

Impurity-Driven Two-Dimensional Spin Relaxation Induced by Intervalley Spin-Flip Scattering in Silicon

Yang Song* and S. Das Sarma

Department of Physics, Condensed Matter Theory Center, University of Maryland, College Park, Maryland 20742, USA and Joint Quantum Institute, University of Maryland, College Park, Maryland 20742, USA

(Received 10 July 2016; revised manuscript received 30 September 2016; published 6 January 2017)

Through the theoretical study of electron spin lifetime in the two-dimensional electron gas (2DEG) confined near the surface of doped Si, we highlight a dominant spin-relaxation mechanism induced by the impurity central-cell potential near an interface via intervalley electron scattering. At low temperatures and with modest doping, this Yafet spin-flip mechanism can become more important than the D'yakonov-Perel' spin relaxation arising from the structural Rashba or Dresselhaus spin-orbit-coupling field. As the leading-order impurity-induced spin flip happens only between two nonopposing valleys in Si, 2DEG systems in Si MOSFETs or SiGe heterostructures are a natural platform to test and utilize this spin-relaxation mechanism due to the valley splitting near the interface and the tunability by electrical gating or applied stress. Our proposed alternative spin-relaxation mechanism may explain part of the spin-relaxation contribution to Si-based 2DEG systems, and it should have spintronic applications in Si-based devices.

DOI: 10.1103/PhysRevApplied.7.014003

I. INTRODUCTION

Silicon takes the unique position in both the conventional mainstream electronic industry (i.e., CMOS) and the emerging fields of quantum-information science and technology such as spintronics [1–3] and quantum computation [4,5]. Its strength derives from the matured capacity of extremely high-purity and low-cost material growth and, perhaps more crucially, the orders of magnitude tunability in electrical conductivity enabled by doping and gating. Silicon (Si) has continued to reveal another crucial property, that is, its long spin lifetime due to the relatively small atomic spin-orbit coupling (SOC), bulk inversion symmetry, and zero nuclear spin in the abundant isotope ^{28}Si . Two-dimensional electron gas (2DEG) occupying the few lowest quantized 2D subbands of various Si surfaces and quantum wells has long been an important playground for fundamental science [6,7] and, more recently, quantum-computing qubit platforms through gate-defined or donor-defined quantum dots [5]. In particular, the long spin lifetime and the ability to control the confined electrons through externally applied electrical voltage (i.e., fast gates) near the Si surface are the main drivers of the great interest and activity in Si-based spintronics and quantum-computing architectures.

With respect to spin relaxation, there are some key differences between 3D and 2D Si systems that are worth emphasizing at the outset. The confining

potential at the interface, where the 2DEG resides, may break the inversion symmetry of the Si crystal. This broken symmetry generally results in spin splitting in the band structure and induces an effective momentum-dependent magnetic field for conduction electrons [8], often referred to as the Rashba field arising purely from the structural asymmetry in real space. In addition, quantum wells with an odd number of Si layers or broken rotoinversion symmetry at the Si-Ge interface [9–12] induce a generalized Dresselhaus field [13,14].

Such structural SOC effects obviously are not present inside the 3D bulk Si and can exist only in the 2DEG. When electrons undergo momentum scattering (e.g., by impurities or phonons) in the presence of Rashba-Dresselhaus effect, their spins precess randomly over time and relax [15]. This D'yakonov-Perel' (DP) process has been the only main spin-relaxation mechanism studied to date in Si 2DEG [16–20]. This limited study has led to the general belief, questioned in this work, that the DP mechanism is the only spin-relaxing mechanism in Si 2DEG that needs to be considered theoretically.

In this work, we bring in a fundamentally different spin-relaxation mechanism which may gradually dominate over the DP mechanism with increasing impurity densities or doping. This alternative impurity-induced spin-relaxation mechanism does not rely on the effective Rashba-Dresselhaus magnetic field between scattering events, but rather flip spins right at the scattering events, through the contact spin-orbit interaction at the impurity core.

*ysong128@umd.edu

Therefore, our present mechanism can be termed a *Yafet* process [21].

While the scattering is driven by impurities in both mechanisms at low temperatures, the difference is that scattering serves to interrupt the spin precession in the DP process, whereas it facilitates spin flip in the Yafet process. As a result, instead of weakening with a higher impurity density or lower mobility as in the DP spin relaxation, our mechanism grows stronger with an increasing (decreasing) impurity density (mobility), and it therefore can be distinguished experimentally from the DP process.

We believe that some contributions of this alternative impurity-induced Yafet process to Si 2D spin relaxation may have already been detected experimentally, as we discuss later in this article. We note that in 3D Si, our process is already known to be the experimentally dominant electronic spin-relaxation mechanism in high-doping situations [22].

The other aspect we highlight is the tunability of the spin lifetime. While the charge transistor builds on the tunable conductivity (i.e., the tunable-carrier momentum-relaxation time), it is desirable that the spin-relaxation time can be controlled as well for spintronic applications. As we show in detail, the leading-order spin flip occurs only during intervalley scattering and between two nonopposing valleys among the six Si conduction valleys (the so-called *f* process [23]), whereas it vanishes during intravalley scattering or intervalley scattering between two opposing valleys (“*g* process”). For various 2DEG plane orientations relative to the Si crystallographic direction (e.g., 100, 110, 111, or arbitrary orientation), it is well known that the resulting 2D electronic ground states have different valley configurations, with the ground-state valley degeneracy varying from 1 to 6 depending on surface orientations and details [6,24]. As such, the spin lifetime determined by this mechanism in 2DEG will be distinguishable owing solely to different plane orientations of the 2D system, producing substantial anisotropy in the 2D spin relaxation. Moreover, the tunability of the relative valley energies by stress and, especially, by gate voltage in Si MOSFETs, can enable fast on-chip spin-lifetime control. Spin-orientation dependence of the spin relaxation, absent for the charge mobility, can also be similarly controlled. Since valleys do not play a central role in the DP process, which is governed entirely by the structural asymmetry, such orientation or gate-voltage dependence of spin relaxation is qualitatively different in the DP mechanism [17,18]. This difference can also distinguish our proposed mechanism from the DP mechanism with respect to Si 2D spin relaxation.

In 3D bulk Si, as we mentioned, this alternative Yafet process has been shown to be the dominant spin-relaxation mechanism when the scattering is caused by donor

impurities [22]. It is caused by the spin-dependent interaction with the impurity core, and it is far more important than the spin flip during intravalley scattering by the long-range Coulomb interaction, or during intervalley scattering by the spin-independent part of the impurity-core potential, neither of which exhibits the empirically strong donor dependence [25–29]. Our goal here is to introduce this important mechanism into the Si 2DEG, build up its primary trend qualitatively and quantitatively, and discuss its experimental relevance and applications. Since the DP process and our process are completely independent spin-relaxation mechanisms, generically, both should be present in Si 2DEG, and their relative quantitative importance will depend on all the details of the specific system and the samples being studied.

We briefly discuss the scenarios where our proposed mechanism can be of importance and can be utilized. The first apparent criterion is low temperature and moderate-to-high impurity density, so phonon-driven spin relaxation is relatively weak. Conversely, our mechanism is most likely overshadowed by the DP spin-relaxation process in intrinsic to low-doped 2DEG systems, typical for Si/SiGe quantum wells and other modulation-doped heterostructures where interface impurity scattering is relatively weak [16,30], except for symmetrically designed wells [9,31] where the structural asymmetry can be reduced. It can be easily shown that DP spin relaxation alone leads to a rapidly diverging spin lifetime once the mobility μ is lowered to a few m^2/Vs [18]. Also, as the present mechanism relies on the SOC between the free electron and the impurity core, for a given impurity density, its quantitative effect is ranked according to the sign of impurity charges: positively charge impurities $>$ neutral impurities \gg negatively charged impurities, the last of which repel the electrons and render little central-cell correction [32–34]. As such, both an *n*-type SiGe quantum well and the accumulation layer in an *n*-type MOSFET are good candidates for study with our mechanism. In low-mobility Si MOSFET samples, however, inversion and accumulation layers can both be relevant due to the dominant interface oxide charges [6,35], which could scatter carriers strongly leading to a strengthening of our mechanism. Noticeably, this mechanism is more effective in 2D than in 3D Si, as the former retains ionized donors under most relevant experimental conditions [36].

In Sec. II, we develop our basic theory of 2D spin relaxation, obtaining detailed results for the relaxation time for different surface orientations and applied external stress in Sec. III. Section IV is devoted to a discussion of our results in the context of experimental implications and the existing 2D spin-relaxation experiments. We conclude in Sec. V with a summary and our outlook. The intervalley-scattering physics and its relevant symmetry analysis is reviewed in the Appendix.

II. THEORETICAL FORMULATION

Our spin-relaxation mechanism arises directly from the impurity SOC, in contrast to the structural SOC effects that emerge from a combination of atomic SOC and broken structural symmetry. The electron spins flip upon scattering. Moreover, the spin flip is governed not by the spin mixing in the conduction electron states but by the SOC of the scattering potential. Our approach to formulating the scattering of electron states in the 2DEG subbands by an impurity potential is to take a doubly applied effective-mass approximation (EMA) [37–39]. One of the EMAs is conventionally applied with the envelope functions of subband states confined in a quantum well or near the surface [6,40] (as discussed in more detail below, we do not consider opposite-valley coupling due to the interface in our leading-order theory). Assuming that the scattering-matrix element for conduction states in a 3D bulk Si of volume V is $U_{v_1, s; v_2, -s}^{3D}$ between valley v_1 , spin s and valley v_2 , spin $-s$, the EMA connects it to that of the 2DEG with an area S ,

$$U_{v_1, n_1; v_2, n_2}^{2D}(z, \mathbf{s}) = \frac{\xi_{v_1, n_1}(z)\xi_{v_2, n_2}(z)}{S} V U_{v_1, s; v_2, -s}^{3D}, \quad (1)$$

for a given impurity located at z along the width direction of the 2DEG, where $\xi_{v, n}(z)$ is the envelope function in the valley v and the quantized 2D subband n , and the normalized $\int dz \xi_{v, n} \xi_{v, m} = \delta_{n, m}$. In the following, we first elaborate on the physics of the spin-flip matrix element $U_{v_1, s; v_2, -s}^{3D}$, where another use of the EMA is crucial to relate a scattering problem with a donor-state problem. Then we study in detail the different specific confinements and the resulting subbands.

The bulk spin-flip scattering is treated rigorously in terms of the general symmetry of the impurity potential [22]. Since the initial and final conduction states are the eigenstates of the bulk Si (one-body) Hamiltonian V_0 , the scattering potential is the difference between the substitutional impurity and the original Si atom, $U = V_{\text{imp}} - V_{\text{Si}}$, and breaks the O_h^7 space symmetry of the diamond-lattice structure. Without going into the details about U (including the screened Coulomb potential and short-range central-cell correction), U obeys the tetrahedral T_d point-group symmetry [41] and one can derive the matrix-element form ($U_{v_1, s; v_2, -s}^{3D} = \langle \psi_{v_2, -s} | U | \psi_{v_1, s} \rangle$) with the correct dependence on valleys and spin orientation of the involved conduction states [22]. We summarize the relevant intervalley scattering in bulk Si and its symmetry analysis in the Appendix. Between conduction states at the valley centers, it turns out that spin flip survives only in intervalley f -process scattering. Its counterparts in intravalley and intervalley g -process scattering are forbidden by the C_2 rotation symmetry of the T_d group and the time-reversal symmetry, respectively. We denote the bare spin angular dependence of $U_{v_1, s; v_2, -s}^{3D}$ without the quantitative prefactor as $\hat{U}_{v_1, s; v_2, -s}$, and the

expression between the $+x$ and $+y$ valleys for the arbitrary spin orientation $\mathbf{s} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta) \equiv (s_x, s_y, s_z)$ reads [22]

$$\begin{aligned} \hat{U}_{+x, s; +y, -s} &= \frac{i}{6} \sin \theta e^{i\phi} \\ &+ \frac{\eta(1-i)}{2\sqrt{3}} \left(\cos^2 \frac{\theta}{2} - i \sin^2 \frac{\theta}{2} e^{2i\phi} \right) \\ &\equiv \frac{is_x - s_y}{6} + \frac{\eta(1-i)}{4\sqrt{3}} \\ &\times \left(1 + s_z - \frac{i(s_x + is_y)^2}{1 + s_z} \right), \end{aligned} \quad (2)$$

where the dimensionless constant η is the ratio between the two symmetry-allowed terms (that is, from the \bar{F} -symmetry states of the T_d group; see the Appendix for details). This expression leads to the anisotropic dependence of spin relaxation on spin orientation.

