Envelope-function theory of inhomogeneous strain in semiconductor nanostructures

Andrea Secchi [*] and Filippo Troiani

Centro S3, CNR-Istituto di Nanoscienze, I-41125 Modena, Italy

Strain represents an ubiquitous feature in semiconductor heterostructures, and can be engineered by different means in order to improve the properties of various devices, including advanced metal-oxide-semiconductor field-effect transistors and spin-based qubits. However, its treatment within the envelope function framework is well established only for the homogeneous case, thanks to the theory of Bir and Pikus. Here, we generalize this theory to the case of inhomogeneous strain. By fully accounting for the relativistic effects and metric aspects of the problem, we derive a complete envelope-function Hamiltonian, including the terms that depend on first and second spatial derivatives of the strain tensor.

I. INTRODUCTION

Strain represents a common feature in semiconductor nanostructures. It develops spontaneously during their fabrication process, because of the lattice mismatch between heterogeneous layers, and can be induced by cooling the system to cryogenic temperatures, due to the presence of materials with different thermal-expansion coefficients [1–3]. As an uncontrolled or unaccounted phenomenon, strain can result in significant differences between the actual and the nominal properties of the nanostructure. On the other hand, strain can be intentionally engineered, in order to modulate the band structure and increase the carrier mobility, an approach that is actively pursued, e.g., with metal-oxide-semiconductor field-effect transistors (MOSFETs) [4,5] or silicon nanowires [6].

These effects are particularly relevant in semiconductor-based implementations of quantum computing. Silicon and germanium quantum dots have emerged as promising hosts of electron- or hole-spin qubits [7–27], whose properties can be strongly affected by strain. In particular, it has been shown that in these systems inhomogeneous strain can modify both the localization of the confined particle and its coupling to external fields, specifically through a modulation of the Rabi frequency [28] and of the g factor [29].

The tool of election for simulating the properties of spin qubits in semiconductor quantum dots is represented by the Luttinger and Kohn (LK)'s envelope-function formalism [30–32]. This applies to crystalline systems subjected to a spatially slow-varying external potential, such as the one generated by the metal gates used in electrostatically defined nanostructures. Describing the effects of strain on the electron and hole states requires an extension of LK's theory, which was developed by Bir and Pikus (BP) for the case where the strain tensor is small and homogeneous [33–37]. Even in these conditions, the absolute displacements of the ions (with respect to the unstrained crystal) may be comparable or larger than the lattice constant. This makes the displacements unsuitable as an expansion parameter for the electron-nuclei potential, unlike for the theory of electron-phonon interactions

[38]. BP's key idea was to introduce a new set of electron coordinates that make the power expansion of the electron-nuclei potential in the strain tensor possible, thus enabling a perturbative calculation of the electron and hole states.

In view of the above, a generalization of BP's theory to the case of inhomogeneous strain would be highly desirable, but is far from trivial. In our understanding, the previous attempts that have been made in this direction are affected by significant shortcomings. These consist either in an incorrect treatment of the Schrödinger equation in the required set of curvilinear coordinates, resulting in the non-Hermiticity of the particle Hamiltonian [39], or in the use of a nonpractical basis set within a nonrelativistic treatment, which precludes from the outset an accurate description of spin-orbit interactions [40].

In this article, we extend BP's theory to the case of inhomogeneous strain in a rigorous and comprehensive way. This is achieved by properly taking into account the modifications to the quantum-mechanical formalism that arise when the metric is non-Cartesian [41,42], and by including relativistic corrections to the Schrödinger equation via a low-energy expansion of the covariant Dirac equation [43]. Our central result — applicable to a variety of semiconductor nanostructures, in the presence of slowly-varying inhomogeneous strain and external electrostatic potential — is a set of equations, whose solution gives the envelope functions within a manifold of arbitrary dimension. From these we derive, as a relevant case, the strain-related six-band Hamiltonian for the hole states in silicon and germanium, and more generally in crystals with diamond structure. For the sake of readability, the main logical steps that have been followed are reflected in the structure of the main text, which contains the main results. The complete derivations are reported in the Appendices and in the Supplemental Material (SM) [44], to which we provide detailed reference at each step.

II. INHOMOGENEOUS STRAIN

The first step consists of the introduction of a curvilinear coordinate system, which allows us to express the nuclei potential in the strained system as a perturbative expansion in the strain tensor. This approach, introduced by BP for the case

[*]Contact author: andrea.secchi@nano.cnr.it

of homogeneous strain and generalized to the inhomogeneous case by Zhang [39], is recalled here for the reader's convenience. In the original Cartesian reference frame, let r_C define the electronic coordinates, while $\boldsymbol{R}_{i,0}$ and $\boldsymbol{R}_i \equiv \boldsymbol{R}_{i,0} + \boldsymbol{u}_i$ are the nuclei positions in the absence and in the presence of strain, respectively. Generalizing BP's approach to the case of inhomogeneous strain, one introduces a set of curvilinear coordinates $\boldsymbol{r}$ to describe the electronic position. These are related to the $\boldsymbol{r}_C$ by the equation [39]:

$$r_C^\alpha = r^\alpha + u^\alpha(\boldsymbol{r}), \tag{1}$$

where the Greek indices label vector components ($\alpha = 1, 2, 3$). The continuous inhomogeneous displacement $\boldsymbol{u}(\boldsymbol{r})$ is assumed to be an invertible and differentiable function of $\boldsymbol{r}$ to all needed orders. It fulfills the conditions $\boldsymbol{u}(\boldsymbol{R}_{i,0}) = \boldsymbol{u}_i$ and $|\boldsymbol{u}(\boldsymbol{r})| \ll |\boldsymbol{r}|$. The former condition allows for the expansion of the nuclei potential in powers of the strain in the transformed coordinate frame, while the latter condition follows from the assumption that the strain tensor is small everywhere.

Given the displacement functions u^α, the components of the strain tensor can be defined as follows:

$$\varepsilon_\beta^\alpha(\boldsymbol{r}) \equiv \partial_\beta u^\alpha(\boldsymbol{r}), \tag{2}$$

where $\partial_\beta \equiv \partial/\partial r^\beta$. Provided that the strain tensor is small and varies slowly over the scale of a unit cell, by applying the transformation in Eq. (1), one can express the potential U_n generated by the nuclei in the strained system in the form:

$$U_n(\boldsymbol{r}) \approx U_{n,0}(\boldsymbol{r}) + \varepsilon_\beta^\alpha(\boldsymbol{r}) U_\alpha^\beta(\boldsymbol{r}). \tag{3}$$

Here, $U_{n,0}$ is the potential in the unstrained system and U_α^β is a strain-independent function that has the same periodicity as the unstrained lattice, while the product $\varepsilon_\beta^\alpha(\boldsymbol{r}) U_\alpha^\beta(\boldsymbol{r})$ is in general not periodic. Further details on the nuclei potential relations are provided in Appendix A. In Appendix B, instead, we give a formal derivation of the relations between Hamiltonians and wave functions in two coordinate systems connected by a point transformation such as Eq. (1). Here and in the remainder of this Article, we use Einstein's summation convention on repeated Greek indexes.

III. THE SCHRÖDINGER PROBLEM IN CURVILINEAR COORDINATES

The second step consists in deriving the general expression of the Schrödinger equation for an electron in curvilinear coordinates, with the inclusion of the spin-orbit term. The adoption of the curvilinear coordinates $\boldsymbol{r}$ implies the introduction of a nontrivial metric tensor, i.e., a $g_{\mu\nu} \neq -\delta_{\mu\nu}$ [41]. As a result, the matrix element of a local operator $\hat{A}$ between two arbitrary electron (spinorial) states is given by

$$\langle \Psi | \hat{A} | \Phi \rangle = \int d\boldsymbol{r} \sqrt{-g(\boldsymbol{r})} \, \Psi^\dagger(\boldsymbol{r}) \cdot A(\boldsymbol{r}) \Phi(\boldsymbol{r}), \tag{4}$$

where $g(\boldsymbol{r}) = \det[g_{\mu\nu}(\boldsymbol{r})]$. The definition of inner products can be obtained from the above equation simply by replacing the generic operator A with the identity. As a technical but crucial point, we note that, in a curvilinear coordinate system, the definition in Eq. (4) should be used in evaluating the Hermiticity of operators and the scalar products between states, rather than its Cartesian counterpart, corresponding to $\sqrt{-g(\boldsymbol{r})} = 1$ [39].

In order to obtain the correct Hamiltonian in curvilinear coordinates and to include spin-orbit coupling, we generalize the covariant formulation of the Schrödinger equation given in Ref. [41], which applies to a non-relativistic Hamiltonian. Starting from the covariant Dirac equation for the four-component electron field in an electromagnetic potential [43], which holds for arbitrary metric tensors, we take the nonrelativistic limit and allow for a nontrivial metric in the spatial sector only. The result is a Schrödinger equation for two spinors that can be augmented with any order of relativistic corrections, while inheriting the covariance of the initial Dirac equation.

In the absence of magnetic field and up to the first order in the relativistic corrections, the Hamiltonian can be written as: $H = H_{kin} + H_{rel} + U$, where the kinetic term reads

$$H_{kin} = -\frac{\hbar^2}{2m}\left[(\nabla_C^2 r^\nu)\partial_\nu - g^{\mu\nu}\partial_\mu \partial_\nu\right], \tag{5}$$

U is a generic scalar potential, and the dominant component in the relativistic term is given by the spin-orbit Hamiltonian

$$H_{so} = -\frac{i\hbar^2}{4m^2c^2}\left(\frac{\partial r^\mu}{\partial r_C^\alpha}\frac{\partial r^\nu}{\partial r_C^\beta}\sigma^{\alpha\beta}\right)(\partial_\mu U)\partial_\nu. \tag{6}$$

Here, we adopt the notation $\nabla_C^2 \equiv \partial^2/(\partial r_C^\alpha \partial r_C^\alpha)$. Besides, $\sigma^{\alpha\beta} = -\widetilde{\epsilon}^{\alpha\beta\gamma}\sigma_\gamma$, where $\widetilde{\epsilon}^{\alpha\beta\gamma}$ is the invariant Levi-Civita symbol and $-\sigma_\gamma$ are the Pauli matrices.

Relying on the generalized expression of the matrix elements and of the inner product [Eq. (4)], one can write the matrix elements of the Hamiltonian between two-component spinors in a manifestly Hermitian way, assuming that the wave functions either vanish at infinity, or satisfy the Born-von Karman boundary conditions. Further details on the relativistic terms of the Hamiltonian and on the boundary conditions are provided in Appendixes C and D, respectively.

