state (and possibly higher \( s \) states with the same symmetry). This coupling results from the variations of the on-site correction on the bismuth impurity, and from the variations of the bismuth-silicon interactions under hydrostatic strain – in other words, from the variations of the depth and shape of the “central cell correction” [12] not accounted for by the VRM.

Indeed, the total potential on the bismuth atom catches contributions from the tails of the atomic potentials of the neighboring silicon atoms, and these contributions depend on the silicon-bismuth bond lengths. As a matter of fact, the present TB model includes a strain-dependent correction for the energy \( E_{i\mu} \) of orbital \( \mu \equiv s, p, d, s^* \) of atom \( i \) [2]:

\[
E_{i\mu} = E_{i\mu}^0 + \frac{3}{4} \alpha_{i\mu} \sum_{j \in \text{NN}(i)} \frac{d_{ij} - d_{ij}^0}{d_{ij}^0} = E_{i\mu}^0 + 3\alpha_{i\mu} \varepsilon_{hs} , \tag{S15}
\]

where the sum runs over the nearest neighbors \( j \) of atom \( i \), \( d_{ij} \) is the distance between atoms \( i \) and \( j \), and \( d_{ij}^0 \) is the relaxed bond length. \( E_{i\mu}^0 \) is the energy of the orbital in the reference, unstrained system and \( \alpha_{i\mu} \) characterizes the deepening of the potential under strain. The interactions between bismuth and the nearest neighbor silicon atoms scale, on the other hand, as \( (d_{ij}/d_{ij}^0)^n \), with \( n \) close to 2.

In the present TB model, the parameters \( \alpha_{i\mu} \) and the exponents \( n \) of bismuth are the same as for silicon. Although this choice is a safe first guess, it can only provide a semi-quantitative description of the dependence of \( A \) on hydrostatic strain. This is why the TB \( K = 29.3 \) is significantly larger than the experimental \( K = 19.1 \). We may, in the spirit of Eqs. (S2), lump all corrections to the TB model into the \( \alpha_{i\mu} \). Therefore, we tentatively set:

\[
\alpha_{i\mu}(\text{Bi}) = \alpha_{i\mu}(\text{Si}) + \Delta \alpha , \tag{S16}
\]

and adjust \( \Delta \alpha \) on the experimental \( K \). This yields \( \Delta \alpha = 3.48 \) eV for bismuth.
TABLE SI. Nature of the Coulomb tail, on-site corrections $U$ [Eq. (S2)] and $\Delta \alpha$ [Eq. (S16)], and value of $K$, $L$ and $N$ for P, As, Sb and Bi donors in silicon.

|        | Coulomb tail | $U_s = U_{s^*}$ (eV) | $U_p$ (eV) | $U_d$ (eV) | $\Delta \alpha$ (eV) | $K$   | $L$   | $N$   |
|--------|--------------|-----------------------|------------|------------|----------------------|-------|-------|-------|
| P      | $-e^2/(\kappa R_i)$ | 4.535 | 2.405 | 2.055 | -19.50 | 79.5 | -103640 | 1277 |
| As     | $-e^2/(\kappa R_i)$ | 5.060 | 2.330 | 0.325 | -2.68 | 37.2 | -33836 | 833  |
| Sb Nara [Eq. (S1)] | 4.629 | 4.448 | 0.000 | -2.54 | 32.6 | -104340 | 1420 |
| Bi Nara [Eq. (S1)] | 5.862 | 3.690 | 0.000 | 3.48  | 19.1 | -9064  | -225 |

We have repeated the same procedure for P, As and Sb. We give in Table SI the on-site parameters $U_s = U_{s^*}$, $U_p$, $U_d$ and $\Delta \alpha$ of each impurity, as well as the values of $K$, $L$ and $N$. The model for P and As [13] is based on a simple Coulomb tail $V(R_i) = -e^2/(\kappa R_i)$ instead of Eq. (S1).

The dependence of the binding energy $E_b$ of As impurities on the hydrostatic pressure $P$ has been measured by Holland and Paul ($dE_b/dP \simeq -0.05$ meV/kbar) [14] and by Samara and Barnes ($dE_b/dP \simeq -0.1$ meV/kbar) [15]. The electron hence gets more loosely bound to the impurity under pressure (or equivalently under compressive hydrostatic strain). This is consistent with the decrease of the hyperfine coupling constant $A$ reported here ($K < 0$). Samara and Barnes explain the decrease of $E_b$ under pressure by the variations of the effective masses and dielectric constant $\kappa$. We point out, though, that there is also a significant contribution from the variations of the central cell correction. The variations of effective masses and dielectric constant actually make little contribution to $K$. In the simplest effective mass approximation, the wave function of the electron bound to the donor is indeed $\Psi(r) = e^{-r/a_B}/(\sqrt{\pi} a_B^{3/2})$, where $a_B = \hbar^2 \kappa/(m^* e^2)$ is the Bohr radius. Hence, $|\Psi(0)|^2 = 1/(\pi a_B^3)$, so that:

$$K = 3 \left( \frac{1}{m^*} \frac{\partial m^*}{\partial \varepsilon_{hs}} - \frac{1}{\kappa} \frac{\partial \kappa}{\partial \varepsilon_{hs}} \right)$$