In order to determine the magnitude of the prefactor in U^{3D} , we make an important connection between it and the spin-split spectrum of the localized impurity states using the essence of EMAs. By comparing the scattering problem and the localized eigenenergy problem of the same impurity, one can realize that the potential is exactly the same for the two problems and that the only difference between the localized state and the conduction state comes from the envelope function in the former due to the Coulomb confinement. As a result, the prefactor in U^{3D} can be related to the spin splitting Δ_{so} of the bound impurity states such that $U_{v_1, s; v_2, -s}^{3D} = (\pi a_B^3 / V) \Delta_{\text{so}} \hat{U}_{v_1, s; v_2, -s}$, where a_B is the impurity Bohr radius and V the bulk volume, an EMA effect not too different from that of Eq. (1) applied for the 2DEG confinement. When the experimental spectrum is available for Δ_{so} , such as those in group-V donors [42,43], this method is most efficient and also likely more accurate than numerical calculations that may miss part of the microscopic contributions. The ratio constant η is estimated to be about 2 from the spin-relaxation data in highly doped n -type Si [22]. For other types of substitutional impurities, η is expected to have a value on the order of unity. In principle, an estimate of η can be obtained by first-principles calculations which are beyond the scope of this work. Again, it is highly preferable to make an empirical comparison with the experiments since a precise quantitative calculation of η is essentially impossible theoretically, particularly in the context of spin relaxation in the 2DEG. As discussed in the Introduction, in general, the SOC scattering strength depends on the type of impurities being considered and would, in general, be smaller for neutral and negatively charged impurities.

$\hat{U}_{+x, s; +y, -s}$ in Eq. (2) describes the leading-order-in-wave-vector spin-flip matrix element in one of all 24 paths of the f -process scattering among six Si conduction valleys. In this work, we group $\hat{U}_{v_1, s; v_2, -s}$ into 12 time-reversal (TR)

related pairs, and then connect them to $\hat{U}_{+x,s;+y,-s}$ by specific spatial-symmetry operations in the T_d group:

$$\begin{aligned}
 |\hat{U}_{v_1,s;v_2,-s}| &\stackrel{\text{TR}}{=} |\hat{U}_{-v_2,s;-v_1,-s}|, \\
 |\hat{U}_{x,s;-y,-s}| &\stackrel{C_{2x}}{=} |\hat{U}_{x,s'=(s_x,-s_y,-s_z);y,-s'}|, \\
 |\hat{U}_{y,s;x,-s}| &\stackrel{\sigma_{x-y}}{=} |\hat{U}_{x,s'=(s_x,-s_x,-s_z);y,-s'}|, \\
 |\hat{U}_{-y,s;x,-s}| &\stackrel{S_4}{=} |\hat{U}_{x,s'=(s_y,-s_x,s_z);y,-s'}|, \\
 |\hat{U}_{x,s;z,-s}| &\stackrel{\sigma_{y-z}}{=} |\hat{U}_{x,s'=(s_x,-s_z,-s_y);y,-s'}|, \\
 |\hat{U}_{x,s;-z,-s}| &\stackrel{\sigma_{y+z}}{=} |\hat{U}_{x,s'=(s_x,s_z,s_y);y,-s'}|, \\
 |\hat{U}_{z,s;x,-s}| &\stackrel{C_3}{=} |\hat{U}_{x,s'=(s_z,s_x,s_y);y,-s'}|, \\
 |\hat{U}_{-z,s;x,-s}| &\stackrel{C_3}{=} |\hat{U}_{x,s'=(s_z,s_x,-s_y);y,-s'}|, \\
 |\hat{U}_{z,s;y,-s}| &\stackrel{\sigma_{x-z}}{=} |\hat{U}_{x,s'=(s_z,-s_y,-s_x);y,-s'}|, \\
 |\hat{U}_{-z,s;y,-s}| &\stackrel{\sigma_{x+z}}{=} |\hat{U}_{x,s'=(s_z,-s_y,s_x);y,-s'}|, \\
 |\hat{U}_{y,s;z,-s}| &\stackrel{C_3}{=} |\hat{U}_{x,s'=(s_y,s_z,s_x);y,-s'}|, \\
 |\hat{U}_{y,s;-z,-s}| &\stackrel{C_3}{=} |\hat{U}_{x,s'=(s_y,-s_z,-s_x);y,-s'}|, \quad (3)
 \end{aligned}$$

where the vector subscript of the reflection operator (σ) marks the normal direction of the reflection plane, and C and S denote the usual proper and improper rotations, respectively, with the given axes (the unspecified axis of the C_3 rotation is along one of the cubic body diagonals). These individual $U_{v_1,s;v_2,-s}$ expressions are important in evaluating spin relaxation in Si 2DEG, where valleys are not all equally occupied and can be subsequently summed together. More specifically, the anisotropy of the effective mass and strain may split the energy degeneracy of the six valleys in different ways, but it always keeps the energy the same for the two opposing valleys. This statement is true without including the small effects from SOC and short-wavelength perturbation beyond the EMA, which could induce splitting on the order of 1 meV or less [5,6,44–47]. This splitting effect is negligible compared to the typical Fermi levels and in the context of the leading-order calculation of the spin-relaxation time. We therefore do not include such small interface coupling between opposing valleys in this work. All in all, we are always allowed to group the 24 f -process paths into three parts, (I) $\pm x \leftrightarrow \pm y$, (II) $\pm x \leftrightarrow \pm z$, and (III) $\pm y \leftrightarrow \pm z$, and we sum $|U_{v_1,s;v_2,-s}|^2$ over each group which shares the same electron statistical distribution factor. Utilizing Eqs. (2) and (3), we have

$$\sum_{8 \in \text{I}} |\hat{U}_{v_1,s;v_2,-s}|^2 = \frac{2}{9} [1 - s_z^2 + 3\eta^2(1 + s_z^2)] \equiv \mathcal{S}(s_z), \quad (4)$$

$$\sum_{8 \in \text{II}} |\hat{U}_{v_1,s;v_2,-s}|^2 = \mathcal{S}(s_y), \quad (5)$$

$$\sum_{8 \in \text{III}} |\hat{U}_{v_1,s;v_2,-s}|^2 = \mathcal{S}(s_x). \quad (6)$$

We will see that Eqs. (4)–(6) directly lead to the strong spin angular dependence of the spin relaxation in the (110)- and (001)-oriented 2DEGs, as well as in the (111) 2DEG under external stress in the next section.

Next, we address the specific confinements and subband envelope functions in Eq. (1). Before doing that, we combine Eq. (1) and $U_{v_1,s;v_2,-s}^{3\text{D}} = (\pi a_B^3/V) \Delta_{\text{so}} \hat{U}_{v_1,s;v_2,-s}$ to give

$$U_{v_1,n_1;v_2,n_2}^{2\text{D}}(z, \mathbf{s}) = \frac{\xi_{v_1,n_1}(z) \xi_{v_2,n_2}(z)}{S} \pi a_B^3 \Delta_{\text{so}} \hat{U}_{v_1,s;v_2,-s}. \quad (7)$$

We stress that the EMA suits our problem especially well, even for the relatively narrow 2DEG. As we showed in Ref. [22], the relevant intervalley-spin-scattering potential comes from the core region of the impurities, evidenced by the strong dependence of the spin-relaxation times on the donor species. The overall 2D confinement is much smoother than the impurity-core potential, whose linear dimension is much less than a lattice constant, and Eq. (7) can be safely used for most of the randomly or uniformly distributed impurities in the 2DEG. The 2DEG system is essentially of a 3D nature with respect to the short-range scattering in the immediate impurity-core region since the 2D confinement length scale (approximately 10 nm or larger) is much larger than the atomic core size (about 0.1 nm). In other words, the weak (as to the influence on impurity cores) symmetry-breaking potential from the 2D confinement is taken into account by the mostly slowly varying envelope $\xi_{v,n}(z)$; otherwise, the scattering interaction and the conduction Bloch functions $\psi_{v,s}$ near the impurity-core region are unchanged to this order of perturbation. Thus, the spin-flip selection rules are still dictated by the bulk symmetry, well retained near the impurity-core region. Under this level of approximation, we also neglect any small change in Δ_{so} and η , and in $\hat{U}_{v_1,s;v_2,-s}$, in going from the 3D bulk to the 2DEG.

We define an effective width $d_{v_1,n_1;v_2,n_2}$ for the scattering between the v_1, n_1 and v_2, n_2 states, in terms of the envelope functions in Eq. (7),

$$\frac{1}{d_{v_1,n_1;v_2,n_2}} \equiv \int dz |\xi_{v_1,n_1}(z) \xi_{v_2,n_2}(z)|^2, \quad (8)$$

which, together with the 2DEG area S , yields an effective volume of the 2DEG $S d_{v_1,n_1;v_2,n_2}$ (taking the role of V in 3D bulk) for a given subband transition. Specifically, we choose two representative confinements for the 2DEG. The first one is a square well, corresponding to the typical

2D-heterostructure quantum well (such as SiGe/Si/SiGe). Focusing on the lowest few levels, we approximate the well potential as an infinite barrier for $0 < z < d$ (where d here is the physical well width) and obtain simple analytical solutions,

$$\xi_{v,n}^{\text{sq}}(z) = \sqrt{\frac{2}{d}} \sin \frac{(n+1)\pi z}{d}, \quad (9)$$

where $n = 0, 1, 2, \dots$ denotes various 2D confined subbands. The corresponding energies at the subband bottoms are

$$E_{v,n} = \frac{[\pi \hbar (n+1)]^2}{2m_{z,v} d^2}, \quad (10)$$

where the effective mass $m_{z,v}$ along the z direction depends on the valley v and the 2DEG plane orientation, as we describe in detail below. Note that $E_{v,n}$ is measured here from the bottom of the v th valley, E_v , in the 3D bulk, as opposed to the lowest subband bottom. Different valley bottoms may shift relative to each other upon various stress configurations (either deliberately applied from the outside or present because of intrinsic interface strain). Following these $\xi_{v,n}(z)$'s, we can obtain the effective width [Eq. (8)] for square wells as

$$d_{n_1;n_2}^{\text{sq}} = \frac{d}{1 + \delta_{n_1,n_2}/2}, \quad (11)$$

which is independent of the involved valleys.

The second representative confinement we use for producing numerical results is a triangular-well potential at the interface $V(z) = eFz$ for $z > 0$ and ∞ for $z < 0$ (F is the electric field including a built-in potential gradient). It corresponds approximately to the inversion (accumulation) layer of a hole- (electron-) doped Si MOSFET, when the 2DEG density is smaller than the saturated charge density of depletion layer per area, N_{depl} [6]. We take the inversion layer as an example in Sec. III, while both types of 2DEG layers are treated in Sec. IV with the variational approach. The envelope function in this model is analytically solved [48],

$$\xi_{v,n}^{\text{tr}}(z) = \alpha_{v,n} \text{Ai} \left[\left(\frac{2m_{z,v}eF}{\hbar^2} \right)^{1/3} \left(z - \frac{E_{v,n}}{eF} \right) \right] \theta(z), \quad (12)$$

$$E_{v,n} \approx \left(\frac{\hbar^2}{2m_{z,v}} \right)^{1/3} \left[\frac{3}{2} \pi eF \left(n + \frac{3}{4} \right) \right]^{2/3}, \quad (13)$$

where Ai denotes the Airy function, and $\alpha_{v,n}$ is the normalization factor [$y = \text{Ai}(x)$ satisfies the original Airy equation $y'' - xy = 0$]. $E_{v,n}$ are the asymptotic values for large n 's but fall within 1% of the exact value, even for $n = 0$. The Airy function depicts the oscillation of state

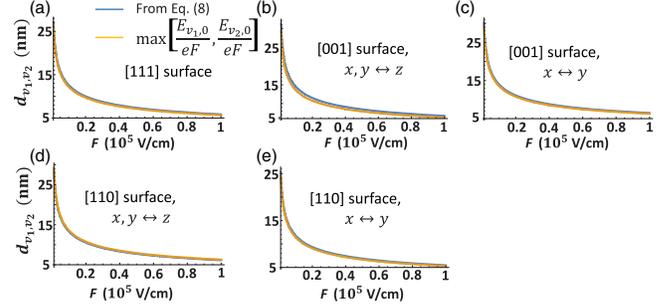


FIG. 1. $d_{v_1,0;v_2,0}$ defined in Eq. (8) (the blue curves) as a function of the effective electrical field F for five representative cases of the triangle-shaped wells: f -process scattering (a) near the [111] surface, (b) between the four- and two-valley groups near the [001] surface, (c) within the four-valley groups near the [001] surface, (d) between the four- and two-valley groups near the [110] surface, and (e) within the four-valley groups near the [110] surface. For comparison, we also plot $\max[E_{v_1,0}, E_{v_2,0}]/eF$ alongside (the yellow curves).

envelopes within the classical turning point ($z_t = E_{v,n}/eF$) and the decay beyond it. The step function $\theta(z)$ [$= 1(0)$, for $z > (<)0$] arises from the one-sided infinite barrier in the model. For the triangular well, $d_{v_1,n_1;v_2,n_2}$ does not have a simple analytical form. In Fig. 1, we plot $d_{v_1,n_1;v_2,n_2}(F)$ in the typical range of electrical field F (10^3 – 10^5 V/cm) numerically in the quantum limit, $n_1 = n_2 = 0$, for all representative surfaces and valley configurations. We find that, for all of these cases, $d_{v_1,0;v_2,0}$ can be well approximated by

$$d_{v_1,0;v_2,0}^{\text{tr}} \approx \max \left[\frac{E_{v_1,0}}{eF}, \frac{E_{v_2,0}}{eF} \right]. \quad (14)$$

The results in both Eq. (11) and Eq. (14), derived from the general definition of the effective width in Eq. (8), can be physically interpreted as follows. First, the volume normalization of the initial and final states scales the scattering matrix element U^{2D} inversely with $\sqrt{d_{v_1,n_1} d_{v_2,n_2}}$, where $d_{v,n}$ is the effective spread of the given subband state in the z direction. Second, the relaxation rate scales with the number of impurities in the overlapping region of the two states, $\propto \min[d_{v_1,n_1}, d_{v_2,n_2}]$. Combining these two factors, the spin relaxation should be roughly proportional to $1/\max[d_{v_1,n_1}, d_{v_2,n_2}]$ [which is, effectively, Eq. (8)]. This deduction is a generic prediction of our theory for impurity-induced 2DEG spin relaxation in Si, which could be directly tested experimentally. Finally, $d_{v,n}$ is basically d for the square well and around the classical turning point $E_{v,n}/eF$ for the triangular well with slope eF .