IV. CURVILINEAR COORDINATES FROM THE STRAIN TENSOR

The equations reported in the previous paragraph provide a general framework, which can be applied to the present problem, where the non-Cartesian character of the coordinates results from the presence of inhomogeneous strain. In fact the strain tensor determines the Jacobian $J_\beta^\alpha \equiv \partial_\beta \partial r_C^\alpha = \delta_\beta^\alpha + \varepsilon_\beta^\alpha$, as can be deduced from Eqs. (1) and (2). As to the metric tensor, to first order in the strain tensor, it can be expressed as

$$g^{\mu\nu} \approx -\delta^{\mu\nu} - \varepsilon_\mu^\nu - \varepsilon_\nu^\mu, \tag{7}$$

under the assumption that $\|\varepsilon_i(\boldsymbol{r})\| \ll 1$, and thus $J^{-1} \approx 1 - \varepsilon$. From this it also follows that $\sqrt{-g} \approx 1 + \text{tr}(\varepsilon)$.

The above relations provide the dependence of H_{kin} and H_{so} on the strain tensor, mediated by the inverse Jacobian, ∇_C^2 and $g^{\mu\nu}$. To the first order in the strain tensor, the kinetic and spin-orbit components of the Hamiltonian can thus be written as follows:

$$H_{kin} = -\frac{\hbar^2}{2m}\left[\frac{\partial^2}{\partial r^\mu \partial r^\mu} + \overleftarrow{\partial}_\mu(\varepsilon_\mu^\nu + \varepsilon_\nu^\mu)\overrightarrow{\partial}_\nu\right], \tag{8}$$

$$H_{so} = -i\eta(f_\mu\sigma^{\mu\nu}\overrightarrow{\partial}_\nu - \overleftarrow{\partial}_\nu\sigma^{\mu\nu}f_\mu + \Sigma^\nu\overrightarrow{\partial}_\nu - \overleftarrow{\partial}_\nu\Sigma^\nu), \tag{9}$$

where

$$\eta \equiv \hbar^2/(8m^2c^2),$$

$$f_\mu \equiv \partial_\mu U_{n,0},$$

$$\Sigma^\nu \equiv \left(\partial_\mu \varepsilon_\beta^\alpha\right) U_\alpha^\beta \sigma^{\mu\nu} + \varepsilon_\beta^\alpha \left(\partial_\mu U_\alpha^\beta\right) \sigma^{\mu\nu} - \varepsilon_\beta^\nu f_\alpha \sigma^{\alpha\beta} - \varepsilon_\mu^\alpha f_\alpha \sigma^{\mu\nu}, \tag{10}$$

and the arrows above the differential operators specify whether these must be applied to the wave function on the left or right sides of the Hamiltonian when evaluating its matrix elements.

As to the potential induced by the nuclei, its expression is given by the sum of the unstrained contribution and of a perturbation that depends linearly on the strain tensor [Eq. (3)].

The derivations of the above equations can be found in Appendixes C and D.

V. GENERALIZED LUTTINGER-KOHN THEORY

In the LK solution scheme, the electron Hamiltonian matrix is derived in an orthonormal basis, and then reduced to a block structure by means of a suitable canonical transformation, which effectively separates the relevant manifold from the others, while perturbatively accounting for the intermanifold coupling. In the present section, this procedure is generalized in order to include the case of a curvilinear set of coordinates, with consistently defined orthonormality relations.

In order to identify a complete basis set, we initially consider the part of the Hamiltonian that is of order zero in the strain. Because this is a periodic function of $\mathbf{r}$, one can apply Bloch's theorem in order to derive its eigenfunctions $\psi_{n,\mathbf{k}} = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n,\mathbf{k}}(\mathbf{r})$ and eigenvalues $E_n(\mathbf{k})$. In view of an expansion around, e.g., the $\mathbf{\Gamma}$ point, it is convenient to introduce also the LK functions [30] $\chi_{n,\mathbf{k}} \equiv e^{i\mathbf{k}\cdot\mathbf{r}} u_{n,0}(\mathbf{r})$. In the curvilinear coordinates neither the Bloch nor the LK functions form an orthonormal basis [45], according to the inner product defined in Eq. (4). However, the orthonormality relations can be recovered by suitably modifying the LK functions, according to the relations [38]:

$$\overline{\chi}_{n,\mathbf{k}} \equiv \frac{\chi_{n,\mathbf{k}}}{[-g(\mathbf{r})]^{1/4}} \approx \left[1 - \frac{1}{2}\text{tr}\,\varepsilon(\mathbf{r})\right] e^{i\mathbf{k}\cdot\mathbf{r}} u_{n,0}(\mathbf{r}). \tag{11}$$

Analogous modifications can be applied in order to recover the orthonormality relations for the Bloch functions.

In the modified LK basis, the matrix elements of the first-order strain-dependent component of the Hamiltonian read:

$$\langle \overline{\chi}_{n,\mathbf{k}} | \hat{H}^{(1)} | \overline{\chi}_{n',\mathbf{k}'} \rangle = -\frac{\hbar^2}{4m} |\mathbf{k} - \mathbf{k}'|^2 \widetilde{\varepsilon}_\mu^\mu(\mathbf{k} - \mathbf{k}') \delta_{n,n'} + \widetilde{\varepsilon}_\nu^\mu(\mathbf{k} - \mathbf{k}') \left(\mathcal{D}_\mu^\nu + k^\alpha \mathcal{L}_{\alpha;\mu}^\nu + k'^\alpha \mathcal{L}_{\alpha;\mu}^{*\nu} + k^\alpha k'^\beta \mathcal{Q}_{\alpha\beta;\mu}^\nu\right)_{n,n'}, \tag{12}$$

where $\widetilde{\varepsilon}_\nu^\mu(\mathbf{q})$ is the Fourier transform of the strain tensor. The last parentheses on the right include the deformation-potential terms; in particular, the $\mathbf{k}$- independent quantities $\mathcal{D}$, $\mathcal{L}$, and $\mathcal{Q}$ are given, respectively, by

$$\mathcal{D}_\mu^\nu \equiv D_\mu^\nu + \Delta D_\mu^\nu, \tag{13a}$$

$$\mathcal{L}_{\alpha;\mu}^\nu \equiv L_{\alpha;\mu}^\nu + \Delta L_{\alpha;\mu}^\nu, \tag{13b}$$

$$\mathcal{Q}_{\alpha\beta;\mu}^\nu \equiv Q_{\alpha\beta;\mu}^\nu + \Delta Q_{\alpha\beta;\mu}^\nu, \tag{13c}$$

where the dominant nonrelativistic components are

$$D_\mu^\nu \equiv U_\mu^\nu - \frac{1}{m} p_\mu p_\nu, \tag{14a}$$

$$L_{\alpha;\mu}^\nu \equiv -\frac{\hbar}{2m} \left(\delta_\alpha^\nu p_\mu + \delta_\alpha^\mu p_\nu\right), \tag{14b}$$

$$Q_{\alpha\beta;\mu}^\nu \equiv -\frac{\hbar^2}{2m} \left(\delta_\alpha^\mu \delta_\beta^\nu + \delta_\alpha^\nu \delta_\beta^\mu\right), \tag{14c}$$

while the relativistic corrections read

$$\Delta D_\mu^\nu \equiv i\eta \left(2 \overleftarrow{\partial}_\alpha \sigma^{\alpha\beta} U_\mu^\nu \overrightarrow{\partial}_\beta + \overleftarrow{\partial}_\beta \sigma^{\nu\mu} f_\beta \right.$$
$$\left. - \overleftarrow{\partial}_\beta \sigma^{\nu\beta} f_\mu - \sigma^{\nu\mu} f_\beta \overrightarrow{\partial}_\beta + \sigma^{\nu\beta} f_\mu \overrightarrow{\partial}_\beta\right), \tag{15a}$$

$$\Delta L_{\alpha;\mu}^\nu \equiv \eta \left(2\sigma^{\alpha\beta} U_\mu^\nu \overrightarrow{\partial}_\beta + \sigma^{\nu\mu} f_\alpha - \sigma^{\nu\alpha} f_\mu\right), \tag{15b}$$

$$\Delta Q_{\alpha\beta;\mu}^\nu \equiv i\eta \sigma^{\alpha\beta} U_\mu^\nu, \tag{15c}$$

in terms of the quantities defined in Eqs. (10). In the above, $p_\mu = -i\hbar\partial_\mu$, and the quantities depending on the band indices

n and n' in Eq. (12) are defined by the relation

$$(A)_{n,n'} \equiv \frac{(2\pi)^3}{\Omega_{\text{cry}}} \int d\mathbf{r}\, u_{n,0}^\dagger(\mathbf{r}) \cdot A\, u_{n',0}(\mathbf{r}), \tag{16}$$

where Ω_{cry} is the crystal (i.e., the normalization) volume. The next step consists of decoupling the low-energy manifold of interest ($n \leqslant N$) from the higher-energy states ($n > N$), using Löwdin partitioning [31,32,39,46]. This amounts to applying a canonical transformation to the Hamiltonian, $\hat{\mathcal{H}} = e^{-\hat{S}} \hat{H} e^{\hat{S}} \equiv \hat{H} + \Delta\hat{H}$, and to its eigenstates, $|\phi\rangle = e^{-\hat{S}} |\psi\rangle$. The transformation is such that $\hat{\mathcal{H}}$ is approximately block diagonal in band space, and specifically displays negligible coupling terms between the relevant manifold and the remote bands. If intermanifold couplings related to the deformation-potential terms can be neglected, then the correction $\Delta\hat{H}$ for the N-dimensional low-energy manifold has the standard effective-mass form [31].

Further technical details on the derivation of Eq. (12) are given in Appendix E and Sec. I of the SM [44]. Details on the manifold decoupling are presented in Sec. II of the SM [44].