($S17$)

Ab-initio calculations within density functional theory (see next section) give $(1/\kappa)(\partial \kappa/\partial \varepsilon_{hs}) = 0.78$, $(1/m^*)(\partial m^*/\partial \varepsilon_{hs}) = -0.19$ for the longitudinal mass, and $(1/m^*)(\partial m^*/\partial \varepsilon_{hs}) = 1.54$ for the transverse mass. Therefore, the variations of the masses and dielectric constant are expected to have little net effect on the hyperfine coupling constant.
III. DENSITY FUNCTIONAL THEORY CALCULATIONS

In order to strengthen the above interpretation, we have also performed first principles calculations using density functional theory (DFT) with the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional [16] and the projector-augmented wave method [17] in the Vienna Ab-initio Simulation Package (VASP) [18], following the methodology described in Ref. [19]. The calculations were carried out on one Bi impurity in a 1728-atom supercell, with a 250 eV plane-wave cutoff energy. DFT describes the central cell correction around the bismuth impurity from first principles and captures the atomic relaxations not accounted for by TB calculations — due to its accuracy in the immediate vicinity of the donor it could be expected to provide a good description of the variation of the hyperfine coupling with strains. Nevertheless, due to the finite size of the supercell, DFT misses the long range Coulomb tail of the potential [20], which (along with over-delocalization arising from the self-interaction error in PBE) contributes to an significant underestimation in the absolute value the hyperfine interaction (1102 MHz).

The ab-initio hyperfine coupling shows the expected linear dependence on hydrostatic strain over the entire range explored here (up to \( \varepsilon_{xx} = \varepsilon_{yy} = \varepsilon_{zz} = 2 \times 10^{-3} \)). We extract a coefficient \( K = 17.5 \), in good agreement with the experimental data. The ab-initio quadratic term is \( L = -1.17 \times 10^4 \), as obtained from a fit to the calculated data with \( \varepsilon_{zz} \leq 10^{-3} \) and \( \varepsilon_{xx} = \varepsilon_{yy} = 0 \). Deviations are observed for higher strains, consistent with the higher-order terms present in the VRM. The equilibrium bismuth-silicon bond length (2.651 Å) is significantly larger than the Si-Si bond length (2.367 Å in the bulk). All Bi-Si and Si-Si bonds are simply scaled by the hydrostatic strain, to within better than 0.001 Å.

These data provide further support for the linear dependence of the hyperfine parameter on the hydrostatic component of strain, and illustrate the complementary strengths of these two computational approaches (DFT and tight-binding) in modelling the behaviour of donors in silicon.

IV. MODELING G-FACTOR ANISOTROPY

The ellipsoidal shape of the Si conduction band minima in \( k \)-space results in differing effective masses for Bloch waves parallel and perpendicular to the valley axis at these points [21].
This leads to a single-valley g-factor which is anisotropic, with \( g_\mu^2 = g_\parallel^2 \cos^2 \phi + g_\perp^2 \sin^2 \phi \), where \( g_\parallel \) and \( g_\perp \) are the g-factors parallel and perpendicular to the valley axis, and \( \phi \) is the angle between the magnetic field and the valley axis. In the case of a donor, the g-factor is found by summing over the relative contribution from each valley such that

\[
g_{\text{donor}} = \sum_{\mu=1}^{6} \alpha_\mu^2 g_\mu.
\]

For the unstrained 1s(A_1) donor ground state, which is an equal superposition of all six equivalent valleys, this summation leads to a cancellation of the anisotropy, leaving \( g_0 = \frac{1}{3} g_\parallel + \frac{2}{3} g_\perp \). Under strain, the valleys repopulate, breaking this cancellation symmetry. In our system with \( \theta \) defined as in figure S1, the resulting g-factor anisotropy can be modeled using the VRM [22] by:

\[
\left. \frac{\partial g}{\partial \varepsilon_{11}} \right|_{\text{VRM}} = \beta_{\text{VRM}} (3 \cos^2 \theta - 1)
\]

where:

\[
\beta_{\text{VRM}} = \frac{2 \Xi_u}{9 \Delta} (g_\parallel - g_\perp)(k_1 + k_2).
\]

(\text{S19})

\( g_\parallel \) and \( g_\perp \) are the parallel and perpendicular g-factors, \( \Xi_u \) is the uniaxial deformation potential, \( \Delta \) is the donor-dependent 1s(A_1)-1s(E) splitting, and \( k_1 \) and \( k_2 \) are defined in Eq. S11.