With the core factors $U_{v_1,s;v_2,-s}$ and $d_{v_1,n_1;v_2,n_2}$ elaborated on, in the following section, we present our calculated spin-relaxation results in all typical Si 2DEG orientations

and stress configurations. By standard time-dependent perturbation theory, one integrates out the periodic time factors of the states resulting in the effective energy conservation in the large time limit, i.e., Fermi's golden rule [49,50], a standard application for relaxation rates,

$$\frac{1}{\tau_s^{2D}(\mathbf{s})} = \frac{4\pi}{\hbar} \left\langle \sum_{v_2, n_2} \int \frac{d^2 k_2}{4\pi^2/S} N_i S \int dz |U_{v_1, n_1; v_2, n_2}^{2D}(z, \mathbf{s})|^2 \delta[\varepsilon_{v_1, n_1}(\mathbf{k}_1) - \varepsilon_{v_2, n_2}(\mathbf{k}_2)] \right\rangle_{\mathbf{k}_1}, \quad (15)$$

where \mathbf{k}_1 and \mathbf{k}_2 are the 2D wave vectors for initial and final states, N_i is the impurity density per volume, and $\langle \mathbf{A} \rangle_{\mathbf{k}_1} \equiv \sum_{v_1, n_1} \int d^2 k_1 \mathbf{A}[\partial \mathcal{F} / \partial \varepsilon_{v_1, n_1}(\mathbf{k}_1)] / \sum_{v_1, n_1} \int d^2 k_1 [\partial \mathcal{F} / \partial \varepsilon_{v_1, n_1}(\mathbf{k}_1)]$ denotes the shortcut for the normalized integration over \mathbf{k}_1 , with \mathcal{F} being the Fermi-Dirac distribution (see, e.g., Ref. [21], p. 73). In our calculation, we neglect the dependence of $U_{v_1, n_1; v_2, n_2}^{2D}$ on the small wave vector measured from its respective valley center (\mathbf{k}_0) [51], which only renders a higher-order relative error [roughly, $|\mathbf{k} - \mathbf{k}_0| / (2\pi/a) \ll 1$, with a being the Si lattice constant], and thus U^{2D} depends only on the valleys and subbands of the involved states. Our leading-order-in-wave-vector theory establishes the first quantitative analysis for impurity-induced 2D spin relaxation beyond those arising from the DP mechanism.

III. NUMERICAL RESULTS FOR DIFFERENT 2DEG ORIENTATIONS AND APPLIED STRESS

A. Without external stress

In the limit of large well width, the number of occupied subbands for a given Fermi level is proportional to the width d , and $1/\tau_s^{2D}$ reduces to the 3D limit independent of d [22], which we have explicitly verified numerically. In this section, we give concrete quantitative results for the opposite 2D limit of only one (the ‘‘quantum limit’’) or the few lowest subbands being populated, in the low-temperature limit of our interest. As we stressed, the EMA is well applied in this limit for our problem where the interaction occurs within the impurity core regions. For more details on the general justification of applying EMA to tightly confined quantum structures, the reader can refer to Ref. [40]. We also note that this theory treats on equal footing the ‘‘weak-field’’ and ‘‘strong-field’’ limits that arise from the study of structural SOC and the oscillation of valley splitting [11,12]. For a square quantum well under a strong electric field, the 2DEG may be modeled in a triangular confinement (or some specific variations) for our mechanism, as exemplified in Sec. II. We remark that, since our spin-flip mechanism draws on f -process intervalley scattering, it is more easily seen for the (111)- and (110)-oriented 2DEGs, where nonopposing valleys coexist in the ground states, than for the (001) one. For the latter case,

multisubband occupation is required in order for our leading-order spin relaxation to play a key role.

The situations with potential confinement but no stress are studied first. τ_s^{2D} as a function of 2DEG electron density N_{2D} for a given well potential $V(z)$, as well as τ_s^{2D} versus $V(z)$ for a given N_{2D} , are computed. Depending on N_{2D} and the subband splitting (or the corresponding well width), different subbands may be populated. Unlike the 3D bulk case, in 2DEG the electron density and the Fermi level are decoupled from the impurity density N_i , as the former can be controlled by the gate voltage. The general relation between N_{2D} and the Fermi level ε_F reads

$$N_{2D} = \frac{1}{2\pi\hbar^2} \sum_v \sum_n \sqrt{m_{1,v} m_{2,v}} (\varepsilon_F - E_{v,n}) \theta(\varepsilon_F - E_{v,n}), \quad (16)$$

where $m_{1,v}$ and $m_{2,v}$ are the in-plane effective masses in the v th valley (the effective mass is anisotropic in Si due to the ellipsoidal forms of the bulk conduction-band minima), and $\theta(x) = 0$ or 1 for $x < 0$ or $x > 0$.

We start with 2DEG of the (111) well orientation. $m_z = 3m_t m_l / (m_t + 2m_l)$ is the same in every valley, so the six energy ‘‘ladders’’ of subbands remain degenerate among different valleys (neglecting any small valley-splitting correction beyond the effective-mass approximation). The spin-relaxation rate, Eq. (15), then becomes

$$\frac{1}{\tau_s^{(111)}(\mathbf{s})} = \frac{\pi^2 a_B^6 \Delta_{so}^2 N_i}{\hbar^3} \mathcal{G}^{(111)}, \quad (17)$$

with the orientation-specific factor assuming the low-temperature limit,

$$\mathcal{G}^{(111)} = \frac{\sum_{n_1, n_2} \frac{4(1+6\eta^2)\sqrt{m_1 m_2}}{9d_{n_1, n_2}} \theta(\varepsilon_F - E_{n_1}) \theta(\varepsilon_F - E_{n_2})}{3 \sum_n \theta(\varepsilon_F - E_n)}, \quad (18)$$

where $m_1 = m_t$ and $m_2 = (m_t + 2m_l)/3$. As in the 3D case, $\tau_s^{(111)}(\mathbf{s})$ is isotropic in spin orientation.

We quantify the spin-relaxation time τ_s for the two basic well types introduced in Sec. II, (1) the infinite square well and (2) the triangular well. For a square well with width d , E_n and d_{n_1, n_2} follow Eqs. (10) and (11), and Eq. (18) reduces to

$$\mathcal{G}_{sq}^{(111)} = \frac{4(1+6\eta^2)\sqrt{m_1 m_2}}{27d} \left[\mathcal{N} \left(\sqrt{\frac{\varepsilon_F}{E_0}} \right) + \frac{1}{2} \right] \theta(\varepsilon_F - E_0), \quad (19)$$

where $\mathcal{N}(x)$ returns the integer part of x . We plot τ_s as a function of ε_F and relate it to the corresponding N_{2D} in Fig. 2 for three different representative well widths, (a) 15 nm, (b) 30 nm, and (c) 45 nm. We also plot τ_s as a function of d for three fixed Fermi levels, $\varepsilon_F = 10, 20$, and 30 meV in Fig. 2(d). Since one can simply scale τ_s with

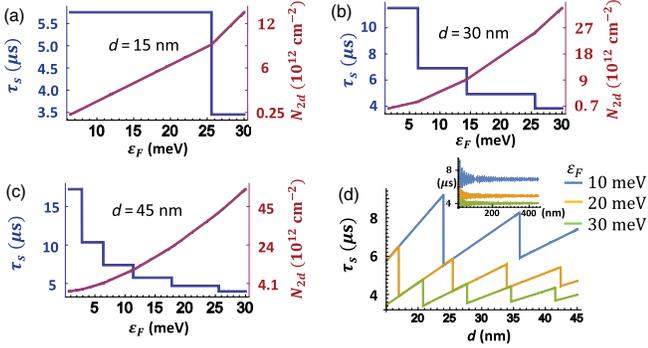


FIG. 2. τ_s in the (111) square well. We plot τ_s as a function of ε_F and the corresponding N_{2D} for three representative well widths d , (a) 15 nm, (b) 30 nm, and (c) 45 nm, and τ_s as a function of d for three fixed Fermi levels, $\varepsilon_F = 10, 20,$ and 30 meV, in (d). The inset of (d) shows the large- d behavior approaching the bulk limit. Note that the energy of the lowest subband bottom is $E_0 = \pi^2 \hbar^2 / 2m_z d^2$ relative to the zero reference energy. $\Delta_{so} = 0.1$ meV and $N_i = 10^{16}$ cm $^{-3}$ are chosen here and for all of the following figures.

Δ_{so}^{-2} and N_i^{-1} , as shown in Eq. (17), we choose the typical parameters $\Delta_{so} = 0.1$ meV and the doping concentration $N_i = 10^{16}$ cm $^{-3}$. The clear kinks in N_{2D} versus ε_F and the jumps in τ_s versus ε_F or d reflect the onset of (de)populating more subbands. τ_s decreases as ε_F or N_{2D} increases for a fixed d since more subbands are available for states at the Fermi level to be scattered into. As d increases towards the bulk limit, denser subbands gradually evolve towards the density of state for the 3D bulk at a given ε_F [see the inset of Fig. 2(d)]. On the other hand, at a small d , τ_s decreases continuously with a decreasing d within the window of a fixed number of occupied subbands, n_{occ} . This general trend is dominated by the volume normalization of the involved state captured in Eq. (8). When n_{occ} reduces by one, the number of available final states decreases and τ_s increases again.

For the triangular well $V(z) = eFz$, with its slope controllable by the gate voltage in an inversion layer, the solution of $\xi_n(z)$ becomes the Airy functions given in Eq. (12), and d_{n_1, n_2}^{\pm} , for $n_1 = n_2 = 0$, can be estimated by Eq. (14). The general changes from the square well are (1) d_{n_1, n_2} in Eq. (17), becoming $n_{1,2}$ dependent, and (2) the different dependence of E_n on n . However, we emphasize that the triangular model is more valid as N_{2D}/N_{depl} decreases (and $N_{2D} < N_{depl}$) [6]. A realistic acceptor density we choose is $N_A = 10^{16}$ cm $^{-3}$. As a result, this situation usually corresponds to one where only the lowest subbands in the inversion layer are occupied. Therefore, the numerical results in Fig. 3 are given in this practical energy window. In this case, τ_s becomes independent of ε_F or N_{2D} due to the constant 2D density of states per subband, and we only need to show the dependence of τ_s on F . For a higher N_{2D} , $V(z)$ is not independent of but is largely determined by N_{2D} .

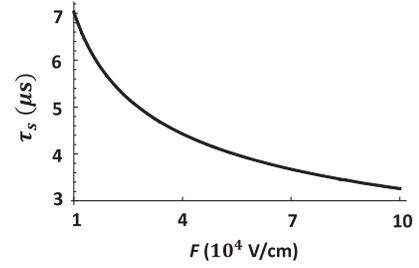


FIG. 3. τ_s in the triangular well $V(z) = eFz$ to the (111) surface for $F \leq 10^5$ V/cm. This value is at the quantum limit where only the lowest subbands are occupied. In this limit, τ_s is independent of ε_F and N_{2D} . We choose the realistic acceptor density in the depletion as well as in the inversion layer, $N_i = N_A = 10^{16}$ cm $^{-3}$.