VI. ENVELOPE FUNCTIONS

The last step consists of the derivation of the confined particle states within the relevant N-dimensional manifold. The electron Hamiltonian includes an external confining potential U_{ext}, such as that induced by metallic gates in electrostatically defined quantum dots, which adds to the nuclear contribution: $U = U_{\text{n}} + U_{\text{ext}}$. We remark that, for consistency, the external

potential must be expressed in the curvilinear coordinates $\boldsymbol{r}$; i.e., if the potential is initially known as a function $U_{\text{ext;C}}(\boldsymbol{r}_{\text{C}})$ of the Cartesian electronic coordinates, then the expression to be used here is $U_{\text{ext}}(\boldsymbol{r}) = U_{\text{ext;C}}[\boldsymbol{r} + \boldsymbol{u}(\boldsymbol{r})]$. Therefore, this quantity depends on the strain tensor through $\boldsymbol{u}(\boldsymbol{r})$. The external potential is assumed to be a slowly varying function of $\boldsymbol{r}$ on the scale of the lattice constant, so as to justify an envelope-function approach. In particular, the eigenfunctions of $\mathcal{H}$ can be written as

$$\phi(\boldsymbol{r}) \equiv \langle \boldsymbol{r} | \phi \rangle = \frac{1}{[-g(\boldsymbol{r})]^{1/4}} \sum_{n \leqslant N} F_n(\boldsymbol{r}) \, u_{n,\boldsymbol{0}}(\boldsymbol{r}), \qquad (17)$$

where the N quantities denoted as $F_n(\boldsymbol{r})$ are the unknown envelope functions. These are determined by diagonalizing in band and position spaces the envelope-function Hamiltonian $\hat{\mathcal{H}}_{\text{EF}} = \hat{\mathcal{H}}_{\text{EF}}^{(0)} + \hat{\mathcal{H}}_{\text{EF}}^{(1)}$, where $\hat{\mathcal{H}}_{\text{EF}}^{(0)}$ is formally the standard $\boldsymbol{k} \cdot \boldsymbol{p}$ term in the $\boldsymbol{r}$ coordinates, and $\hat{\mathcal{H}}_{\text{EF}}^{(1)}$ is given by

$$\hat{\mathcal{H}}_{\text{EF}}^{(1)} = \varepsilon_\nu^\mu(\boldsymbol{r}) \big[\mathcal{D}_\mu^\nu + (\mathcal{L}_{\alpha;\mu}^\nu + \mathcal{L}_{\alpha;\mu}^{*\nu}) \hat{k}_\alpha + \mathcal{Q}_{\alpha\beta;\mu}^\nu \hat{k}_\alpha \hat{k}_\beta \big]$$
$$- \mathrm{i} \big[\partial_\alpha \varepsilon_\nu^\mu(\boldsymbol{r}) \big] \big(\mathcal{L}_{\alpha;\mu}^\nu + \mathcal{Q}_{\alpha\beta;\mu}^\nu \hat{k}_\beta \big) + \frac{\hbar^2}{4m} \big[\nabla^2 \varepsilon_\mu^\mu(\boldsymbol{r}) \big] \mathbf{1},$$
$$(18)$$

with $\hat{k}_\alpha \equiv -\mathrm{i} \partial_\alpha$. Here $\mathcal{D}$, $\mathcal{L}$, and $\mathcal{Q}$ are matrices in band space, whose elements are defined according to Eqs. (16) and (13)-(15); $\mathbf{1}$ is the identity matrix in band space.

It should be emphasized that Eq. (17) is the spinor wave function in the curvilinear reference frame $\boldsymbol{r}$, while the wave function in the Cartesian frame is given by $\phi_{\text{C}}(\boldsymbol{r}_{\text{C}}) = \phi[\boldsymbol{r}(\boldsymbol{r}_{\text{C}})]$, where $\boldsymbol{r}(\boldsymbol{r}_{\text{C}})$ is the inverse of the transformation Eq. (1). Further details on the derivation of the envelope-function Hamiltonian are given in Appendix F.

Equation (18) is the main result of this work. It contains terms that have not been considered in the literature, and that cannot be inferred from the homogeneous case by replacing a constant strain tensor with a position-dependent one. These terms can be either intra- ($n = n'$) or interband ($n \neq n'$), and depend on the first or on the second spatial derivatives of the strain tensor. Besides, the terms linear in $\varepsilon(\boldsymbol{r})$ and $\propto \hat{k}_\alpha$ or $\propto \hat{k}_\alpha \hat{k}_\beta$ are non-zero also in the case of homogeneous strain, but have been neglected in previous analyses.

VII. VALENCE STATES IN DIAMOND STRUCTURES

As a specific but practically relevant application, we consider the valence bands of a crystal with diamond structure, such as silicon or germanium. In this case, the three relevant orbital states are built from p-type atomic orbitals, and thus carry an angular momentum $l = 1$. This, combined with the $s = 1/2$ spin of the electron, gives rise to a $j = 3/2$ quartet and a $j = 1/2$ doublet ($N = 6$). In this basis, the dominant part of the spin-orbit is diagonal, and gives rise to a splitting Δ_{SO} between the $j = 3/2$ and $j = 1/2$ states at the $\boldsymbol{\Gamma}$ point [47]. We here discuss the Hamiltonian that is obtained after neglecting the other spin-orbit contributions. Then: $\mathcal{L}$ coincides with L, which vanishes at the band maximum; $\mathcal{D}$ reduces to the nonrelativistic deformation potentials D; $\mathcal{Q}$ equals Q, which consists of purely intraband ($n = n'$) contributions. The strain-dependent component of the envelope-function

Hamiltonian matrix, Eq. (18), thus becomes

$$\hat{\mathcal{H}}_{\text{EF}}^{(1)} = \left\{ \frac{\hbar^2}{4m} \big[\nabla^2 \varepsilon_\mu^\mu(\boldsymbol{r}) \big] - \frac{\hbar^2}{m} \varepsilon_{\alpha\beta}^{\text{sym}}(\boldsymbol{r}) \hat{k}_\alpha \hat{k}_\beta \right.$$
$$\left. - \frac{\hbar^2}{m} \big[\hat{k}_\alpha \varepsilon_{\alpha\beta}^{\text{sym}}(\boldsymbol{r}) \big] \hat{k}_\beta \right\} \mathbf{1} + \varepsilon_\nu^\mu(\boldsymbol{r}) \boldsymbol{D}_\mu^\nu. \qquad (19)$$

Here, $\varepsilon_{\alpha\beta}^{\text{sym}} \equiv (\varepsilon_\beta^\alpha + \varepsilon_\alpha^\beta)/2$ is the symmetric part of the strain tensor, and $\boldsymbol{D}_\mu^\nu$ is the matrix of nonrelativistic deformation potentials.

The three main new terms in Eq. (19) are diagonal in the band index. The first one, $\propto \nabla^2 \varepsilon_\mu^\mu(\boldsymbol{r})$, is a function of coordinates only, hence it represents an effective correction to the electrostatic potential, which might affect the hole confinement in a nanostructure. Now, the trace of the strain tensor is proportional to that of the stress tensor, and it can be shown [48] that the Laplacian of the latter vanishes in an *isotropic* system. However, crystals are not isotropic, and the deviation from isotropy is responsible for the existence of this term. The second term, $\propto \hat{k}_\alpha \hat{k}_\beta$, represents a spatially dependent correction to the effective-mass terms, where the spatial dependence is due to the inhomogeneity of the strain tensor. This term is also nonzero in the case of an homogeneous strain; however, it has been ignored so far. The third term, $\propto \hat{k}_\beta$, has a completely new form, being linear in the momentum operator. It is not analogous to a spin-orbit coupling, since it does not couple different bands.

In order to provide an order-of-magnitude estimate of the new terms, we consider the stress tensor used in Ref. [49], adapted to describe a MOSFET with two stressors placed on top of the source and drain regions [22] (details are given in Appendix G). We find that, in such a system, the typical energy scales characterizing the new terms is 10^{-2} meV, corresponding to frequencies of the order of $2 - 5$ GHz. In particular, the terms quadratic in the momentum induce a spatial modulation of the hole effective masses in Si of approximately $1\% - 2\%$. The hard-to-control modulation of the confinement induced by the term $\propto \nabla^2 \varepsilon_\mu^\mu(\boldsymbol{r})$ is small and mostly relevant far from the center of the channel (close to the stressors); therefore, it likely has a small impact on the effective confinement. However, the impact of all the derived terms is ultimately device and material dependent: they should thus be accounted for in order to obtain an accurate modeling of all nanostructures where the strain tensor varies on the length scale of the particle wave function.

VIII. CONCLUSIONS

By combining solid-state theory and relativistic quantum mechanics in a non-Cartesian geometry, we have derived the envelope-function Hamiltonian for a general semiconductor nanostructure subjected to a small and slowly varying inhomogeneous strain. Our theory requires, as an input, the strain tensor, which can be computed for each given device via finite-element methods based on the minimization of the elastic energy density [28]. Numerical calculations of the electron/hole states based on our theory are expected to provide an accurate modeling of the effects of inhomogeneous strain on quantum-dot spin qubits. In particular, they will

allow to engineer spin-orbit interactions and *g*-tensor modulations aimed at improving the qubits' manipulability.

APPENDIX A: INHOMOGENEOUS STRAIN AND EXPANSION OF THE NUCLEI POTENTIAL

In the literature, there are two definitions of the strain tensor. The first one, which we adopt in the present work, is given by Eq. (2) and is consistent with that used in other envelope-function treatments [31,33,39]. The strain tensor reported in the second definition is the symmetrized version of that given in the first one:

$$\varepsilon_{\alpha\beta}^{\text{sym}}(\boldsymbol{r}) \equiv \frac{1}{2}\left[\frac{\partial u^{\alpha}(\boldsymbol{r})}{\partial r^{\beta}} + \frac{\partial u^{\beta}(\boldsymbol{r})}{\partial r^{\alpha}}\right] = \frac{1}{2}\left[\varepsilon_{\beta}^{\alpha}(\boldsymbol{r}) + \varepsilon_{\alpha}^{\beta}(\boldsymbol{r})\right]. \quad (A1)$$

This is the quantity that enters the expression of the infinitesimal variation in the distance between two points, in going from an unstrained to a strained system [4]. The two definitions do not necessarily coincide, since in general $\varepsilon_{\beta}^{\alpha}(\boldsymbol{r}) \neq \varepsilon_{\alpha}^{\beta}(\boldsymbol{r})$. Within the second convention, the quantity introduced in Eq. (2) is termed the *deformation tensor*, the symmetric quantity in Eq. (A1) is termed the *strain tensor*, while the antisymmetric combination

$$\varepsilon_{\alpha\beta}^{\text{antisym}}(\boldsymbol{r}) \equiv \frac{1}{2}\left[\varepsilon_{\beta}^{\alpha}(\boldsymbol{r}) - \varepsilon_{\alpha}^{\beta}(\boldsymbol{r})\right]. \quad (A2)$$

is called the *rotation tensor*.