It is also known that the g-factor of a single valley is changed in the presence of shear strain by the coupling with the opposite valley. Following Wilson & Feher [22], the effective Hamiltonian for this mechanism can be written [23]:

\[
\hat{H}_{\text{shear}} = C \mu B \varepsilon_{xy} (B_x S_y + B_y S_x)
\]

(\text{S20})

where \( C \) is a coefficient involving spin-orbit coupling matrix elements [22]. Rewriting the corresponding g-factor contribution in terms of \( \varepsilon_{11} \) results in a second anisotropic term:

\[
\left. \frac{\partial g}{\partial \varepsilon_{11}} \right|_{\text{shear}} = \beta_{\text{shear}} (3 \cos^2 \theta - 3)
\]

(\text{S21})

where:

\[
\beta_{\text{shear}} = \frac{C(1 - k_1)}{3}.
\]

(\text{S22})

Then, the derivative of the ESR transition frequency with respect to the uniaxial strain reads:

\[
\frac{df}{d\varepsilon_{11}} = \frac{\partial A}{\partial \varepsilon_{11}} \frac{df}{\partial A} + \left[ \left. \frac{\partial g}{\partial \varepsilon_{11}} \right|_{\text{VRM}} + \left. \frac{\partial g}{\partial \varepsilon_{11}} \right|_{\text{shear}} \right] \frac{\partial f}{\partial g}.
\]

(\text{S23})
$K = \frac{\partial (A/A_0)}{\partial \epsilon_{hs}}$ (exp.)

| Donor | $\beta_{\text{VRM}} \times 10^{-3}$ (theory) | $\beta_{\text{VRM}} \times 10^{-3}$ (exp.) | $\beta_{\text{shear}} \times 10^{-3}$ (theory) | $\beta_{\text{shear}} \times 10^{-3}$ (exp.) |
|-------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| $^{31}\text{P}$ | $79.2 \pm 25.7$ | $126.7$ | $206 \pm 51$ | $78$ | $173 \pm 33$ |
| $^{75}\text{As}$ | $37.4 \pm 3.3$ | $77.5$ | $165.1 \pm 24.6$ | $78$ | $96 \pm 15.6$ |
| $^{121}\text{Sb}$ | $32.8 \pm 2.3$ | $146.0$ | $197.6 \pm 25.8$ | $78$ | $90.3 \pm 16.4$ |
| $^{209}\text{Bi}$ | $19.1 \pm 0.3$ | $42.5$ | - | $78$ | - |

TABLE SII. Parameters $K = \partial (A/A_0)/\partial \epsilon_{hs}$, $\beta_{\text{VRM}}$ and $\beta_{\text{shear}}$ extracted from the experimental data, along with theoretical values for $\beta_{\text{VRM}}$ and $\beta_{\text{shear}}$. The theoretical values are calculated using $C = 0.44$ [22], $\Delta$ values from Ref. 8, and $(g_\parallel - g_\perp)$ values from Ref. [22]. We assume $(g_\parallel - g_\perp) \approx 1.1 \times 10^{-3}$ for Si:Bi.

$\partial f/\partial g$ can be calculated for each transition in the same manner as $\partial f/\partial A$ by solving the spin Hamiltonian while varying the value of $g$. We use a linear least squares regression to fit this model to the experimental data. We introduce three fitting parameters characterising the strength of the different effects for each donor: $K = \partial (A/A_0)/\partial \epsilon_{hs}$, $\beta_{\text{VRM}}$ and $\beta_{\text{shear}}$. The results of these fits are reported in table SII and figure 4 in the main text for Si:Sb, Si:As, and Si:P. The absence of a clear anisotropy for Si:Bi could be explained by the relatively small magnitude of the predicted g-factor effects in comparison with the absolute shifts due to the modified hyperfine interaction.