Next, we study the (001) well orientation, where, additionally, we have the relative shift between different subbands belonging to the two-valley group ($\pm z$ valleys in the 3D limit) and the four-valley group ($\pm x$ and $\pm y$ valleys in the 3D limit). The subband edges in these two groups are determined by the different m_z 's, $m_z = m_l(m_t)$ for the two- (four-) valley group. This symmetry breaking between the six otherwise equivalent ladders of subbands results in spin-orientation dependence, absent in the (111) well case. In the quantum limit, only the two-valley group is occupied and the spin relaxation due to impurities vanishes in the leading order. The general spin-relaxation rate follows Eq. (17), with $\mathcal{G}^{(111)}$ replaced by

$$\begin{aligned} \mathcal{G}^{(001)}(\mathbf{s}) = & \sum_{n_1, n_2} \theta(\varepsilon_F - E_{x, n_2}) \left\{ \frac{\mathcal{S}(s_z) \sqrt{m_l m_t} \theta(\varepsilon_F - E_{x, n_1})}{d_{x, n_1; y, n_2}} \right. \\ & \left. + \frac{\left[\frac{4}{9} (1 + 6\eta^2) - \mathcal{S}(s_z) \right] m_t \theta(\varepsilon_F - E_{z, n_1})}{d_{x, n_1; z, n_2}} \right\} \\ & / \sum_n \left[\sqrt{\frac{m_l}{m_t}} \theta(\varepsilon_F - E_{z, n}) + 2\theta(\varepsilon_F - E_{x, n}) \right] \end{aligned} \quad (20)$$

after utilizing the spin-orientation form factors in Eqs. (4)–(6). We use the anisotropic in-plane effective masses: $m_1 = m_2 = m_l$ for the two-valley group, and $m_1 = m_t, m_2 = m_l$ for the four-valley group. For a square well with a width d , we have

$$\begin{aligned} \mathcal{G}_{sq}^{(001)}(\mathbf{s}) = & \frac{\theta(\varepsilon_F - E_{x,0})}{d} \left[\mathcal{N} \left(\sqrt{\frac{\varepsilon_F}{E_{x,0}}} \right) + \frac{1}{2} \right] \\ & \times \left\{ \mathcal{S}(s_z) \sqrt{m_l m_t} \mathcal{N} \left(\sqrt{\frac{\varepsilon_F}{E_{x,0}}} \right) \right. \\ & \left. + \left[\frac{4}{9} (1 + 6\eta^2) - \mathcal{S}(s_z) \right] m_t \mathcal{N} \left(\sqrt{\frac{\varepsilon_F}{E_{z,0}}} \right) \right\} \\ & / \left[\sqrt{\frac{m_l}{m_t}} \mathcal{N} \left(\sqrt{\frac{\varepsilon_F}{E_{z,0}}} \right) + 2\mathcal{N} \left(\sqrt{\frac{\varepsilon_F}{E_{x,0}}} \right) \right]. \end{aligned} \quad (21)$$

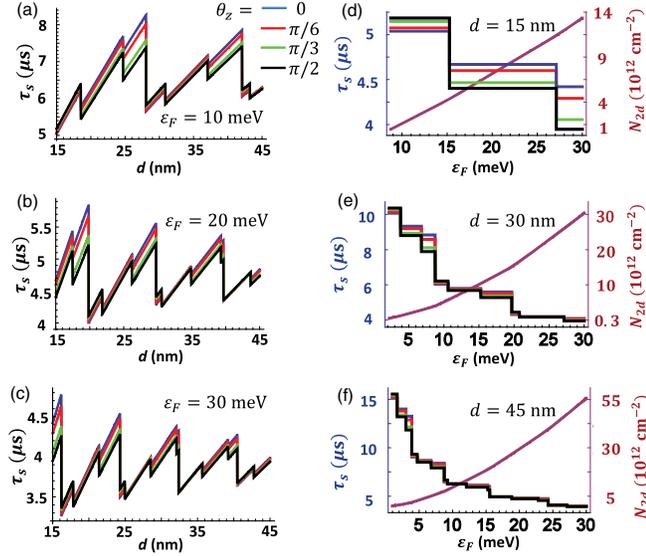


FIG. 4. $\tau_s(\mathbf{s})$ in the (001) square well, anisotropic in spin orientation \mathbf{s} . τ_s depends on the polar angle of \mathbf{s} , θ_z , with respect to the well normal, z . In (a)–(c), we vary the well width d from 15 to 45 nm for three Fermi levels, $\varepsilon_F =$ (a) 10, (b) 20, and (c) 30 meV. In (d)–(f), we plot $\tau_s(\mathbf{s})$ as a function of ε_F , $E_{x,0} < \varepsilon_F < 30$ meV (and also a function of N_{2D} by relating ε_F to N_{2D}), for three different well widths, $d =$ (d) 15, (e) 30 and (f) 45 nm. In each subplot, we detail four spin polar angles: $\theta_z = 0, \pi/6, \pi/3$, and $\pi/2$.

Plots of $\tau_s(\mathbf{s})$ with $\mathcal{G}_{\text{sq}}^{(001)}(\mathbf{s})$ in Eq. (21) as functions of d and ε_F (and the corresponding N_{2D}) are given in Fig. 4. An apparently new feature is the generally smaller steps in comparison with the (111) well results, resulting from the consecutive fillings of the two-valley subbands which have smaller interband splitting. As previously mentioned, the important consequence of the inequivalency between the two- and four-valley groups is the spin-orientation dependence of the spin lifetime. The anisotropy is the strongest when the occupied states in the two groups differ the most, which happens right before one more four-valley subband begins to be filled. We must note that $n_{\text{occ},z} > n_{\text{occ},x}$ is not sufficient on its own to guarantee spin anisotropy; rather, it is necessary that $n_{\text{occ},z}/n_{\text{occ},x} > \sqrt{m_{1,x}m_{2,x}/m_{1,z}m_{2,z}} = \sqrt{m_l/m_t}$, based on Eq. (21). This value is most appreciable preceding the filing of the second subband in the four-valley group, $n_{\text{occ},z}/n_{\text{occ},x} = 4:1$. In the large ε_F (or large d) limit, on the other hand, $n_{\text{occ},z}/n_{\text{occ},x} \rightarrow \sqrt{m_{z,z}/m_{z,x}} = \sqrt{m_l/m_t} \approx 2.27$. We see that this ratio exactly cancels out the effective-mass difference in the 2D (x - y) plane, owing to the fact that $m_1m_2m_z \equiv m_l m_t^2$ is orientation independent for a given ellipsoid.

For the triangular well, Eqs. (8), (12), and (13) are substituted into Eq. (20), and $\tau_s(\mathbf{s})$ is shown in Fig. 5. After the onset of f -process scattering ($\varepsilon_F > E_{x,0}$) follows the second subband of the two-valley group, as shown clearly in Figs. 5(c) and 5(d). $(E_{z,1} - E_{x,0})/E_{x,0}$ is a small fixed

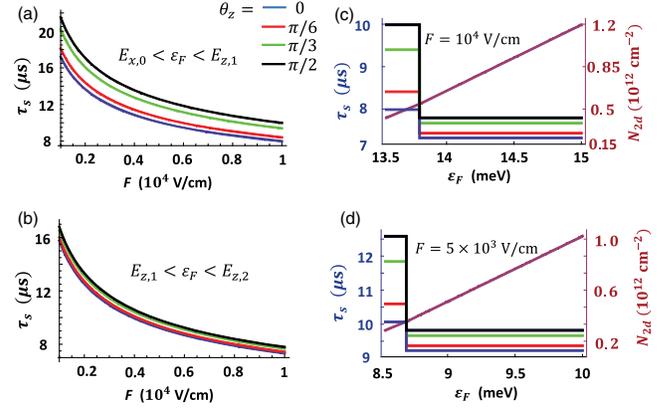


FIG. 5. $\tau_s(\mathbf{s})$ in the triangular well $V(z) = eFz$ to the (001) surface. In (a) and (b), $\tau_s(\mathbf{s})$ is plotted as a function of the electric field $F < 10^4$ V/cm for the Fermi level ε_F (a) just above the onset of f -process scattering ($E_{x,0} < \varepsilon_F < E_{z,1}$) and (b) in the next energy window ($E_{z,1} < \varepsilon_F < E_{z,2}$), where $\tau_s(\mathbf{s})$ is independent of ε_F . In (c) and (d), $\tau_s(\mathbf{s})$ is plotted as a function of ε_F and N_{2D} for a range $E_{x,0} < \varepsilon_F < E_{z,2}$, at the electric field $F =$ (c) 10^4 and (d) 5×10^3 V/cm. At $F = 10^4$ V/cm, the onset of f -process scattering already requires $N_{2D} \approx 5 \times 10^{11}$ cm $^{-2}$, a value comparable to the typical depletion-layer impurity density $N_{\text{depl}} = \sqrt{2E_G\kappa N_A}/e^2$ [6], where E_G and κ are the Si band gap and permittivity ($N_{\text{depl}} \sim 4 \times 10^{11}$ cm $^{-2}$ at $N_A = 10^{16}$ cm $^{-3}$).

ratio for any electric field F , according to Eq. (13). Since we work in the regime where the triangular-well model is valid and multiple subbands are occupied, ε_F and F should not be too large. As a result, we focus on two energy intervals, $E_{x,0} < \varepsilon_F < E_{z,1}$ and $E_{z,1} < \varepsilon_F < E_{z,2}$, below $F = 10^4$ V/cm (in practice, our mechanism works well under a higher electrical field, as long as the Fermi level can reach the x and y valleys and the triangular-well approximation is relaxed).

For scattering involving subband $n = 1$, we obtain $d_{z,1;x,0} \approx 1.15E_{z,1}/eF$ using Eqs. (8) and (12). An interesting behavior is the large anisotropy (approximately 20%) of $\tau_s(\mathbf{s})$ right at the onset of the f -process scattering [Figs. 5(a), 5(c), and 5(d)], which drops (approximately 5%) at the second energy window [cf. Fig. 5(b)]. In the first energy window, $E_{x,0} < \varepsilon_F < E_{z,1}$, the number of occupied subbands in each valley is one for both the two- and four-valley groups. The large anisotropy here is the sole consequence of effective-mass anisotropy in the 2D plane and is opposite in sign to that in the square well before filling $E_{x,1}$.

The change of anisotropy may be sharply *tuned by the gate voltage* in the Si inversion layer. Since $E_{z,1} - E_{x,0}$ and $\varepsilon_F - E_{x,0}$ depend on the electric field F , by simply tuning F , the chemical potential (i.e., the Fermi energy) can cross $E_{z,1}$ and, therefore, induce a switch between Figs. 5(a) and 5(b). An even more important application of our results may be the sharply *gate-voltage-modulated* spin lifetime in

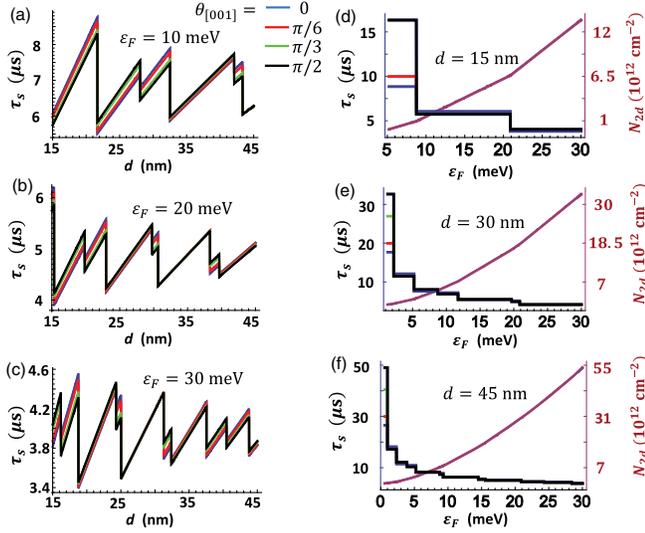


FIG. 6. τ_s (s) in the (110) square well, anisotropic in spin orientation \mathbf{s} . τ_s depends on the polar angle of \mathbf{s} , $\theta_{[001]}$, with respect to the [001] crystallographic direction. In (a)–(c), we vary the well width d from 15 to 45 nm for three Fermi levels, $\varepsilon_F =$ (a) 10, (b) 20, and (c) 30 meV. In (d)–(f), we plot τ_s (s) as a function of $\varepsilon_F < 30$ meV and N_{2D} , for three different well widths, $d =$ (d) 15, (e) 30, and (f) 45 nm.

the Si inversion layer following similar reasoning: the crossover of the chemical potential to $E_{x,0}$, which is the threshold of finite leading-order spin relaxation, could be achieved solely by the top gate voltage. This gate voltage is the on-chip real-time electrical switch for a substantial change of the spin lifetime. We emphasize that both gate-voltage modulations envisioned above should be very robust in general Si inversion layers, not relying on the triangular-well approximation we used in producing Fig. 5.

Last, we study the case of the (110) well. The difference of the (110) well from the (001) one lies in the effective masses, which result in a quantitative difference in the valley splitting, the subband splitting, and the density of states, all playing roles in determining τ_s . For the two-valley group, $m_z = m_t$, $m_1 = m_t$, and $m_2 = m_l$; for the four-valley group, $m_z = 2m_l m_t / (m_t + m_l)$, $m_1 = m_t$, and $m_2 = (m_t + m_l)/2$. The most important distinction is in the quantum limit, where the lowest subbands in the (110) case are from the four-valley group, rather than from the two-valley group in the (001) case. Consequently, the spin lifetime due to the leading-order spin relaxation is finite even for the lowest electron density in the (110) well.