The formal solution of Eq. (2), as noticed also in Ref. [39], is

$$u^{\alpha}(\boldsymbol{r}) = u^{\alpha}(\boldsymbol{r}_0) + \int_{\boldsymbol{r}_0}^{\boldsymbol{r}} \varepsilon_{\beta}^{\alpha}(\boldsymbol{r}')dr'^{\beta}, \quad (A3)$$

where the line integral is performed over any path connecting $\boldsymbol{r}_0$ to $\boldsymbol{r}$. In the specific case of an homogeneous strain, considered by BP [33], this reduces to

$$u^{\alpha}(\boldsymbol{r}) = \varepsilon_{\beta}^{\alpha} r^{\beta}. \quad (A4)$$

In the absence of strain, the potential generated by the nuclei is periodic, and is given by

$$U_{\text{n},0}(\boldsymbol{r}_{\text{C}}) = \sum_{i} U_{1\text{n}}(\boldsymbol{r}_{\text{C}} - \boldsymbol{R}_{i,0}). \quad (A5)$$

The same quantity can be rewritten as

$$U_{\text{n},0}(\boldsymbol{r}_{\text{C}}) = \sum_{i} \overline{U}_{1\text{n}}(\boldsymbol{r}_{\text{C}} - \boldsymbol{R}_{i,0})\,\Theta[\boldsymbol{r}_{\text{C}} \in \mathcal{C}_0(\boldsymbol{R}_{i,0})], \quad (A6)$$

where $\overline{U}_{1\text{n}}$ is a pseudopotential, and the contribution due to the nucleus at $\boldsymbol{R}_{i,0}$ goes to zero outside the unit cell $\mathcal{C}_0(\boldsymbol{R}_{i,0})$, centered on the same nucleus. This amounts to a mere resummation of contributions due to all nuclei, and can be done in the strained system as well. In the rigid-ion approximation,

with reference to the strained unit cells $\mathcal{C}[\boldsymbol{R}_{i,0} + \boldsymbol{u}(\boldsymbol{R}_{i,0})]$, the nuclei potential is written as

$$U_{\text{n;C}}(\boldsymbol{r}_{\text{C}}) = \sum_{i} \overline{U}_{1\text{n}}[\boldsymbol{r}_{\text{C}} - \boldsymbol{R}_{i,0} - \boldsymbol{u}(\boldsymbol{R}_{i,0})]$$

$$\times \Theta\{\boldsymbol{r}_{\text{C}} \in \mathcal{C}[\boldsymbol{R}_{i,0} + \boldsymbol{u}(\boldsymbol{R}_{i,0})]\}. \quad (A7)$$

After the coordinate transformation $\boldsymbol{r}_{\text{C}} = \boldsymbol{r} + \boldsymbol{u}(\boldsymbol{r}) \equiv \boldsymbol{t}(\boldsymbol{r})$, setting $U_{\text{n;C}}(\boldsymbol{r}_{\text{C}}) = U_{\text{n;C}}[\boldsymbol{t}(\boldsymbol{r})] \equiv U_{\text{n}}(\boldsymbol{r})$, one obtains

$$U_{\text{n}}(\boldsymbol{r}) = \sum_{i} \overline{U}_{1\text{n}}[\boldsymbol{r} + \boldsymbol{u}(\boldsymbol{r}) - \boldsymbol{R}_{i,0} - \boldsymbol{u}(\boldsymbol{R}_{i,0})]$$

$$\times \Theta\{\boldsymbol{t}(\boldsymbol{r}) \in \mathcal{C}[\boldsymbol{t}(\boldsymbol{R}_{i,0})]\}. \quad (A8)$$

Since the strain, besides being small, varies slowly within a unit cell, the condition imposed by the function Θ is satisfied only when $\boldsymbol{r}$ is close to $\boldsymbol{R}_{i,0}$ and, therefore, $\boldsymbol{u}(\boldsymbol{r}) - \boldsymbol{u}(\boldsymbol{R}_{i,0})$ is small. Therefore, the following approximation holds:

$$u^{\nu}(\boldsymbol{r}) - u^{\nu}(\boldsymbol{R}_{i,0}) \approx \frac{\partial u^{\nu}(\boldsymbol{r})}{\partial r^{\mu}}\left(r^{\mu} - R_{i,0}^{\mu}\right) \equiv \varepsilon_{\mu}^{\nu}(\boldsymbol{r})\left(r^{\mu} - R_{i,0}^{\mu}\right). \quad (A9)$$

From this it follows that

$$\overline{U}_{1\text{n}}[\boldsymbol{r} + \boldsymbol{u}(\boldsymbol{r}) - \boldsymbol{R}_{i,0} - \boldsymbol{u}(\boldsymbol{R}_{i,0})]$$

$$\approx \overline{U}_{1\text{n}}\left\{\left[\delta_{\beta}^{\alpha} + \varepsilon_{\beta}^{\alpha}(\boldsymbol{r})\right]\left(r^{\beta} - R_{i,0}^{\beta}\right)\right\}$$

$$\approx \overline{U}_{1\text{n}}(\boldsymbol{r} - \boldsymbol{R}_{i,0}) + \frac{\partial \overline{U}_{1\text{n}}(\boldsymbol{r} - \boldsymbol{R}_{i,0})}{\partial\left(r^{\alpha} - R_{i,0}^{\alpha}\right)}\,\varepsilon_{\beta}^{\alpha}(\boldsymbol{r})\left(r^{\beta} - R_{i,0}^{\beta}\right). \quad (A10)$$

Summing over the atoms, one obtains the total potential as in Eq. (3), where

$$U_{\alpha}^{\beta}(\boldsymbol{r}) \equiv \sum_{i} \frac{\partial \overline{U}_{1\text{n}}(\boldsymbol{r} - \boldsymbol{R}_{i,0})}{\partial\left(r^{\alpha} - R_{i,0}^{\alpha}\right)}\left(r^{\beta} - R_{i,0}^{\beta}\right) \quad (A11)$$

has the same periodicity as the unstrained lattice.

APPENDIX B: CONCEPTS RELATED TO POINT TRANSFORMATIONS

The Schrödinger equation,

$$\hat{H}|\psi\rangle = E|\psi\rangle, \quad (B1)$$

can be written in position representation after introducing a spatial coordinate frame. The purpose of this Appendix is to compare the representations corresponding to two different spatial coordinate frames $\mathbb{R}$ and $\mathbb{R}'$, connected by a spatial transformation. In the first (unprimed) frame, positions are measured by the coordinates $\boldsymbol{r}$, and a position eigenstate is defined by

$$\hat{\boldsymbol{r}}|\mathbb{R};\boldsymbol{r}\rangle = \boldsymbol{r}|\mathbb{R};\boldsymbol{r}\rangle, \quad (B2)$$

where $\hat{\boldsymbol{r}}$ is the position operator in $\mathbb{R}$. The decomposition of the identity in coordinate space is [41]

$$\int d\Omega|\mathbb{R};\boldsymbol{r}\rangle\langle\mathbb{R};\boldsymbol{r}| = \hat{1}, \quad (B3)$$

where $d\Omega \equiv d\boldsymbol{r}\sqrt{-g(\boldsymbol{r})}$ is the elementary volume in coordinate space, $g(\boldsymbol{r})$ being the determinant of the metric tensor.

The orthonormality relation between the position eigenstates in the $\mathbb{R}$ reference frame is

$$\langle \mathbb{R}; \boldsymbol{r}_1 | \mathbb{R}; \boldsymbol{r}_2 \rangle = \delta(\boldsymbol{r}_1, \boldsymbol{r}_2) \equiv \frac{1}{\sqrt{-g(\boldsymbol{r}_1)}} \delta(\boldsymbol{r}_1 - \boldsymbol{r}_2). \quad \text{(B4)}$$

Multiplying Eq. (B1) by $\langle \mathbb{R}; \boldsymbol{r} |$, one obtains the usual coordinate-space Schrödinger problem:

$$H(\boldsymbol{r}) \, \psi(\boldsymbol{r}) = E \, \psi(\boldsymbol{r}), \quad \text{(B5)}$$

where

$$\psi(\boldsymbol{r}) \equiv \langle \mathbb{R}; \boldsymbol{r} | \psi \rangle, \quad \langle \mathbb{R}; \boldsymbol{r} | \hat{H} | \psi \rangle \equiv H(\boldsymbol{r}) \, \psi(\boldsymbol{r}),$$

$$\langle \mathbb{R}; \boldsymbol{r}_1 | \hat{H} | \mathbb{R}; \boldsymbol{r}_2 \rangle \equiv \delta(\boldsymbol{r}_1, \boldsymbol{r}_2) \, H(\boldsymbol{r}_2), \quad \text{(B6)}$$

and the eigenstate of the Hamiltonian in Hilbert space representation is

$$|\psi\rangle = \int d\Omega \, |\mathbb{R}; \boldsymbol{r}\rangle \, \psi(\boldsymbol{r}). \quad \text{(B7)}$$

To solve the same Schrödinger problem as in Eq. (B1), one can equivalently choose the different coordinate frame $\mathbb{R}'$, with coordinates $\boldsymbol{r}'$, connected to the previous representation via $\boldsymbol{r} = \boldsymbol{t}(\boldsymbol{r}')$. It should be noted that this is a relation between the position *eigenvalues* measured in different reference frames on the same position eigenstate. This means that the following relations between the position eigenstates in the two reference frames hold:

$$|\mathbb{R}; \boldsymbol{t}(\boldsymbol{r}')\rangle = |\mathbb{R}'; \boldsymbol{r}'\rangle, \quad |\mathbb{R}; \boldsymbol{r}\rangle = |\mathbb{R}'; \boldsymbol{t}^{-1}(\boldsymbol{r})\rangle. \quad \text{(B8)}$$

In other words, if a position measurement on a position eigenstate gives the result $\boldsymbol{r}'$ in the reference frame $\mathbb{R}'$, then the result of the position measurement on the same state gives the result $\boldsymbol{t}(\boldsymbol{r}')$ in the reference frame $\mathbb{R}$:

$$\hat{\boldsymbol{r}} | \mathbb{R}; \boldsymbol{t}(\boldsymbol{r}') \rangle = \boldsymbol{t}(\boldsymbol{r}') | \mathbb{R}; \boldsymbol{t}(\boldsymbol{r}') \rangle, \quad \hat{\boldsymbol{r}}' | \mathbb{R}'; \boldsymbol{r}' \rangle = \boldsymbol{r}' | \mathbb{R}'; \boldsymbol{r}' \rangle, \quad \text{(B9)}$$

where $\hat{\boldsymbol{r}}'$ is the position operator in $\mathbb{R}'$.