We compare the extracted parameters with theoretical predictions for $\beta_{\text{VRM}}$ and $\beta_{\text{shear}}$ [Eqs. (S19) and (S22)] calculated using $C = 0.44$ [22], $\Delta$ values from Ref. 8, and $g_\parallel - g_\perp$ values from Ref. [22]. The measured strengths of the g-factor effects agree with theory within approximately a factor of 2. Tight binding simulations with the present model (see also Ref. 24) predict that the very small $g_\parallel - g_\perp$ is approximately an order of magnitude larger than given by Ref. [22]. It is interesting to note that the experimental data sit between Ref. [22] and the TB predictions.
V. FULL EXPERIMENTAL DATASET FOR ALL DONORS

FIG. S5. Full dataset for Si:Bi showing frequency shifts for all ten ESR transitions as a function of strain $\varepsilon_{11}$ and $\theta$. Linear fits are shown for each transition.
FIG. S6. Full dataset for Si:Sb showing frequency shifts for all six ESR transitions as a function of strain $\varepsilon_{11}$ and $\theta$. Linear fits are shown for each transition.
FIG. S7. Full dataset for Si:As showing frequency shifts for all four ESR transitions as a function of strain $\varepsilon_{11}$ and $\theta$. Linear fits are shown for each transition.
FIG. S8. Full dataset for Si:P showing frequency shifts for both ESR transitions as a function of strain $\varepsilon_{11}$ and $\theta$. Linear fits are shown for each transition.
[1] A. Schweiger and G. Jeschke, *Principles of pulse electron paramagnetic resonance* (Oxford University Press, 2005).

[2] Y. M. Niquet, D. Rideau, C. Tavernier, H. Jaouen, and X. Blase, Physical Review B 79, 245201 (2009).

[3] B. Roche, E. Dupont-Ferrier, B. Voisin, M. Cobian, X. Jehl, R. Wacquez, M. Vinet, Y.-M. Niquet, and M. Sanquer, Physical Review Letters 108, 206812 (2012).

[4] M. Usman, R. Rahman, J. Salfi, J. Bocquel, B. Voisin, S. Rogge, G. Klimeck, and L. L. C. Hollenberg, Journal of Physics: Condensed Matter 27, 154207 (2015).

[5] H. Nara and A. Morita, Journal of the Physical Society of Japan 21, 1852 (1966).

[6] J. Bernholc and S. T. Pantelides, Physical Review B 15, 4935 (1977).

[7] W. E. Krag, W. H. Kleiner, and H. J. Zieger, in *Proceedings of the 10th International Conference on the Physics of Semiconductors (Cambridge, Massachusetts, 1970)*, edited by S. P. Keller, J. C. Hensel, and F. Stern (USAEC Division of Technical Information, Washington D.C., 1970) p. 271.

[8] A. K. Ramdas and S. Rodriguez, Reports on Progress in Physics 44, 1297 (1981).

[9] R. K. Zhukavin, K. A. Kovalevsky, V. V. Tsyplenkov, V. N. Shastin, S. G. Pavlov, H.-W. Hübers, H. Riemann, N. V. Abrosimov, and A. K. Ramdas, Applied Physics Letters 99, 171108 (2011).

[10] D. K. Wilson and G. Feher, Physical Review 124, 1068 (1961).

[11] D. Pines, J. Bardeen, and C. P. Slichter, Phys. Rev. 106, 489 (1957).

[12] N. Lipari, A. Baldereschi, and M. Thewalt, Solid State Communications 33, 277 (1980).

[13] M. Diarra, Y.-M. Niquet, C. Delerue, and G. Allan, Physical Review B 75, 045301 (2007).

[14] M. G. Holland and W. Paul, Physical Review 128, 30 (1962).

[15] G. A. Samara and C. E. Barnes, Physical Review B 35, 7575 (1987).

[16] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. 77, 3865 (1996).

[17] P. E. Blöchl, Phys. Rev. B 50, 17953 (1994).

[18] G. Kresse and J. Furthmüller, Phys. Rev. B 54, 11169 (1996).

[19] P. E. Blöchl, Phys. Rev. B 62, 6158 (2000).

[20] Y. M. Niquet, L. Genovese, C. Delerue, and T. Deutsch, Physical Review B 81, 161301
[21] P. Y. Yu and M. Cardona, *Fundamentals of semiconductors: physics and materials properties* (Springer, 1999).

[22] D. K. Wilson and G. Feher, Physical Review **124**, 1068 (1961).

[23] The form of Eq. (9) of Ref. [22] suggests that $\epsilon_{xy}$ in Ref. [22] is the “engineering” shear strain $\epsilon_{xy} = 2\varepsilon_{xy}$, hence the factor 2 difference between Eq. (S20) and Eq. (8) of Ref. [22].

[24] R. Rahman, S. H. Park, T. B. Boykin, G. Klimeck, S. Rogge, and L. C. L. Hollenberg, Physical Review B **80**, 155301 (2009).