We show the key results for the (110) square well in Fig. 6 and the triangular well in Fig. 7. We use similar parameters (Fermi level ε_F , well width d , electric field F) as those in the (001) case, so that we can focus on the differences between the (110) and (001) wells. First, τ_s is finite in the (110) 2DEG, even for the lowest ε_F in Figs. 6(d)–6(f) and Figs. 7(c) and 7(d), as mentioned above. Owing to the smaller number of available f -process paths

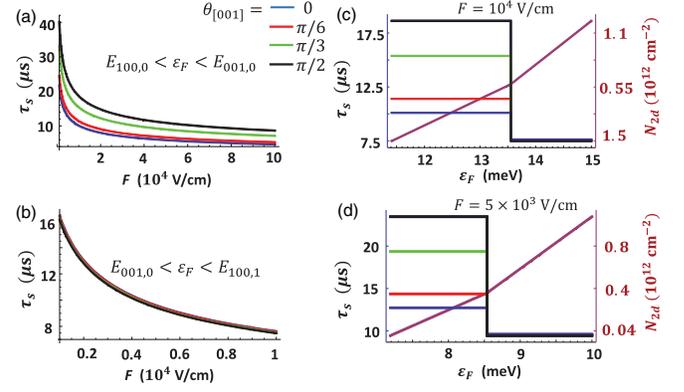


FIG. 7. τ_s (s) in the triangular well $V(z) = eFz$ to the (110) surface. In (a) and (b), τ_s is plotted as a function of the electric field F for the Fermi level ε_F (a) in the lowest energy interval ($E_{100,0} < \varepsilon_F < E_{001,0}$) with $F < 10^5$ V/cm, and (b) in the next energy window ($E_{001,0} < \varepsilon_F < E_{100,1}$), with $F < 10^4$ V/cm, where τ_s (s) is independent of ε_F . We work in a smaller field in (b) for the same reason in Figs. 5(a) and 5(b). In (c) and (d), τ_s is plotted as a function of ε_F and N_{2D} for a range where $\varepsilon_F < E_{100,1}$, at an electric field $F =$ (c) 10^4 and (d) 5×10^3 V/cm.

(two as opposed to four for each state), though, τ_s is larger in the (110) case than the longest finite τ_s in the (001) case. This finding is a clearly verifiable sharp prediction of our theory. Note that here the plane normal (z) is along the crystallographic direction [110] (not [001]). To avoid ambiguity, we use the scripts 001 and 100 rather than z and x to denote directions.

The most significant feature in this quantum limit for the (110) wells is the nearly 50% variation of τ_s (s) on spin orientation (more specifically, on the projection of \mathbf{s} along the [001] crystallographic direction). This extreme case represents only one type of f -process scattering [Eq. (4)], with zero weight from the other two. This feature is clearly seen on the left-hand sides of Figs. 6(d)–6(f) and Figs. 7(c) and 7(d), as well as the entire range in Fig. 7(a). Therefore, the idea of the gate-tuned anisotropic τ_s (s), discussed in the context of (001) wells, is even more prominent in (110) wells. Note that this dependence on the polar angle around the [001] direction, $\theta_{[001]}$, is opposite in sign to that of the (001) square well when the anisotropy is the strongest and is the same as that in the (001) triangular well. The anisotropy and the orientation dependence of 2D Si spin relaxation arising in the impurity-induced spin flip is an important prediction of our theory.

B. With external stress

When external stress is applied, different ladders of subbands may undergo relative shift with each other [52]. On the other hand, no leading-order effect comes from the slight variation of interband splitting within each ladder, as the effective masses are fixed under stress to the leading

order [52]. We study τ_s individually for all three well orientations under uniaxial stress, as described below.

(111) well.—Out-of-plane [111] stress does not induce additional symmetry breaking or shift subband ladders, while the in-plane stress in the $[1\bar{1}0]$ ($\bar{1} \equiv -1$) or $[1 - \sqrt{3}, 1 + \sqrt{3}, -2]$ direction does. These latter two stress directions are in the plane of the 2D well and are realizable experimentally [24]. The $[1\bar{1}0]$ (or $[11\bar{2}]$) stress results in a relative shift $\Delta_V = E_{001} - E_{100}$ between the two- and four-valley groups, and the $[1 - \sqrt{3}, 1 + \sqrt{3}, -2]$ one separates the valleys into three groups, such that $\Delta_V = E_{010} - E_{001} = E_{001} - E_{100}$.

Under the $[1\bar{1}0]$ stress, the form factor \mathcal{G} changes from Eq. (18) for the unstrained (111) well to

$$\begin{aligned} \mathcal{G}_{1\bar{1}0}^{(111)}(\mathbf{s}) &= \sqrt{\frac{m_t(m_t + 2m_l)}{3}} \\ &\times \sum_{n_1, n_2} \frac{\theta(\varepsilon_F - E_{100, n_2})}{d_{n_1, n_2}} \left\{ \mathcal{S}(s_z) \theta(\varepsilon_F - E_{100, n_1}) \right. \\ &+ \left. \left[\frac{4}{9} (1 + 6\eta^2) - \mathcal{S}(s_z) \right] \theta(\varepsilon_F - E_{100, n_1} - \Delta_V) \right\} \\ &/ \sum_n [\theta(\varepsilon_F - E_{100, n} - \Delta_V) + 2\theta(\varepsilon_F - E_{100, n})]. \end{aligned} \quad (22)$$

Focusing on the near quantum limit with only the $n = 0$ subbands occupied, $d_{0,0} = 2d/3$ for the square well and about $(9\pi\hbar)^{2/3}/[4(2m_{111}eF)^{1/3}]$ for the triangular well, from Eqs. (11) and (14), respectively. Equation (22) has only a few discrete outcomes for a given well width or electric field: (1) when $\varepsilon_F > \{E_{100,0}, E_{100,0} + \Delta_V\}$, lines 2–4 of Eq. (22) reduce to $4(1 + 6\eta^2)/27d_{0,0}$, and τ_s recovers the no-strain result (Figs. 2 and 3); (2) when $E_{100,0} < \varepsilon_F < \Delta_V + E_{100,0}$, the same factor decreases to $[1 - s_z^2 + 3\eta^2(1 + s_z^2)]/9d_{0,0}$, with a strong s_z dependence; and (3) when $\Delta_V + E_{100,0} < \varepsilon_F < E_{100,0}$, $1/\tau_s = 0$.

Under the $[1 - \sqrt{3}, 1 + \sqrt{3}, -2](\equiv \gamma)$ stress, the three groups of f -process scattering vary independently. Utilizing Eqs. (4)–(6), the form factor $\mathcal{G}_\gamma^{(111)}$ reads

$$\begin{aligned} \mathcal{G}_\gamma^{(111)}(\mathbf{s}) &= \sqrt{\frac{m_t(m_t + 2m_l)}{3}} \\ &\times \frac{\sum_{n_1, n_2, i} \frac{\mathcal{S}(s_i)}{d_{n_1, n_2}} \theta(\varepsilon_F - E_{i+1, n_1}) \theta(\varepsilon_F - E_{i+2, n_2})}{\sum_{n, i} \theta(\varepsilon_F - E_{i, n})}, \end{aligned} \quad (23)$$

where i denotes the three cyclic directions 100, 010, and 001, $E_{010, n} = E_{001, n} + \Delta_V$, and $E_{100, n} = E_{001, n} - \Delta_V$. Although the three f -process groups depend on spin projection along different directions, shown in

Eqs. (4)–(6), $\tau_s(\mathbf{s})$ can be associated with a fixed projection direction in each energy window of ε_F , thanks to the constant density of state per subband. Focusing on the $n = 0$ limit, the second line in Eq. (23) has several possible outcomes: (1) $4(1 + 6\eta^2)/27d_{0,0}$ when $\varepsilon_F > \{E_{100,0}, E_{010,0}, E_{001,0}\}$, recovering the no-strain result; (2) $\mathcal{S}(s_{100})/2d_{0,0}$ when $\{E_{010,0}, E_{001,0}\} < \varepsilon_F < E_{100,0}$; (3) $\mathcal{S}(s_{010})/2d_{0,0}$ when $\{E_{100,0}, E_{001,0}\} < \varepsilon_F < E_{010,0}$; and, finally, (4) 0 when $E_{100,0} < \varepsilon_F < \{E_{010,0}, E_{001,0}\}$ or $E_{010,0} < \varepsilon_F < \{E_{100,0}, E_{001,0}\}$.

(001) well.—Out-of-plane [001] stress keeps the two- and four-valley degeneracies and tunes the energy distance $E_z - E_x$ between them. In-plane [100] (or [010]) stress breaks the four-valley degeneracy into two groups and tunes $E_x - E_y$, while keeping the splitting $E_{y(x)} - E_z > 0$ unchanged.

Under the [001] stress, the spin relaxation keeps the unstrained form in Eq. (20) with $E_z - E_x$ tunable. In the $n = 0$ limit, $d_{v_1,0;v_2,0} = 2d/3$ for the square well, and about $(9\pi\hbar)^{2/3}/\{4(2\min[m_{001,v_1}, m_{001,v_2}]eF)^{1/3}\}$ for the triangular well. In this limit, we have the additional possibility that $E_{x,0} < \varepsilon_F < E_{z,0}$, which results in a strong s -anisotropic spin lifetime.

Under the [100] (or [010]) stress, one may have f -process scattering available between the $n = 0$ subbands by pushing $x(y)$ valleys lower towards z valleys and $y(x)$ valleys farther away. This action may require a relatively large compressive stress for narrow wells. In this situation, the form factor goes to $\mathcal{G}_{100}^{(001)}(\mathbf{s}) = \mathcal{S}(s_y)m_t/(1 + \sqrt{m_t/m_l})$ and $s_y \leftrightarrow s_x$ for $\mathcal{G}_{010}^{(001)}(\mathbf{s})$.

(110) well.—The unique feature for a (110) well under stress is that the two- and four-valley groups remain degenerate for, respectively, the out-of-plane [110] stress and all of the in-plane stress directions, including [001] and $[\bar{1}10]$. Therefore, the spin-relaxation rate expression follows the unstrained one, except with the energy distance between E_z and E_x being tunable by stress.

The stress dependence of 2D spin relaxation, arising entirely from the valley-dependent subband structure of the 2DEG, is a characteristic feature of the Yafet impurity process being considered in this work, which is completely absent in the Rashba-Dresselhaus-based DP-relaxation mechanism.

IV. EXPERIMENTAL IMPLICATIONS

In this section, we discuss the potential experimental implications of the impurity-driven spin-relaxation mechanism. Variational calculations for the 2DEG subband structure are carried out which are valid for a broader range of MOSFET parameters, including both inversion and accumulation layers, and they use variables that are convenient to compare with experiments. Possible experimental proposals are discussed to differentiate major

contributions to the spin relaxation in Si 2DEG. The few existing 2D Si spin-relaxation measurements available in the literature (see our discussion below), in particular, have all been interpreted using the DP relaxation mechanism, although the quantitative agreement between theory and experiment is, in general, not satisfactory.

Far fewer spin-relaxation measurements have been made on the Si 2DEG than on the bulk Si. In n -type bulk Si, with the inversion-symmetric lattice structure, Elliott-Yafet (EY) spin relaxation is the dominant mechanism for conduction electrons. It is established that the scattering is caused mainly by the electron-phonon interaction at elevated temperatures [21,53–57], or by various processes involving impurities under high impurity density and low temperatures (see, e.g., the review in Ref. [1] and Fig. 2 in Ref. [57]). Despite the weak SOC in Si, the spin lifetime τ_s is only on the order of 10 ns at room temperature without high impurity density, and 0.1–100 ns at low temperatures and high donor concentrations, depending on the specific donor type [22,25–29,58]. By comparison, in 2DEG, several additional features emerge with respect to spin relaxation. Aside from the tunability of spin lifetime and anisotropy showed in Sec. III, the DP spin relaxation becomes relevant, caused by the inversion-breaking structure and interfaces and the associated Rashba-Dresselhaus field. A third feature is the large amount of interface disorder, especially in MOSFETs, which may produce valley-spin-flip scattering. As the 2DEG quantum limit is typically studied at low temperatures, the spin relaxation is determined by scattering with impurities, where EY and DP mechanisms happen, respectively, during and between the scattering events. In the following, we quantify the spin lifetime in the quantum limit based on our studied Yafet mechanism, taking into account the experimentally measured parameters and the uncertainty in the interface disorder.