In the primed representation, the identity decomposition is

$$\int d\Omega' \, |\mathbb{R}'; \boldsymbol{r}'\rangle \langle \mathbb{R}'; \boldsymbol{r}'| = \hat{1}, \quad \langle \mathbb{R}'; \boldsymbol{r}'_1 | \mathbb{R}'; \boldsymbol{r}'_2 \rangle = \delta'(\boldsymbol{r}'_1, \boldsymbol{r}'_2). \quad \text{(B10)}$$

The Schrödinger problem is represented in $\mathbb{R}'$ as

$$H'(\boldsymbol{r}') \, \psi'(\boldsymbol{r}') = E \, \psi'(\boldsymbol{r}'), \quad \text{(B11)}$$

with

$$|\psi\rangle = \int d\Omega' |\mathbb{R}'; \boldsymbol{r}'\rangle \psi'(\boldsymbol{r}'). \quad \text{(B12)}$$

One should now derive the relation between $\psi(\boldsymbol{r})$ and $\psi'(\boldsymbol{r}')$, and that between $H(\boldsymbol{r})$ and $H'(\boldsymbol{r}')$. The most direct way to do so is to consider the scalar product between any two states $|\Phi_1\rangle$ and $|\Phi_2\rangle$, which must be the same independently of the reference frame where it is evaluated:

$$\langle \Phi_1 | \Phi_2 \rangle = \int d\Omega \, [\Phi_1(\boldsymbol{r})]^* \, \Phi_2(\boldsymbol{r}) = \int d\Omega' \, [\Phi'_1(\boldsymbol{r}')]^* \, \Phi'_2(\boldsymbol{r}'). \quad \text{(B13)}$$

We now change variables in the first equality according to $\boldsymbol{r} = \boldsymbol{t}(\boldsymbol{r}')$, using the transformation property of the determinant of

the metric tensor,

$$\sqrt{-g'(\boldsymbol{r}')} = |J(\boldsymbol{r}')| \sqrt{-g[\boldsymbol{t}(\boldsymbol{r}')]}, \quad \text{(B14)}$$

where $J(\boldsymbol{r}')$ is the determinant of the Jacobian matrix. Therefore, under this change of coordinates it holds that

$$\int d\Omega \, \{\Phi_1(\boldsymbol{r})\}^* \, \Phi_2(\boldsymbol{r}) = \int d\Omega' \, \{\Phi_1[\boldsymbol{t}(\boldsymbol{r}')]\}^* \, \Phi_2[\boldsymbol{t}(\boldsymbol{r}')]$$

$$= \int d\Omega' \, [\Phi'_1(\boldsymbol{r}')]^* \, \Phi'_2(\boldsymbol{r}'). \quad \text{(B15)}$$

Since this must hold for any couple of states $|\Phi_1\rangle$ and $|\Phi_2\rangle$, one concludes that

$$\Phi'(\boldsymbol{r}') = \Phi[\boldsymbol{t}(\boldsymbol{r}')] \quad \text{if } \boldsymbol{r} = \boldsymbol{t}(\boldsymbol{r}'). \quad \text{(B16)}$$

Therefore,

$$\psi'(\boldsymbol{r}') = \psi[\boldsymbol{t}(\boldsymbol{r}')] \quad \text{if } \boldsymbol{r} = \boldsymbol{t}(\boldsymbol{r}'), \quad \text{(B17)}$$

which provides the relation between the wave functions in the two reference frames.

In order to derive an analogous relation between the Hamiltonians, let us consider again Eq. (B8). Multiplying both sides of the first equation by $\langle \mathbb{R}; \boldsymbol{r} |$, and using the orthonormality of the position eigenstates in $\mathbb{R}$, one obtains

$$\langle \mathbb{R}; \boldsymbol{r} | \mathbb{R}'; \boldsymbol{r}' \rangle = \delta[\boldsymbol{r}, \boldsymbol{t}(\boldsymbol{r}')] = \frac{\delta[\boldsymbol{r} - \boldsymbol{t}(\boldsymbol{r}')]}{\sqrt{-g(\boldsymbol{r})}}. \quad \text{(B18)}$$

Analogously, multiplying both sides of the second equation by $\langle \mathbb{R}'; \boldsymbol{r}' |$, and using the orthonormality of the position eigenstates in $\mathbb{R}'$, one obtains

$$\langle \mathbb{R}'; \boldsymbol{r}' | \mathbb{R}; \boldsymbol{r} \rangle = \delta'[\boldsymbol{r}', \boldsymbol{t}^{-1}(\boldsymbol{r})] = \frac{\delta[\boldsymbol{r}' - \boldsymbol{t}^{-1}(\boldsymbol{r})]}{\sqrt{-g'(\boldsymbol{r}')}}. \quad \text{(B19)}$$

The two transformation laws between position eigenstates in the two reference frames are equivalent, due to the composition law of a Dirac delta with a function. Equipped with Eqs. (B18) and (B19), it is now possible to compare the quantities

$$H(\boldsymbol{r}_1, \boldsymbol{r}_2) \equiv \langle \mathbb{R}; \boldsymbol{r}_1 | \hat{H} | \mathbb{R}; \boldsymbol{r}_2 \rangle \quad \text{(B20)}$$

and

$$H'(\boldsymbol{r}'_1, \boldsymbol{r}'_2) \equiv \langle \mathbb{R}'; \boldsymbol{r}'_1 | \hat{H} | \mathbb{R}'; \boldsymbol{r}'_2 \rangle. \quad \text{(B21)}$$

Inserting twice the decomposition of the identity in the original reference frame in Eq. (B21) and using the scalar product relations derived above, one obtains

$$H'(\boldsymbol{r}'_1, \boldsymbol{r}'_2)$$

$$= \int d\Omega_1 \int d\Omega_2 \, \langle \mathbb{R}'; \boldsymbol{r}'_1 | \mathbb{R}; \boldsymbol{r}_1 \rangle \, H(\boldsymbol{r}_1, \boldsymbol{r}_2) \, \langle \mathbb{R}; \boldsymbol{r}_2 | \mathbb{R}'; \boldsymbol{r}'_2 \rangle$$

$$= H[\boldsymbol{t}(\boldsymbol{r}'_1), \boldsymbol{t}(\boldsymbol{r}'_2)], \quad \text{(B22)}$$

which provides the relation between the Hamiltonians in the Cartesian and in the curvilinear coordinates.

APPENDIX C: DERIVATION OF THE COVARIANT SCHRÖDINGER EQUATION, WITH RELATIVISTIC CORRECTIONS

In this Appendix, Greek indexes refer to the components of vectors and tensors in a curvilinear reference frame, while

Latin indexes refer to such components in a Cartesian (rectilinear) reference frame, which coincides with the laboratory reference frame of the main text (note that in the main text Greek indexes are used for both frames). The four-component space-time coordinate is denoted as x in the curvilinear frame, and as x_C in the Cartesian frame. The metric tensor in the x frame is denoted as $g_{\mu\nu}$, while the metric tensor in the x_C frame is $\eta_{ab} = \text{diag}(1, -1, -1, -1)$. The relation between the two metric tensors is [43]

$$g_{\mu\nu} = e_\mu^a \, e_\nu^b \, \eta_{ab}, \tag{C1}$$

where the tetrad fields e_a^μ satisfy

$$e_a^\mu e_\nu^a = \delta_\nu^\mu, \quad e_a^\mu e_\mu^b = \delta_a^b. \tag{C2}$$

In the case at hand, since the two coordinate frames are connected by a point transformation, one has

$$e_\mu^a = \frac{\partial x_C^a}{\partial x^\mu}, \quad e_a^\mu = \frac{\partial x^\mu}{\partial x_C^a}, \tag{C3}$$

i.e., the tetrad vectors coincide with the Jacobian matrix of the transformation between the two coordinate systems. This ensures the conservation of the infinitesimal arc length squared,

$$\eta_{ab} \, dx_C^a \, dx_C^b = g_{\mu\nu} \, dx^\mu \, dx^\nu, \tag{C4}$$

which must be independent of the chosen coordinates.

1. Covariant Dirac equation

We start from the covariant Dirac equation [43] for the four-component electron field Ψ,

$$\gamma^\mu (i\hbar \nabla_\mu - e A_\mu)\Psi - mc\Psi = 0, \tag{C5}$$

where A_μ is the four-potential of the electromagnetic field, and ∇_μ is the covariant derivative; the latter acts on the four-spinor as follows:

$$\nabla_\mu \Psi \equiv \partial_\mu \Psi + \Gamma_\mu \Psi, \tag{C6}$$

where ∂_μ is the ordinary derivative, and

$$\Gamma_\mu \equiv \frac{1}{2}\eta_{ac} e_\nu^c \big(\partial_\mu e_b^\nu + e_b^\rho \Gamma_{\rho\mu}^\nu\big) G^{ab},$$
$$G^{ab} = \frac{1}{4}(\gamma^a \gamma^b - \gamma^b \gamma^a). \tag{C7}$$

The coordinate-invariant Dirac matrices satisfy

$$\gamma^a \gamma^b + \gamma^b \gamma^a = 2\eta^{ab}, \tag{C8}$$

while the spacetime-dependent Dirac matrices (entering the covariant Dirac equation) are defined as $\gamma^\mu = e_a^\mu \gamma^a$; thus, they satisfy

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}. \tag{C9}$$

In terms of the tetrads, the Christoffel symbols are written as:

$$\Gamma_{\rho\mu}^\nu \equiv \frac{1}{2} g^{\nu\beta}\big(\partial_\mu g_{\beta\rho} + \partial_\rho g_{\beta\mu} - \partial_\beta g_{\rho\mu}\big)$$
$$= \frac{1}{2} e_i^\nu \big(\partial_\mu e_\rho^i + \partial_\rho e_\mu^i\big) + \frac{1}{2}\eta^{kl}\eta_{ij} e_k^\nu e_l^\beta$$
$$\times \big[e_\rho^i\big(\partial_\mu e_\beta^j - \partial_\beta e_\mu^j\big) + e_\mu^i\big(\partial_\rho e_\beta^j - \partial_\beta e_\rho^j\big)\big]. \tag{C10}$$

In the case at hand, since Eq. (C3) holds, one has

$$\partial_\mu e_\rho^i = \frac{\partial^2 x_C^i}{\partial x^\mu \partial x^\rho} = \partial_\rho e_\mu^i, \tag{C11}$$

and Eq. (C10) simplifies as

$$\Gamma_{\rho\mu}^\nu = e_i^\nu\big(\partial_\rho e_\mu^i\big). \tag{C12}$$

The combinations needed for the covariant derivative are

$$\frac{1}{2}\eta_{ac} e_\nu^c e_b^\rho \Gamma_{\rho\mu}^\nu G^{ab} = \frac{1}{2}\eta_{ac} e_b^\rho\big(\partial_\rho e_\mu^c\big) G^{ab}, \tag{C13}$$

and

$$\gamma^\mu \Gamma_\mu = \frac{1}{2}\eta_{ac}\big[e_\nu^c (\partial_d e_b^\nu) + e_d^\mu\big(\partial_b e_\mu^c\big)\big]\gamma^d G^{ab}$$
$$= \frac{1}{2}\eta_{ac}\big[e_\mu^c\big(\partial_d e_b^\mu\big) - \big(\partial_b e_d^\mu\big)e_\mu^c\big]\gamma^d G^{ab} = 0, \tag{C14}$$

where we have used the fact that $\partial_\mu(e_\nu^c e_b^\nu) = 0$, and thus $e_\nu^c \partial_\mu e_b^\nu = -e_b^\nu \partial_\mu e_\nu^c$. Therefore, the covariant Dirac equation, when the tetrad coincides with the Jacobian matrix, reduces to

$$i\hbar e_n^\mu \gamma^n \partial_\mu \Psi - e e_n^\mu A_\mu \gamma^n \Psi - mc\Psi = 0. \tag{C15}$$

In the following, we choose the coordinate-invariant Dirac matrices in the Dirac form,

$$\gamma^0 = \begin{pmatrix} 1_{2\times2} & 0_{2\times2} \\ 0_{2\times2} & -1_{2\times2} \end{pmatrix}; \quad \gamma^a = \begin{pmatrix} 0_{2\times2} & \sigma^a \\ -\sigma^a & 0_{2\times2} \end{pmatrix}, \tag{C16}$$

where σ^a, with $a \in \{1, 2, 3\}$, are the Pauli matrices.