To make a comparison with the experiments, we take the acceptor concentration N_A and the 2D electron density N_{2D} as two independent variables. Furthermore, we take the effective impurity density N_i in the 2DEG as a separate variable, as it can be significantly different from that in the bulk. To be applicable to a wide range of parameter values and to both inversion and accumulation layers, we relax the triangular approximation and adopt the variational subband wave function [48,59,60]

$$\xi(z) = \sqrt{\frac{b^3}{2}} z \exp\left(-\frac{bz}{2}\right), \quad (24)$$

with a variational variable b . After the numerical energy minimization to find b , taking into account both the depletion- and 2DEG-layer potentials (the band bending takes 1.1 eV and 45 meV for inversion and accumulation, respectively), simple expressions for the effective width $d_{v_1,0,v_2,0}$ [see Eq. (8)] can be obtained in terms of b :

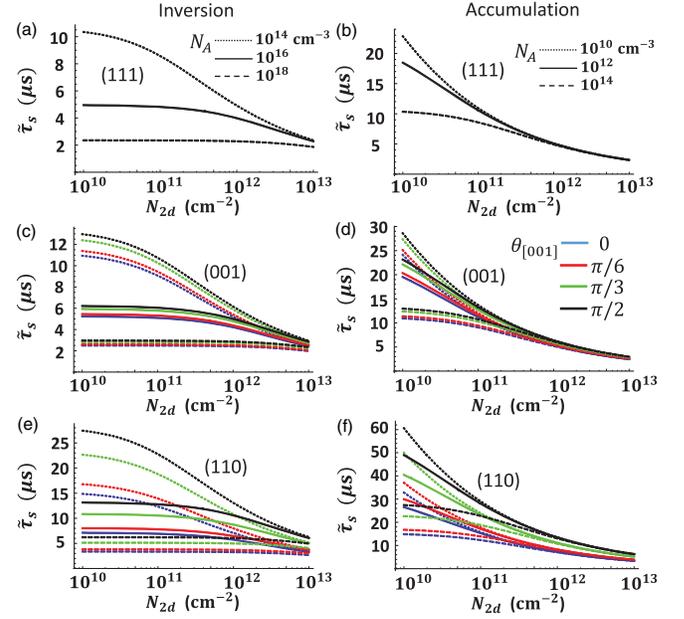


FIG. 8. Normalized spin-relaxation time independent of impurity density, $\tilde{\tau}_s = (N_i/10^{16} \text{ cm}^{-3})\tau_s$, for an inversion [(a), (c), and (e)] or an accumulation [(b), (d), and (f)] layer in the Si MOSFET setup. All three 2DEG plane orientations are shown: (111) in (a) and (b), (001) in (c) and (d), and (110) in (e) and (f). Only the lowest subbands are occupied [for the (001) case, the lowest subbands in the four higher valleys are also occupied]. They are plotted as functions of 2D electron density, N_{2D} , and for acceptor density (related to potential from the depletion layer), $N_A = 10^{14}, 10^{16}, 10^{18}$ for the inversion-layer and $10^{10}, 10^{12}, 10^{14}$ for the accumulation-layer case. As previous calculations, we also give spin-orientation dependence for the (001) and (110) cases.

$$d_{v_1,0,v_2,0} = \frac{(b_1 + b_2)^5}{6b_1^3 b_2^3}, \quad (25)$$

where b_1 and b_2 are from valleys v_1 and v_2 , respectively, which may have different m_z masses. Substituting the obtained effective widths into Eqs. (4), (17), (18), and (20), one can obtain τ_s for three 2DEG orientations, (111), (001), and (110), for any arbitrary spin orientation, as a function of N_A , N_{2D} , and N_i . We collect all of the useful information in Fig. 8. For visual clarity, we leave out the reciprocal linear dependence on N_i and plot $\tilde{\tau}_s = (N_i/10^{16} \text{ cm}^{-3})\tau_s$. For the inversion- (accumulation-) layer case, we choose the majority (minority) acceptor density N_A as $10^{14}, 10^{16}$, and 10^{18} cm^{-3} ($10^{10}, 10^{12}$, and 10^{14} cm^{-3}), and $10^{10} < N_{2D} < 10^{13} \text{ cm}^{-2}$, covering typical experimental choices. As expected, $\tilde{\tau}_s$ decreases slightly with an increasing N_{2D} due to the steeper confinement. Similar trends occur with N_A . More importantly, of course, the absolute time $\tau_s \propto \tilde{\tau}_s/N_i$ decreases much faster with N_i . For clean interfaces, $N_i \approx N_A(N_D)$ for the inversion (accumulation) layer. However, for some highly disordered Si/SiO₂ interfaces, including oxide charges on the SiO₂ side, N_i may be much larger than the majority dopant density. In the high

N_{2D} limit, the potential confinement is dominated by N_{2D} over $N_{\text{depl}} \propto \sqrt{N_A}$ [6] and, as a result, $\tilde{\tau}_s$ converges for different N_A 's.

Finally, we should note that the SOC parameter Δ_{so} used here is 0.1 meV, the value for As dopants. It is 0.03 or 0.3 meV for P or Sb [43]. For interface disorder or majority acceptors, Δ_{so} and η [see Eq. (2)] must be studied separately. This future study can be important, and we come back to this issue at the end. In fact, the effective SOC parameters of the interface impurities are an important unknown in the theory, which can be adjusted to get agreement between our theory and all existing experimental data. We refrain from doing so, however, emphasizing that if the measured 2D spin-relaxation time shows a positive correlation with the quality of the interface (i.e., improving interface quality leads to a longer spin-relaxation time), then it is likely that the impurity-induced Yafet mechanism discussed in this paper is playing a dominant role, in contrast to the DP mechanism, which mostly leads to a lower spin-relaxation time with higher mobility.

Now we briefly discuss some available experimentally measured T_1 's ($\equiv \tau_s$ in our notation) in Si 2DEG. We stress that these samples are not particularly highly doped and, therefore, our mechanism is not expected to be dominant unless the extrinsic disorder associated with interface impurities plays a crucial role. Reference [30] measured T_1 in Si/SiGe (001) quantum wells with a relatively high-quality interface and mobility μ . Their device I, with $\mu = 9 \text{ m}^2/\text{V s}$, 2D electron density $N_{2D} = 3 \times 10^{11} \text{ cm}^{-2}$, and well thickness $d = 20 \text{ nm}$, yields a $T_1 = 2.0 \mu\text{s}$. Their device II, with $\mu = 19 \text{ m}^2/\text{V s}$, $N_{2D} = 1.7 \times 10^{11} \text{ cm}^{-2}$, and $d = 15 \text{ nm}$, yields a $T_1 = 0.95 \mu\text{s}$. These N_{2D} and d combinations indicate that only the ground subbands in the $\pm z$ valley are likely occupied.

We show in the following that our mechanism cannot quantitatively account for the measured T_1 , even if we assume that the lowest subbands in the $\pm x$ and $\pm y$ valleys are occupied. Using our calculation that leads to Fig. 4, one needs a 3D impurity density $N_i \approx 3.3 \times 10^{16} \text{ cm}^{-3}$ or an effective 2D impurity density $n_i \approx N_i d \approx 6.6 \times 10^{10} \text{ cm}^{-2}$ for device I, and $N_i \approx 5 \times 10^{16} \text{ cm}^{-3}$ or $n_i \approx N_i d \approx 7.5 \times 10^{10} \text{ cm}^{-2}$ for device II. The precise n_i depends weakly on the detailed impurity distribution in the 2DEG. To estimate the experimental impurity density residing in the quantum well, we use the theoretical result of Ref. [61] which relates mobility with the charged impurity. From Fig. 1 of that paper, for the (001) well orientation with two ground valleys, one needs $n_i \approx 10^{10} \text{ cm}^{-2}$ for device I and $n_i \approx 4 \times 10^9 \text{ cm}^{-2}$ for device II. These n_i 's make our spin-relaxation mechanism too weak to yield the measured T_1 time. The spin anisotropy in our calculation has the same sign as in the measurement but is not as large in magnitude [$T_1(\theta = \pi/2)/T_1(\theta = 0) = 1.1$ versus the measured 1.5].

We do, however, mention that our mechanism using these estimated experimental impurity densities gives T_1

values within an order of magnitude of the measured T_1 values. Given the uncertainties associated with the impurity SOC parameters, the possibility that the Yafet mechanism is perhaps playing a (minor) role in the experiment cannot be ruled out, although it does appear that the main spin-relaxation mechanism in these high-mobility Si/SiGe quantum wells is likely to be the DP mechanism.

Reference [62] measured T_1 time for 2DEG in a Si/SiO₂ (001) accumulation layer doped with 10^{14} P donors. At a gate voltage of 2 V, $T_1 = 0.33 \mu\text{s}$, while $\mu = 1 \text{ m}^2/\text{V s}$ and $N_{2D} = 4 \times 10^{11} \text{ cm}^{-2}$ [63]. Once again, we check the effect of our mechanism by assuming for a moment that the $\pm x$ and $\pm y$ valleys are reached. From Fig. 8, one needs about $N_i \approx 3 \times 10^{17} \text{ cm}^{-3}$, or $n_i \approx 4 \times 10^{11} \text{ cm}^{-2}$. Note that the unknown N_A value affects the result only slightly (a factor less than 2). From the measured mobility at 5 K, we can deduce $n_i \approx 10^{11} \text{ cm}^{-2}$ [61]. As a result, the impurity density is again too small to induce the measured T_1 time by our mechanism, even if the finite mobility is entirely caused by impurities in the 2DEG region and $\Delta_{\text{so}} = 0.1 \text{ meV}$. However, our mechanism now gives a T_1 which is within a factor of 4 of the measured value, indicating that, for the spin relaxation in disordered Si MOSFETs, our impurity-driven mechanism is perhaps playing a more important quantitative role. This finding is not unexpected since Si MOSFETs typically have larger impurity densities than Si/SiGe quantum wells, leading to possibly stronger spin relaxation due to the Yafet mechanism. Note that the DP mechanism does not find agreement with the experimental data either, which can be verified by the calculation in Ref. [18] in combination with the experimental parameters. It is possible that, in Si MOSFETs, both the DP and Yafet mechanisms are operational in producing the observed low value of T_1 in the experiment. Obviously, more experimental measurements are essential in understanding this important puzzle.

We propose several experimental ways to properly investigate the nature of spin relaxation in Si 2DEG. To begin with, (111) and (110) orientations are better suited for our intervalley spin-flip mechanism to make an important contribution, as we mentioned earlier. Moreover, a lowering of the x and y valleys in the (001) 2DEG may also show a sudden jump of the spin-relaxation rate, which serves as a turn-on signal of our mechanism. This valley tuning can be achieved by external stress or gating, as emphasized in Sec. III. A similarly sudden change in τ_s anisotropy can also occur for the (110) 2DEG due to our mechanism.

Apart from the 2DEG plane orientation, a number of aspects are important in the experimental verification of our proposed spin-relaxation mechanism. First, it is crucial (and we urge future experiments) to measure a series of samples with different mobilities (μ) at the same carrier densities (and all other parameters). The (anti)correlation of τ_s with μ is a characteristic signature for the Yafet (DP)

spin-relaxation mechanism [1]. The crossover occurs at modest doping levels, as the spin lifetime from the DP mechanism rises rapidly past tens of microseconds already around the mobility $5 \times 10^4 \text{ cm}^2/\text{Vs}$ [18]. Specifically, our mechanism should become dominant when the 2DEG region is heavily doped.

Second, there is a positive correlation between conduction electron density and spin relaxation in the DP mechanism. The Rashba or generalized Dresselhaus field scales linearly with the wave vector (\mathbf{k}) measured from the valley bottom, yet our spin-flip matrix elements depend little on \mathbf{k} . Third, for the MOSFET setup, it is useful to measure τ_s separately for both the bulk Si and the 2DEG to deduce the contribution of the interface disorder to the 2DEG spin relaxation.

It is also possible to deduce the distribution of 2DEG impurities from the gate-voltage dependence of τ_s : from Eqs. (15) and (7), $\tau_s \propto d$ for a uniformly distributed $N_i(z)$, while $\tau_s \propto d^2$ if all impurities are concentrated at the interface [$N_i(z) \propto \delta(z)$]. For the Si/SiGe setup, making two-sided symmetric confinement may separate out the contribution from the DP spin relaxation, as the change of the interfacial symmetry property greatly affects the DP mechanism through the envelope functions but leaves the Yafet one the same. In addition, as our calculated spin-orientation dependence [rooted in Eqs. (4)–(6)] is distinct from that of the DP mechanism due to the Rashba or Dresselhaus field, the τ_s anisotropy measurement can also help to disentangle the two contributions (for the large magnetic-field limit, we note that the DP mechanism is partially suppressed, which is similar to the bulk case [1,17,64]).

Parenthetically, while this work does not focus on the DP mechanism, we point out the existing studies concerning its various contributions [65]. Different views have emerged to account for the same experimental measurement in a Si/SiGe quantum well [16], whether it be dominated by the Rashba field [16–18] or the Dresselhaus one [11,12]. They lead to different SOC anisotropy but a similar overall spin-relaxation rate, as both SOC fields scale linearly in a wave vector in the 2DEGs. DP spin relaxation and its anisotropy has also been studied in Si/SiGe quantum dots [66]. To date, the relative magnitude of Rashba- and Dresselhaus-like SOC has yet to be verified experimentally [67].

Finally, our mechanism relies on the short-range interaction with the impurity core and directly measures the SOC strength of the impurity atoms. The spin-relaxation rate scales quadratically with Δ_{so} [see Eqs. (7) and (15)], and it increases significantly by switching from low- to high-atomic-number dopants for the same density. Therefore, different types of impurities that lead to a similar mobility may yield very different τ_s times according to their SOC strengths, a unique signature of this spin-relaxation process.