2. Restriction to spatial-only transformation

In the problem at hand, the coordinate transformation only involves the spatial coordinates ($\mu \in \{1, 2, 3\}$), affecting the corresponding sector of the metric tensor, while the time coordinate ($\mu = 0$) is untouched. Therefore, we specialize our treatment to the cases where $e_\mu^0 = \delta_\mu^0$ and $e_0^a = \delta_0^a$. All tetrads are independent of time.

It is then convenient to write the Dirac equation in a way that explicitly separates time from the spatial coordinates. Equation (C15) can thus be transformed into

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}i\hbar\partial_t\Psi = \begin{pmatrix} 0 & c\sigma^\mu P_\mu \\ -c\sigma^\mu P_\mu & 0 \end{pmatrix}\Psi$$
$$+ \begin{pmatrix} mc^2 + eV & 0 \\ 0 & mc^2 - eV \end{pmatrix}\Psi, \tag{C17}$$

where $x_0 = ct$, $\partial_0 = \frac{1}{c}\partial_t$, $A_0 = \frac{1}{c}V$, $P_\mu \equiv -i\hbar\partial_\mu + eA_\mu$, and the coordinate-dependent Pauli matrices are defined as

$$\sigma^\mu \equiv e_n^\mu \sigma^n. \tag{C18}$$

This is rephrased as an eigenproblem, by setting

$$\Psi \equiv e^{-iE_D t/\hbar}\begin{pmatrix} \Xi \\ \Phi \end{pmatrix}, \tag{C19}$$

where E_D is the (Dirac) energy eigenvalue, while Ξ and Φ are two-component spinors depending only on the spatial coordinates (eigenstates). This results in the following coupled equations for the two-component spinor eigenstates:

$$(E_D - mc^2 - eV)\Xi = c\sigma^\mu P_\mu \Phi,$$
$$(E_D + mc^2 - eV)\Phi = c\sigma^\mu P_\mu \Xi. \tag{C20}$$

The spinor Φ is obtained as a function of Ξ from the second equation; substituting the resulting expression in the first equation, one obtains an eigenvalue equation for the spinor Ξ alone:

$$\left[c\sigma^\mu P_\mu \frac{1}{2mc^2 + E - eV} c\sigma^\nu P_\nu + eV \right] \Xi = E\,\Xi, \quad (C21)$$

where the Schrödinger eigenenergy is $E \equiv E_D - mc^2$. This equation is exact.

3. Schrödinger equation with relativistic corrections, in curvilinear coordinates

In order to recover the Schrödinger equation and the lowest-order relativistic corrections, one must perform an approximation based on the assumption that $2mc^2 \gg |E - eV|$, namely,

$$\frac{1}{2mc^2 + E - eV} = \frac{1}{2mc^2} \left(1 + \frac{E - eV}{2mc^2} \right)^{-1}$$

$$\approx \frac{1}{2mc^2} - \frac{E - eV}{4m^2c^4}. \quad (C22)$$

Inserting this into Eq. (C21), one obtains

$$\left[\frac{(\sigma^\mu P_\mu)^2}{2m} - \sigma^\mu P_\mu \frac{E - eV}{4m^2c^2} \sigma^\nu P_\nu + eV \right] \Xi \approx E\,\Xi. \quad (C23)$$

To remove E from the left-hand side of the above equation (to lowest order in v/c), we notice that

$$(E - eV)\sigma^\nu P_\nu \Xi = \sigma^\nu P_\nu (E - eV)\Xi - i\hbar\sigma^\nu(\partial_\nu eV)\Xi$$

$$\approx \frac{(\sigma^\nu P_\nu)^3}{2m} \Xi - i\hbar\sigma^\nu(\partial_\nu eV)\Xi. \quad (C24)$$

Equation (C23) can thus be written as an eigenvalue equation

$$H\,\Xi \approx E\,\Xi, \quad (C25)$$

where the Hamiltonian is given by

$$H = \underbrace{\frac{(\sigma^\mu P_\mu)^2}{2m} + eV}_{\text{nonrelativistic}} \underbrace{- \frac{(\sigma^\mu P_\mu)^4}{8m^3c^2} + \frac{i\hbar(\sigma^\mu P_\mu)\sigma^\nu(\partial_\nu eV)}{4m^2c^2}}_{\text{relativistic corrections}}$$

$$\equiv H_{\text{nonrel}} + H_{\text{rel}}. \quad (C26)$$

4. Hamiltonian in curvilinear coordinates

Finally, by making the electromagnetic potentials explicit, and using the properties of the Pauli matrices, one can rewrite the terms H_{nonrel} and H_{rel} appearing in Eq. (C26) as follows:

$$H_{\text{nonrel}} = \frac{1}{2m} \big\{ -\hbar^2(-g^{\mu\nu})\partial_\mu\partial_\nu - \hbar^2(\nabla_C^2 x^\nu)\partial_\nu$$

$$+ i\hbar e\big[e_a^\mu(\partial_\mu e_\nu^a)A^\nu + (\partial_\mu A^\mu) + 2A^\mu\partial_\mu \big]$$

$$- e^2 A_\mu A^\mu - e\hbar\,\sigma_c B^c \big\} + eV, \quad (C27)$$

$$H_{\text{rel}} = -\frac{1}{8m^3c^2} \big\{ -\hbar^2(-g^{\mu\nu})\partial_\mu\partial_\nu - \hbar^2(\nabla_C^2 x^\nu)\partial_\nu$$

$$+ i\hbar e\big[e_a^\mu(\partial_\mu e_\nu^a)A^\nu + (\partial_\mu A^\mu) + 2A^\mu\partial_\mu \big]$$

$$- e^2 A_\mu A^\mu - e\hbar\,\sigma_c B^c \big\}^2$$

$$+ \frac{\hbar^2}{4m^2c^2}(\nabla_C^2 eV) - \frac{i\hbar}{4m^2c^2}(\partial_\mu eV)g^{\mu\nu}P_\nu$$

$$+ \frac{\hbar}{4m^2c^2}\big(e_a^\mu e_b^\nu \sigma^{ab} \big)(\partial_\mu eV)P_\nu. \quad (C28)$$

In the above expressions, $\sigma_a = \eta_{ab}\sigma^b$, where the implicit summation on the Latin indices only involves the three spatial coordinates. Besides, we have introduced $\sigma^{ab} \equiv -\widetilde{\epsilon}^{abc}\sigma_c$ and the components of the magnetic field with respect to the rectilinear reference frame, $B^c \equiv \widetilde{\epsilon}^{abc}(\partial_a A_b)$; here, $\widetilde{\epsilon}^{abc}$ is the totally antisymmetric Levi-Civita symbol, which, unlike the Levi-Civita tensor, takes the same values in all reference frames. Finally, we have used the symbol ∇_C^2, which can be converted to curvilinear coordinates through:

$$\nabla_C^2 = \partial_a\partial_a = -\partial_a\partial^a = -e_a^\alpha\partial_\alpha e_\beta^a\partial^\beta$$

$$= -e_a^\alpha\big(\partial_\alpha e_\beta^a\big)\partial^\beta - \partial_\alpha\partial^\alpha$$

$$= -e_a^\alpha\big(\partial_\alpha e_\beta^a\big)g^{\beta\gamma}\partial_\gamma - (\partial_\alpha g^{\alpha\beta})\partial_\beta - g^{\alpha\beta}\partial_\alpha\partial_\beta. \quad (C29)$$

Alternatively (and more simply), the relevant derivatives can be computed in the rectilinear coordinate frame first, and then the resulting expressions can be converted in curvilinear coordinates.

In the main text, we consider the case where there is no magnetic field ($A^\mu = 0$), and only spin-orbit coupling is retained among the relativistic corrections. In this case, the Hamiltonian simplifies to the sum of the kinetic term given in Eq. (5), the scalar potential $U \equiv eV$, and the spin-orbit term given in Eq. (6). For notational convenience and self-containedness, in the main text we have written the tetrads explicitly in terms of the inverse Jacobian matrix, and we have used Greek indices also for the components of the Cartesian coordinates; since such coordinates themselves are indicated explicitly by means of the subscript C, there is no ambiguity.

APPENDIX D: MATRIX ELEMENTS OF THE HAMILTONIAN IN A MANIFESTLY HERMITIAN FORM

Using the definition of the matrix elements in the presence of an arbitrary metric tensor [Eq. (4)], one can write the matrix elements of the Hamiltonian between two-component spinors $|\Xi_n\rangle$ in a manifestly Hermitian way, assuming that the wave functions either vanish at infinity, or satisfy the Born–von Karman boundary conditions (BvKBCs). The three terms of the Hamiltonian [see Eqs. (5) and (6)] give the following contributions:

$$\langle \Xi_n | \hat{H}_{\text{kin}} | \Xi_m \rangle = -\frac{\hbar^2}{2m} \int d\mathbf{r}\sqrt{-g}\,g^{\mu\nu}\frac{\partial \Xi_n^\dagger}{\partial r^\mu} \cdot \frac{\partial \Xi_m}{\partial r^\nu}, \quad (D1)$$

$$\langle \Xi_n | \hat{U} | \Xi_m \rangle = \int d\mathbf{r}\sqrt{-g}\,U\,\Xi_n^\dagger \cdot \Xi_m, \quad (D2)$$

$$\langle \Xi_n | \hat{H}_{\text{so}} | \Xi_m \rangle = -\frac{i\hbar^2}{8m^2c^2} \int d\mathbf{r}\sqrt{-g}\,\frac{\partial r^\mu}{\partial r_C^\alpha}\frac{\partial r^\nu}{\partial r_C^\beta}(\partial_\mu U)$$

$$\times [\Xi_n^\dagger \cdot \sigma^{\alpha\beta} \cdot (\partial_\nu \Xi_m) - (\partial_\nu \Xi_n^\dagger) \cdot \sigma^{\alpha\beta} \cdot \Xi_m]. \quad (D3)$$

These forms are obtained by applying partial integration, using the boundary conditions, and exploiting the properties of the metric tensor.