We suggest conducting future spin-relaxation measurements in 2D Si systems as a systematic function of

mobility, carrier density, impurity type, surface and spin orientation, and applied stress in order to develop a complete understanding of the mechanisms controlling the spin relaxation of free carriers near Si surfaces. The few existing measurements simply do not provide enough information for a definitive conclusion.

Last, in order to establish the relative strength of our spin-relaxation rate in comparison to the momentum-relaxation rate, which determines the device charge mobility, we calculate their relative ratio (ν) for a few representative cases where impurities are the dominant source of scattering (i.e., at low temperatures, where phonons are unimportant). We take a simplified uniform distribution of the highly doped 2DEG. Define $\nu = \tau_m^{2D}/\tau_s^{2D}$, where τ_s^{2D} follows from Eq. (15) and the momentum-relaxation time τ_m^{2D} takes the form appropriate for mobility calculations in 2D transport studies [36,61].

The momentum scattering matrix elements are governed by the well-known screened Coulomb interaction in the intravalley scattering, as appropriate for scattering by the random charged impurities. For this interaction, the impurity-distribution profile can be approximated as a δ function normal to the 2DEG plane. Under the 2D RPA screening, the momentum-relaxation rate in the quantum limit is given as [36,61]

$$\frac{1}{\tau_m^{2D}} = \frac{4\pi e^4 m_{\text{eff}} n_i}{\hbar^3 \kappa^2 k_F^2} \int_0^1 \frac{dx x^2}{(x + q_{\text{TF}}/2k_F)^2 \sqrt{1-x^2}}, \quad (26)$$

where m_{eff} is the conductivity effective mass different for each specific 2DEG orientation [6,61], n_i is the 2D impurity density, permittivity κ is the Si permittivity, k_F is the 2D Fermi wave number, $q_{\text{TF}} = m_{\text{eff}} e^2 g / \hbar^2 \kappa$ is the 2D Thomas-Fermi wave number, and $g = g_v g_s$ is the number of populated valleys including the spin degree of freedom (with v and s denoting valley and spin, respectively).

The dependence on the Fermi level (similarly, on k_F or N_{2D}) is very slow for both τ_s^{2D} and τ_m^{2D} within a given Fermi-energy window between 2D subbands [36]. This slow dependence can be clearly seen in Fig. 8 for τ_s^{2D} over many orders of magnitude of N_{2D} . For τ_m^{2D} , we plot its explicit dependence on k_F in Fig. 9 for (111) and (110) 2DEG orientations for which both leading-order momentum and spin-relaxation rates are nonvanishing in the quantum limit. They are both nearly constant over the large region $k_F \leq 0.1 \text{ \AA}^{-1}$, i.e., about 10% of the length of the Brillouin zone.

This near independence of the Fermi level allows us to obtain a simple estimation for ν for each specific 2DEG orientation. To get the leading-order estimate, it is sufficient to replace $\tau_s^{(111)}$ and $\tau_s^{(110)}$ with the square-well results through Eqs. (17) and (19), as well as a variation of Eq. (21). We get the following results:

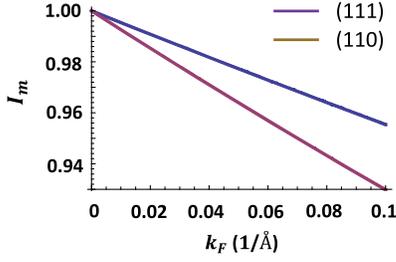


FIG. 9. The k_F -dependent factor in the momentum scattering rate [Eq. (26)], $I_m \equiv (1/\pi)(k_F/q_{\text{TF}})^2 \int_0^1 (dx x^2 / (x + q_{\text{TF}}/2k_F)^2 \sqrt{1-x^2})$, for the (111) and (110) 2DEG orientations, respectively, over $0 < k_F < 0.1 \text{ \AA}^{-1}$.

$$\nu_{111} = \frac{400a_B^6 \Delta_{\text{so}}^2}{\hbar^4 d^2} \sqrt{\frac{m_t(m_t + 2m_l)}{3}} / \left(\frac{1}{m_t} + \frac{3}{m_t + 2m_l} \right), \quad (27)$$

$$\nu_{110} = \frac{128a_B^6 \Delta_{\text{so}}^2}{\hbar^4 d^2} \sqrt{m_t m_l} / \left(\frac{1}{m_t} + \frac{2}{m_t + m_l} \right) \quad (28)$$

for the (111) and (110) 2DEGs, respectively. With respect to typical Δ_{so} and d parameters, $\nu_{111} = 1.5 \times 10^{-5} (\Delta_{\text{so}}/0.1 \text{ meV})^2 (20 \text{ nm}/d)^2$ and $\nu_{110} = 5.6 \times 10^{-6} (\Delta_{\text{so}}/0.1 \text{ meV})^2 (20 \text{ nm}/d)^2$. Obviously, this ratio depends quadratically on the impurity SOC constant Δ_{so} . For the expected typical values of the impurity SOC Δ_{so} in Si and the 2DEG width d , this ratio ν varies between 10^{-6} and 10^{-4} . By comparison, we note that, for intrinsic phonon-induced spin and momentum-relaxation rates in 3D bulk Si, this ratio is around 10^{-5} [54], which is determined completely by the host Si SOC.

V. SUMMARY AND OUTLOOK

We have introduced in Si 2DEG a previously overlooked yet important spin-relaxation mechanism due to electron-impurity scattering. This mechanism dominates over other spin relaxations in the multivalley Si conduction band as impurity density increases, and it can be significantly suppressed when electrons are transferred into two opposing ground valleys by specific 2DEG orientations and stress configurations. We provide the general expression for obtaining the leading-order spin-relaxation rate under arbitrary confinement potential, applied stress, and subband occupation. We calculate quantitatively the (T_1) spin-relaxation time τ_s (s) as a function of spin orientation \mathbf{s} , as well as of the conduction electron density and confinement strength for the representative square and triangular wells.

Moreover, the consequences of various stress configurations have been worked out in detail. Importantly, this newly discovered spin-relaxation mechanism combined with the Si 2DEG setup provides interesting possibilities to tune the spin lifetime as well as its dependence on spin

orientation (or the applied magnetic-field direction) substantially by on-chip gate voltages, and possibly by local stress. Such a tunability of spin relaxation in MOSFET-type Si devices could have potential spintronic applicability.

Also crucially, we provide experimental ways (elaborated on in Sec. IV) to verify our spin-relaxation mechanism and distinguish it from the DP spin-relaxation effect from the generalized Rashba-Dresselhaus field in Si 2DEGs, by exploiting their different dependence on impurity densities and types, on the interface symmetry properties, and on 2DEG plane, spin, and stress orientations.

Regarding a general expansion of this model, we point out that, for 2DEG near the interface with a considerable amount of disorder, a variation of our impurity-driven intervalley spin-flip process may become quantitatively important in determining the spin-relaxation rate. As mentioned in the Introduction, the DP spin-relaxation mechanism alone leads to a much longer spin lifetime for low-mobility 2DEG than has been observed experimentally [18]. However, spin lifetime is apparently shorter in 2DEG near typical Si/SiO₂ interfaces, indicating impurity-driven Elliott-Yafet spin relaxation. While our spin-flip matrix elements [Eq. (2)] apply specifically to substitutional impurities in Si with their given symmetry, it is a basic rule that lower-symmetry disorder inherits the allowed transition matrix elements. Thus, the key idea of zeroth-order intervalley spin-flip scattering [22] robustly holds for irregular defects, with additional scattering channels potentially being open depending on the specific defects. It is, therefore, possible that interface impurities (even when they are completely nonmagnetic—a restriction that our theory holds itself to—any magnetic interface impurities will, of course, very strongly affect spin relaxation near the surface through direct magnetic spin-flip scattering) are playing a strong role in determining the 2D spin-relaxation time in disordered Si/SiO₂ MOSFETs by participating in the Yafet process identified and analyzed in this work. Obviously, figuring this problem out presents an open and important future experimental challenge in Si spintronics.

ACKNOWLEDGMENTS

This work is supported by LPS-MPO-CMTC.

APPENDIX: PHYSICS OF INTERVALLEY COUPLING IN SI AND SYMMETRY ANALYSIS

To be self-contained, we provide the essential physical picture of intervalley coupling in bulk Si and the relevant symmetry analysis and selection rules for Sec. II.

Bulk Si has a crystal structure [Fig. 10(a) in the absence of the impurity substitution] consisting of two sets of interpenetrating face-centered-cubic lattices and a space symmetry group O_h^7 . Its lowest conduction band in the wave-vector space has its bottom not at the center of the

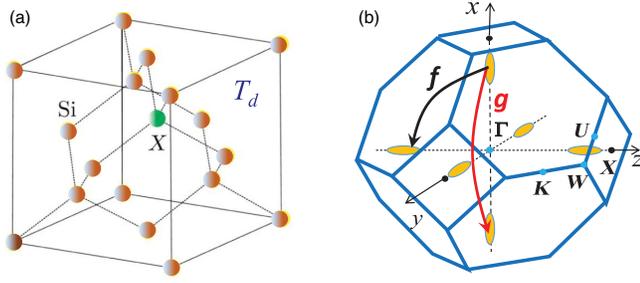


FIG. 10. (a) The Si crystal lattice, with one of its Si atoms replaced by an impurity denoted by X. As a result, the symmetry of the whole Hamiltonian system is reduced to that of the T_d point group. (b) The Brillouin zone of the Si crystal. The yellow ellipsoids mark the low-energy surface of the six conduction valleys. Two representative examples are marked—for the intervalley g and f processes.

Brillouin zone but along the cubic-axis directions. Crystal symmetry determines that six energy valleys reside cylindrically along the $\pm x$, $\pm y$, and $\pm z$ axes. This well-known multivalley picture of Si supplies relevant information for the electron states involved in this work. The transitions between the electron states residing near the bottom of the conduction valleys can obviously be classified into three groups [see Fig. 10(b)]: (I) within the same valley (“intravalley”), (II) between two opposing valleys (“intervalley g process”), and (III) between two nonopposing valleys (“intervalley f process”).

The specific scattering potential we deal with comes from the impurity which replaces one of the Si atoms. This impurity immediately invalidates the translational symmetry of the Si crystal, and as a result the symmetry of the Hamiltonian system falls into a point group around the impurity [Fig. 10(a)]. This point group has the same symmetry operations as a tetrahedron molecule: C_2 rotation about the x , y , or z axis, C_3 rotation about body diagonals, σ reflection about the face-diagonal planes, and S_4 (C_4 followed by reflection) about the x , y , or z axis, and the point group is called the T_d group.

To utilize the symmetry property of this system for selection rules, we work with symmetrized electrons states by linearly combining six different valleys, rather than states in each individual valley as one is used to. The six combinations are as follows [41]:

$$\psi_{A_1} = \frac{1}{\sqrt{6}}(1, 1, 1, 1, 1, 1), \quad (\text{A1})$$

$$\psi_{E^I} = \frac{1}{2}(1, 1, -1, -1, 0, 0), \quad (\text{A2})$$

$$\psi_{E^{II}} = \frac{1}{2\sqrt{3}}(1, 1, 1, 1, -2, -2); \quad (\text{A3})$$

$$\psi_{T_2^I} = \frac{1}{\sqrt{2}}(1, -1, 0, 0, 0, 0), \quad (\text{A4})$$

$$\psi_{T_2^{II}} = \frac{1}{\sqrt{2}}(0, 0, 1, -1, 0, 0), \quad (\text{A5})$$

$$\psi_{T_2^{III}} = \frac{1}{\sqrt{2}}(0, 0, 0, 0, 1, -1), \quad (\text{A6})$$

where the ordering of the six components of the state vectors applies to the valley bottom states along the $+x$, $-x$, $+y$, $-y$, $+z$, and $-z$ axes, respectively. Each new state is given a name at the subscript of ψ , following the well-established naming system (see the T_d -group character table in Ref. [68] or [22]). The selection rules immediately follow since only the same-symmetry states can couple, while different-symmetry states are not mixed by the scatterer potential, which transforms as the identity in this group. Once we get the scattering matrix elements that do not vanish, we can easily make linear combinations between them to transfer back to the familiar intravalley and intervalley g and f processes [22].

Thus far, we have not considered spin degrees of freedom or SOC. To include spin, we can expand the basis to be the product space of six valleys and two spins. It turns out that two \bar{F} states emerge from this valley-spin coupling. To be concrete, the multiplication expressions are as follows. The pure spin transforms as \bar{E}_1 , and we then have

$$A_1 \times \bar{E}_1 = \bar{E}_1, \quad (\text{A7})$$

$$E \times \bar{E}_1 = \bar{F}, \quad (\text{A8})$$

$$T_2 \times \bar{E}_1 = \bar{E}_2 + \bar{F}. \quad (\text{A9})$$

We may follow a similar procedure as the spinless case to obtain spin-dependent scattering selection rules [22]. Only states with the same symmetry can be coupled. Among all five nonvanishing couplings [each of the four states in Eqs. (A7)–(A9) coupling to itself, as well as the intercoupling of the two \bar{F} states from Eqs. (A8) and (A9)], we find that there are spin-flip terms in two of them: the difference between the \bar{E}_2 and \bar{F} self-coupling matrix elements from Eq. (A9), and the intercoupling matrix element between two different \bar{F} 's. This finding leads to the two terms in Eq. (2). After transforming back to the intravalley, intervalley g , and intervalley f processes, we find [22] that both terms contribute to the f -process spin flip.