As an example, we show the explicit derivation of the first contribution. One starts from

$$\langle \Xi_n | \hat{H}_{\text{kin}} | \Xi_m \rangle$$

$$= -\frac{\hbar^2}{2m} \int d\mathbf{r} \sqrt{-g}\, \Xi_n^\dagger \cdot \left[(\nabla_{\text{C}}^2 r^\nu) - g^{\mu\nu} \frac{\partial}{\partial r^\mu} \right] \frac{\partial \Xi_m}{\partial r^\nu}, \quad \text{(D4)}$$

which follows from Eq. (5). Now, partial integration with respect to r^μ is applied to the second term of this integral; the boundary term vanishes due to the boundary conditions, and the remaining term is

$$\langle \Xi_n | \hat{H}_{\text{kin}} | \Xi_m \rangle = -\frac{\hbar^2}{2m} \int d\mathbf{r} \left[\sqrt{-g}(\nabla_{\text{C}}^2 r^\nu)\, \Xi_n^\dagger \right.$$
$$\left. - \frac{\partial(-g^{\mu\nu}\sqrt{-g}\, \Xi_n^\dagger)}{\partial r^\mu} \right] \cdot \frac{\partial \Xi_m}{\partial r^\nu}. \quad \text{(D5)}$$

The derivative with respect to r^μ at the second term inside the square brackets is carried on by applying the following identities:

$$\partial_\mu(-g^{\mu\nu}) = -\eta^{ab} e_\mu^c \left(\partial_c e_a^\mu \right) e_b^\nu + \left(\nabla_{\text{C}}^2 r^\nu \right),$$
$$\partial_\mu \sqrt{-g} = \sqrt{-g}\, e_a^\alpha \left(\partial_\mu e_\alpha^a \right),$$
$$g^{\mu\nu} e_a^\alpha \left(\partial_\mu e_\alpha^a \right) + \eta^{ab} \left(\partial_c e_a^\mu \right) e_\mu^c e_b^\nu = 0, \quad \text{(D6)}$$

which follow directly from the definitions of the metric tensor and of the tetrads, and from the fact that $\partial_a e_c^\alpha = \partial_c e_a^\alpha$ in the case at hand, because the tetrads are defined via a point transformation from a Cartesian reference frame. After this, one directly obtains Eq. (D1). Equation (D3) follows from a similar derivation.

The symmetry of the matrix elements in Eqs. (D1)–(D3) can be made explicit by formally rewriting the operators themselves as follows:

$$\hat{H}_{\text{kin}} = -\frac{\hbar^2}{2m} \overleftarrow{\partial}_\mu g^{\mu\nu} \overrightarrow{\partial}_\nu, \quad \text{(D7)}$$

$$\hat{H}_{\text{so}} = -\frac{i\hbar^2}{8m^2c^2} \left[\frac{\partial r^\mu}{\partial r_{\text{C}}^\alpha} \frac{\partial r^\nu}{\partial r_{\text{C}}^\beta} (\partial_\mu U) \sigma^{\alpha\beta} \overrightarrow{\partial}_\nu \right.$$
$$\left. - \overleftarrow{\partial}_\nu \frac{\partial r^\mu}{\partial r_{\text{C}}^\alpha} \frac{\partial r^\nu}{\partial r_{\text{C}}^\beta} (\partial_\mu U) \sigma^{\alpha\beta} \right], \quad \text{(D8)}$$

where the arrow above a derivative operator indicates the direction along which the derivative operator acts, when evaluating a matrix element; it is intended that, within this convention, the derivatives do *not* act on the metric factor $\sqrt{-g}$ inside the integrals. These two definitions are equivalent to Eqs. (5) and (6); the same convention is used in Eqs. (8) and (9), which are the expansions of Eqs. (D7) and (D8), respectively, to the first order in the strain tensor.

APPENDIX E: SECULAR EQUATION UP TO FIRST ORDER IN THE STRAIN TENSOR

We now derive the expressions of the matrix elements of the Hamiltonian up to first order in the strain components on the $\overline{\chi}$ basis, i.e. we derive Eq. (12) by evaluating the matrix elements of Eqs. (8) and (9). By construction, the $\overline{\chi}$

are orthonormal, i.e., they satisfy

$$\langle \overline{\chi}_{n,\mathbf{k}} | \overline{\chi}_{n',\mathbf{k}'} \rangle = \int d\mathbf{r} \sqrt{-g(\mathbf{r})}\, \overline{\chi}_{n,\mathbf{k}}^\dagger(\mathbf{r}) \cdot \overline{\chi}_{n',\mathbf{k}'}(\mathbf{r})$$
$$= \delta_{n,n'}\, \delta(\mathbf{k} - \mathbf{k}'). \quad \text{(E1)}$$

with the representation of the scalar product in a curvilinear reference frame, given by Eq. (4).

It is convenient to introduce the quantity

$$(\Psi | \hat{A} | \Phi) \equiv \int d\mathbf{r}\, \Psi^\dagger(\mathbf{r}) \cdot A(\mathbf{r})\, \Phi(\mathbf{r}), \quad \text{(E2)}$$

which, as mentioned in the main text, is *not* a scalar product in the curvilinear reference frame, which is given instead by Eq. (4). The quantity in Eq. (E2), nevertheless, will appear in the following derivation, due to the fact that the accuracy of the theory up to the first order in the strain tensor also requires the expansion of $\sqrt{-g(\mathbf{r})}$ in Eq. (4).

For the purposes of the present derivation, it is convenient to rewrite the Hamiltonian $H = H_{\text{kin}} + U_{\text{n}} + H_{\text{so}} + U_{\text{ext}}$ as $H = H_0 + H_1 + U_{\text{ext}}$, where

$$H_0 \equiv H_{\text{kin},0} + U_{\text{n},0} + H_{\text{so},0}$$
$$= -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial r^\mu \partial r_\mu} + U_{\text{n},0}$$
$$- \frac{i\hbar^2}{8m^2c^2} [(\partial_\mu U_{\text{n},0}) \sigma^{\mu\nu} \overrightarrow{\partial}_\nu - \overleftarrow{\partial}_\nu \sigma^{\mu\nu} (\partial_\mu U_{\text{n},0})] \quad \text{(E3)}$$

and

$$H_1 \equiv H_{\text{kin},1} + U_{\text{n},1} + H_{\text{so},1}$$
$$= \underbrace{-\frac{\hbar^2}{2m} \overleftarrow{\partial}_\mu \left(\varepsilon_\mu^\nu + \varepsilon_\nu^\mu \right) \overrightarrow{\partial}_\nu + \varepsilon_\beta^\alpha U_\alpha^\beta}_{H_{1,\text{nonrel}}}$$
$$\underbrace{- \frac{i\hbar^2}{8m^2c^2} (\Sigma^\nu \overrightarrow{\partial}_\nu - \overleftarrow{\partial}_\nu \Sigma^\nu)}_{H_{1,\text{so}}} \quad \text{(E4)}$$

collect the terms which are, respectively, independent of and linear in the strain tensor; U_{ext} is left untouched.

Using the equation

$$\frac{\partial \overline{\chi}_{n',\mathbf{k}'}}{\partial r^\nu} \approx \frac{\partial \chi_{n',\mathbf{k}'}}{\partial r^\nu} - \frac{1}{2} \frac{\partial \varepsilon_\gamma^\gamma}{\partial r^\nu} \chi_{n',\mathbf{k}'} - \frac{1}{2} \varepsilon_\gamma^\gamma \frac{\partial \chi_{n',\mathbf{k}'}}{\partial r^\nu}, \quad \text{(E5)}$$

and keeping the BvKBCs into account, one obtains:

$$\langle \overline{\chi}_{n,\mathbf{k}} | \hat{H}_0 | \overline{\chi}_{n',\mathbf{k}'} \rangle = \int d\mathbf{r} \sqrt{-g}\, \overline{\chi}_{n,\mathbf{k}}^\dagger H_0 \overline{\chi}_{n',\mathbf{k}'}$$
$$\approx (\chi_{n,\mathbf{k}} | \hat{H}_0 | \chi_{n',\mathbf{k}'}) - \frac{\hbar^2}{4m} \int d\mathbf{r} \frac{\partial \varepsilon_\gamma^\gamma}{\partial r^\mu} \frac{\partial (\chi_{n,\mathbf{k}}^\dagger \cdot \chi_{n',\mathbf{k}'})}{\partial r^\mu}$$
$$= (\chi_{n,\mathbf{k}} | \hat{H}_0 | \chi_{n',\mathbf{k}'}) + \frac{\hbar^2}{4m} \int d\mathbf{r} \frac{\partial^2 \varepsilon_\gamma^\gamma}{\partial r^\mu \partial r_\mu} (\chi_{n,\mathbf{k}}^\dagger \cdot \chi_{n',\mathbf{k}'})$$
$$= (\chi_{n,\mathbf{k}} | \hat{H}_0 | \chi_{n',\mathbf{k}'}) + \frac{\hbar^2}{4m} (\chi_{n,\mathbf{k}} | (\nabla^2 \text{tr}\varepsilon) | \chi_{n',\mathbf{k}'}), \quad \text{(E6)}$$

accurately to the first order in the strain tensor. Then, the equation

$$\langle \overline{\chi}_{n,\mathbf{k}} | \hat{U}_{\text{ext}} | \overline{\chi}_{n',\mathbf{k}'} \rangle = (\chi_{n,\mathbf{k}} | \hat{U}_{\text{ext}} | \chi_{n',\mathbf{k}'}) \quad \text{(E7)}$$

holds exactly, because $\hat{U}_{\text{ext}}$ is a function of position only, and

$$\langle \overline{\chi}_{n,\boldsymbol{k}} | \hat{H}_1 | \overline{\chi}_{n',\boldsymbol{k}'} \rangle \approx (\chi_{n,\boldsymbol{k}} | \hat{H}_1 | \chi_{n',\boldsymbol{k}'}), \tag{E8}$$

because $\hat{H}_1$ is of order 1 in the components of the strain tensor.