-
- [1] I. Žutić, J. Fabian, and S. Das Sarma, Spintronics: Fundamentals and applications, *Rev. Mod. Phys.* **76**, 323 (2004).
 [2] R. Jansen, Silicon spintronics, *Nat. Mater.* **11**, 400 (2012).
 [3] V. Sverdlov and S. Selberherr, Silicon spintronics: Progress and challenges, *Phys. Rep.* **585**, 1 (2015).
 [4] B. E. Kane, A silicon-based nuclear spin quantum computer, *Nature (London)* **393**, 133 (1998).
 [5] F. A. Zwanenburg, A. S. Dzurak, A. Morello, M. Y. Simmons, L. C. L. Hollenberg, G. Klimeck, S. Rogge,

- S. N. Coppersmith, and M. A. Eriksson, Silicon quantum electronics, *Rev. Mod. Phys.* **85**, 961 (2013).
- [6] T. Ando, A. B. Fowler, and F. Stern, Electronic properties of two-dimensional systems, *Rev. Mod. Phys.* **54**, 437 (1982).
- [7] L. von Klitzing, The quantized Hall effect, *Rev. Mod. Phys.* **58**, 519 (1986).
- [8] Yu. L. Bychkov and E. I. Rashba, Oscillatory effects and the magnetic susceptibility of carriers in inversion layers, *J. Phys. C* **17**, 6039 (1984).
- [9] L. E. Golub and E. L. Ivchenko, Spin splitting in symmetrical SiGe quantum wells, *Phys. Rev. B* **69**, 115333 (2004).
- [10] M. O. Nestoklon, L. E. Golub, and E. L. Ivchenko, Spin and valley-orbit splittings in SiGe/Si heterostructures, *Phys. Rev. B* **73**, 235334 (2006).
- [11] M. O. Nestoklon, E. L. Ivchenko, J.-M. Jancu, and P. Voisin, Electric field effect on electron spin splitting in SiGe/Si quantum wells, *Phys. Rev. B* **77**, 155328 (2008).
- [12] M. Prada, G. Klimeck, and R. Joynt, Spinorbit splittings in Si/SiGe quantum wells: From ideal Si membranes to realistic heterostructures, *New J. Phys.* **13**, 013009 (2011).
- [13] G. Dresselhaus, Spin-orbit coupling effects in zinc blende structures, *Phys. Rev.* **100**, 580 (1955).
- [14] O. Krebs and P. Voisin, Giant Optical Anisotropy of Semiconductor Heterostructures with No Common Atom and the Quantum-Confined Pockels Effect, *Phys. Rev. Lett.* **77**, 1829 (1996).
- [15] M. I. D'yakonov and V. I. Perel', Spin relaxation of conduction electrons in noncentrosymmetric semiconductors, *Fiz. Tverd. Tela (Leningrad)* **13**, 3581 (1971) [*Sov. Phys. Solid State* **13**, 3023 (1972)].
- [16] Z. Wilamowski, W. Jantsch, H. Malissa, and U. Rössler, Evidence and evaluation of the Bychkov-Rashba effect in SiGe/Si/SiGe quantum wells, *Phys. Rev. B* **66**, 195315 (2002).
- [17] Z. Wilamowski and W. Jantsch, Suppression of spin relaxation of conduction electrons by cyclotron motion, *Phys. Rev. B* **69**, 035328 (2004).
- [18] C. Tahan and R. Joynt, Rashba spin-orbit coupling and spin relaxation in silicon quantum wells, *Phys. Rev. B* **71**, 075315 (2005).
- [19] Y. V. Pershin, Optically induced suppression of spin relaxation in two-dimensional electron systems with Rashba interaction, *Phys. Rev. B* **75**, 165320 (2007).
- [20] V. K. Dugaev, E. Ya. Sherman, V. I. Ivanov, and J. Barnaś, Spin relaxation and combined resonance in two-dimensional electron systems with spin-orbit disorder, *Phys. Rev. B* **80**, 081301(R) (2009).
- [21] Y. Yafet, in *Solid State Physics*, Vol. 14, edited by F. Seitz and D. Turnbull (Academic, New York, 1963), p. 1.
- [22] Y. Song, O. Chalaev, and H. Dery, Donor-Driven Spin Relaxation in Multivalley Semiconductors, *Phys. Rev. Lett.* **113**, 167201 (2014).
- [23] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors*, 3rd ed. (Springer, Berlin, 2005), pp. 215–217.
- [24] G. Dorda, I. Eisele, and H. Gesch, Many-valley interactions in *n*-type silicon inversion layers, *Phys. Rev. B* **17**, 1785 (1978).
- [25] H. Ue and S. Maekawa, Electron-spin-resonance studies of heavily phosphorus-doped silicon, *Phys. Rev. B* **3**, 4232 (1971).
- [26] J. D. Quirt and J. R. Marko, Absolute spin susceptibilities and other ESR parameters of heavily doped *n*-type silicon. I. Metallic samples, *Phys. Rev. B* **5**, 1716 (1972).
- [27] J. H. Pifer, Microwave conductivity and conduction-electron spin-resonance linewidth of heavily doped Si:P and Si:As, *Phys. Rev. B* **12**, 4391 (1975).
- [28] Y. Ochiai and E. Matsuura, Spin-lattice relaxation at high temperatures in heavily doped *n*-type silicon, *Phys. Status Solidi (a)* **45**, K101 (1978).
- [29] V. Zarifis and T. G. Castner, Observation of the conduction-electron spin resonance from metallic antimony-doped silicon, *Phys. Rev. B* **57**, 14600 (1998).
- [30] A. M. Tyryshkin, S. A. Lyon, W. Jantsch, and F. Schäffler, Spin Manipulation of Free Two-Dimensional Electrons in Si/SiGe Quantum Wells, *Phys. Rev. Lett.* **94**, 126802 (2005).
- [31] E. Y. Sherman, Minimum of spin-orbit coupling in two-dimensional structures, *Phys. Rev. B* **67**, 161303(R) (2003).
- [32] H. I. Ralph, G. Simpson, and R. J. Elliott, Central-cell corrections to the theory of ionized-impurity scattering of electrons in silicon, *Phys. Rev. B* **11**, 2948 (1975).
- [33] H. M. A. El-Ghanem and B. K. Ridley, Impurity scattering of electrons in non-degenerate semiconductors, *J. Phys. C* **13**, 2041 (1980).
- [34] B. K. Ridley, *Quantum Processes in Semiconductors*, 5th ed. (Oxford University Press, New York, 2013).
- [35] A. Hartstein, T. H. Ning, and A. B. Fowler, Electron scattering in silicon inversion layers by oxide and surface roughness, *Surf. Sci.* **58**, 178 (1976).
- [36] S. Das Sarma and E. H. Hwang, Universal density scaling of disorder-limited low-temperature conductivity in high-mobility two-dimensional systems, *Phys. Rev. B* **88**, 035439 (2013).
- [37] G. H. Wannier, The structure of electronic excitation levels in insulating crystals, *Phys. Rev.* **52**, 191 (1937).
- [38] J. M. Luttinger and W. Kohn, Motion of electrons and holes in perturbed periodic fields, *Phys. Rev.* **97**, 869 (1955).
- [39] G. L. Bir and G. E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors* (Halsted Press, Jerusalem, 1974), pp. 187–294.
- [40] M. G. Burt, The justification for applying the effective-mass approximation to microstructures, *J. Phys. Condens. Matter* **4**, 6651 (1992).
- [41] W. Kohn, *Solid State Physics*, Vol. 5, edited by F. Seitz and D. Turnbull (Academic Press, New York, 1957), p. 257.
- [42] R. L. Aggarwal and A. K. Ramdas, Optical determination of the symmetry of the ground states of group-V donors in silicon, *Phys. Rev.* **140**, A1246 (1965).
- [43] T. G. Castner, Orbach spin-lattice relaxation of shallow donors in silicon, *Phys. Rev.* **155**, 816 (1967).
- [44] A. Campo and R. Kummel, Valley splitting in (001), (011) and (111) Si-MOSFETS, *Solid State Commun.* **37**, 433 (1981).
- [45] M. Friesen, S. Chutia, C. Tahan, and S. N. Coppersmith, Valley splitting theory of SiGe/Si/SiGe quantum wells, *Phys. Rev. B* **75**, 115318 (2007).
- [46] A. L. Saraiva, M. J. Calderón, X. Hu, S. Das Sarma, and B. Koiller, Physical mechanisms of interface-mediated intervalley coupling in Si, *Phys. Rev. B* **80**, 081305(R) (2009).
- [47] A. L. Saraiva, M. J. Calderón, R. B. Capaz, X. Hu, S. Das Sarma, and B. Koiller, Intervalley coupling for

- interface-bound electrons in silicon: An effective mass study, *Phys. Rev. B* **84**, 155320 (2011).
- [48] F. Stern, Self-consistent results for n -type Si inversion layers, *Phys. Rev. B* **5**, 4891 (1972).
- [49] P. A. M. Dirac, The quantum theory of the emission and absorption of radiation, *Proc. R. Soc. A* **114**, 243 (1927).
- [50] F. Fermi, *Nuclear Physics* (University of Chicago Press, Chicago, 1950).
- [51] O. Chalaev, Y. Song, and H. Dery, Suppressing spin relaxation in silicon, [arXiv:1609.07077](https://arxiv.org/abs/1609.07077).
- [52] S. Das Sarma, R. K. Kalia, M. Nakayama, and J. J. Quinn, Stress and temperature dependence of subband structure in silicon inversion layers, *Phys. Rev. B* **19**, 6397 (1979).
- [53] R. J. Elliott, Theory of the effect of spin-orbit coupling on magnetic resonance in some semiconductors, *Phys. Rev.* **96**, 266 (1954).
- [54] J. L. Cheng, M. W. Wu, and J. Fabian, Theory of the Spin Relaxation of Conduction Electrons in Silicon, *Phys. Rev. Lett.* **104**, 016601 (2010).
- [55] P. Li and H. Dery, Spin-Orbit Symmetries of Conduction Electrons in Silicon, *Phys. Rev. Lett.* **107**, 107203 (2011).
- [56] J.-M. Tang, B. T. Collins, and M. E. Flatté, Electron spin-phonon interaction symmetries and tunable spin relaxation in silicon and germanium, *Phys. Rev. B* **85**, 045202 (2012).
- [57] Y. Song and H. Dery, Analysis of phonon-induced spin relaxation processes in silicon, *Phys. Rev. B* **86**, 085201 (2012).
- [58] M. Shiraishi, Y. Honda, E. Shikoh, Y. Suzuki, T. Shinjo, T. Sasaki, T. Oikawa, K. Noguchi, and T. Suzuki, Spin transport properties in silicon in a nonlocal geometry, *Phys. Rev. B* **83**, 241204(R) (2011).
- [59] F. F. Fang and W. E. Howard, Negative Field-Effect Mobility on (100) Si Surfaces, *Phys. Rev. Lett.* **16**, 797 (1966).
- [60] F. Stern, Calculated Energy Levels and Optical Absorption in n -Type Si Accumulation Layers at Low Temperature, *Phys. Rev. Lett.* **33**, 960 (1974).
- [61] E. H. Hwang and S. Das Sarma, Valley-dependent two-dimensional transport in (100), (110), and (111) Si inversion layers at low temperatures and carrier densities, *Phys. Rev. B* **87**, 075306 (2013).
- [62] S. Shankar, A. M. Tyryshkin, Jianhua He, and S. A. Lyon, Spin relaxation and coherence times for electrons at the Si/SiO₂ interface, *Phys. Rev. B* **82**, 195323 (2010).
- [63] S. Shankar, Ph.D. thesis, Princeton University, 2010.
- [64] M. M. Glazov, Magnetic field effects on spin relaxation in heterostructures, *Phys. Rev. B* **70**, 195314 (2004).
- [65] S. D. Ganichev and L. E. Golub, Interplay of Rashba/Dresselhaus spin splittings probed by photogalvanic spectroscopy A review, *Phys. Status Solidi B* **251**, 1801 (2014).
- [66] M. Raith, P. Stano, and J. Fabian, Theory of single electron spin relaxation in Si/SiGe lateral coupled quantum dots, *Phys. Rev. B* **83**, 195318 (2011).
- [67] C. Tahan and R. Joynt, Relaxation of excited spin, orbital, and valley qubit states in ideal silicon quantum dots, *Phys. Rev. B* **89**, 075302 (2014).
- [68] C. J. Bradley and A. P. Cracknell, *The Mathematical Theory of Symmetry in Solids: Representation Theory for Point Groups and Space Groups* (Clarendon Press, Oxford, 1972), p. 433.