As a result, one has that

$$\langle \overline{\chi}_{n,\boldsymbol{k}} | \hat{H} | \overline{\chi}_{n',\boldsymbol{k}'} \rangle \approx (\chi_{n,\boldsymbol{k}} | \hat{H}_0 | \chi_{n',\boldsymbol{k}'}) + (\chi_{n,\boldsymbol{k}} | \hat{U}_{\text{ext}} | \chi_{n',\boldsymbol{k}'})$$
$$+ \frac{\hbar^2}{4m} (\chi_{n,\boldsymbol{k}} | (\nabla^2 \text{tr}\varepsilon) | \chi_{n',\boldsymbol{k}'})$$
$$+ (\chi_{n,\boldsymbol{k}} | \hat{H}_1 | \chi_{n',\boldsymbol{k}'}). \tag{E9}$$

The detailed evaluation of the matrix elements appearing in the right-hand side of Eq. (E9) is presented in Secs. I and II of the SM [44].

APPENDIX F: ENVELOPE-FUNCTION HAMILTONIAN

Since the transformed Hamiltonian, obtained after the manifold decoupling, is block-diagonal, its eigenstates are combinations of the basis states belonging to a single block. For the low-energy block, an eigenstate is written as

$$|\phi\rangle \equiv \sum_{n \leqslant N} \int_{1\text{BZ}} d\boldsymbol{k} \, \mathcal{C}_n(\boldsymbol{k}) | \overline{\chi}_{n,\boldsymbol{k}} \rangle, \tag{F1}$$

where the coefficients $\mathcal{C}_n(\boldsymbol{k})$ satisfy

$$\sum_{n' \leqslant N} \int_{1\text{BZ}} d\boldsymbol{k}' \, \mathcal{H}_{n,n'}^{(N)}(\boldsymbol{k}, \boldsymbol{k}') \mathcal{C}_{n'}(\boldsymbol{k}') = E \mathcal{C}_n(\boldsymbol{k}). \tag{F2}$$

The slowly-varying *envelope functions* are defined as

$$F_n(\boldsymbol{r}) \equiv \int_{1\text{BZ}} d\boldsymbol{k} \, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \mathcal{C}_n(\boldsymbol{k}), \tag{F3}$$

and they allow to write the spinorial wave function $\langle \boldsymbol{r} | \phi \rangle$ in the form given by Eq. (17) of the main text. Equation (F2) is rewritten in terms of the envelope functions as

$$\sum_{n' \leqslant N} \int_{1\text{BZ}} d\boldsymbol{k} \, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \int_{1\text{BZ}} d\boldsymbol{k}' \, \mathcal{H}_{n,n'}^{(N)}(\boldsymbol{k}, \boldsymbol{k}') \mathcal{C}_{n'}(\boldsymbol{k}') = E F_n(\boldsymbol{r}). \tag{F4}$$

To simplify the left-hand side of Eq. (F4), we distinguish three types of matrix elements of the Hamiltonian:

(1) Those having the form $\delta(\boldsymbol{k} - \boldsymbol{k}')\Lambda_{n,n'}(\boldsymbol{k})$: these give

$$\sum_{n' \leqslant N} \Lambda_{n,n'}(-i\nabla) F_{n'}(\boldsymbol{r}), \tag{F5}$$

analogously to standard envelope-function theories;

(2) Those having the form $\widetilde{\mathcal{U}}_{n,n'}(\boldsymbol{k} - \boldsymbol{k}')$: these include the external potential and formally analogous terms, for which the standard treatment is applicable; under the assumption that the envelope functions are slowly varying, they contribute terms

$$\sum_{n' \leqslant N} \mathcal{U}_{n,n'}(\boldsymbol{r}) F_{n'}(\boldsymbol{r}); \tag{F6}$$

(3) Those having the form $\widetilde{\varepsilon}_\nu^\mu(\boldsymbol{k} - \boldsymbol{k}')[X_\mu^\nu(\boldsymbol{k}, \boldsymbol{k}')]_{n,n'}$. This is a new category of terms, which do not map onto those related to homogeneous strain, because the Fourier transform of the strain tensor is not a Dirac delta.

The contributions to the left-hand side of Eq. (F4) which include the formally new terms are written as

$$\sum_{n' \leqslant N} \int_{1\text{BZ}} d\boldsymbol{k} \, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \int_{1\text{BZ}} d\boldsymbol{k}' \, \widetilde{\varepsilon}_\nu^\mu(\boldsymbol{k} - \boldsymbol{k}')[X_\mu^\nu(\boldsymbol{k}, \boldsymbol{k}')]_{n,n'} \mathcal{C}_{n'}(\boldsymbol{k}')$$
$$= \sum_{n' \leqslant N} \int d\boldsymbol{r}' \varepsilon_\nu^\mu(\boldsymbol{r}') \frac{1}{(2\pi)^3} \int_{1\text{BZ}} d\boldsymbol{k} \, e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')} \int_{1\text{BZ}} d\boldsymbol{k}' e^{i\boldsymbol{k}'\cdot\boldsymbol{r}'} \left\{ (\mathcal{D}_\mu^\nu)_{n,n'} + k^\alpha (\mathcal{L}_{\alpha;\mu}^\nu)_{n,n'} + k'^\alpha (\mathcal{L}_{\alpha;\mu}^{*\nu})_{n',n} + k^\alpha k'^\beta (\mathcal{Q}_{\alpha\beta;\mu}^\nu)_{n,n'} \right\} \mathcal{C}_{n'}(\boldsymbol{k}'). \tag{F7}$$

The first contribution to the right-hand side of Eq. (F7) is

$$\sum_{n' \leqslant N} (\mathcal{D}_\mu^\nu)_{n,n'} \int d\boldsymbol{r}' \varepsilon_\nu^\mu(\boldsymbol{r}') \frac{1}{(2\pi)^3} \int_{1\text{BZ}} d\boldsymbol{k} \, e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')} \int_{1\text{BZ}} d\boldsymbol{k}' e^{i\boldsymbol{k}'\cdot\boldsymbol{r}'} \mathcal{C}_{n'}(\boldsymbol{k}') \approx \varepsilon_\nu^\mu(\boldsymbol{r}) \sum_{n' \leqslant N} (\mathcal{D}_\mu^\nu)_{n,n'} F_{n'}(\boldsymbol{r}), \tag{F8}$$

the second contribution is

$$\sum_{n' \leqslant N} (\mathcal{L}_{\alpha;\mu}^\nu)_{n,n'} \int d\boldsymbol{r}' \varepsilon_\nu^\mu(\boldsymbol{r}') \frac{1}{(2\pi)^3} \int_{1\text{BZ}} d\boldsymbol{k} \, e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')} k^\alpha \int_{1\text{BZ}} d\boldsymbol{k}' e^{i\boldsymbol{k}'\cdot\boldsymbol{r}'} \mathcal{C}_{n'}(\boldsymbol{k}')$$
$$\approx -i \sum_{n' \leqslant N} (\mathcal{L}_{\alpha;\mu}^\nu)_{n,n'} \int d\boldsymbol{r}' \varepsilon_\nu^\mu(\boldsymbol{r}') F_{n'}(\boldsymbol{r}') \frac{\partial \delta(\boldsymbol{r} - \boldsymbol{r}')}{\partial r^\alpha} = -i \sum_{n' \leqslant N} (\mathcal{L}_{\alpha;\mu}^\nu)_{n,n'} \frac{\partial [\varepsilon_\nu^\mu(\boldsymbol{r}) F_{n'}(\boldsymbol{r})]}{\partial r^\alpha}, \tag{F9}$$

the third contribution is

$$\sum_{n' \leqslant N} (\mathcal{L}_{\alpha;\mu}^{*\nu})_{n',n} \int d\boldsymbol{r}' \varepsilon_\nu^\mu(\boldsymbol{r}') \frac{1}{(2\pi)^3} \int_{1\text{BZ}} d\boldsymbol{k} \, e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')} \int_{1\text{BZ}} d\boldsymbol{k}' e^{i\boldsymbol{k}'\cdot\boldsymbol{r}'} k'^\alpha \mathcal{C}_{n'}(\boldsymbol{k}') \approx -i \sum_{n' \leqslant N} (\mathcal{L}_{\alpha;\mu}^{*\nu})_{n',n} \varepsilon_\nu^\mu(\boldsymbol{r}) \frac{\partial F_{n'}(\boldsymbol{r})}{\partial r^\alpha}, \tag{F10}$$

and the fourth contribution is

$$\sum_{n' \leqslant N} \left(\mathcal{Q}_{\alpha\beta;\mu}^{\nu}\right)_{n,n'} \int d\mathbf{r}' \varepsilon_{\nu}^{\mu}(\mathbf{r}') \frac{1}{(2\pi)^3} \int_{\text{1BZ}} d\mathbf{k}\, e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} k^{\alpha} \int_{\text{1BZ}} d\mathbf{k}'\, e^{i\mathbf{k}'\cdot\mathbf{r}'} k'^{\beta} \mathcal{C}_{n'}(\mathbf{k}')$$

$$\approx -\sum_{n' \leqslant N} \left(\mathcal{Q}_{\alpha\beta;\mu}^{\nu}\right)_{n,n'} \int d\mathbf{r}' \varepsilon_{\nu}^{\mu}(\mathbf{r}') \frac{\partial F_{n'}(\mathbf{r}')}{\partial r'^{\beta}} \frac{\partial \delta(\mathbf{r}-\mathbf{r}')}{\partial r^{\alpha}} = -\sum_{n' \leqslant N} \left(\mathcal{Q}_{\alpha\beta;\mu}^{\nu}\right)_{n,n'} \frac{\partial}{\partial r^{\alpha}}\left[\varepsilon_{\nu}^{\mu}(\mathbf{r}) \frac{\partial F_{n'}(\mathbf{r})}{\partial r^{\beta}}\right]. \tag{F11}$$

In deriving the expressions above, we have used

$$\frac{1}{(2\pi)^3} \int_{\text{1BZ}} d\mathbf{k}\, e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \approx \delta(\mathbf{r}-\mathbf{r}'), \tag{F12}$$

which is an approximate relation, only valid when this quantity is multiplied by a slowly-varying spatial function, such as envelope functions and components of the strain tensor.

Collecting all terms, one obtains

$$\sum_{n' \leqslant N} \left[\left(\hat{\mathcal{H}}_{\text{EF}}^{(0)}\right)_{n,n'} + \left(\hat{\mathcal{H}}_{\text{EF}}^{(1)}\right)_{n,n'}\right] F_{n'}(\mathbf{r}) = E F_n(\mathbf{r}), \tag{F13}$$

where

$$\left(\hat{\mathcal{H}}_{\text{EF}}^{(0)}\right)_{n,n'} \equiv \left[E_n(\mathbf{0}) + \frac{\hbar^2 \hat{\mathbf{k}}^2}{2m} + U_{\text{ext}}(\mathbf{r})\right]\delta_{n,n'} + \frac{\hbar \pi_{n,n'}^{\alpha}}{m}\hat{k}_{\alpha} + \frac{\hbar^2 \Pi_{n,n'}^{\alpha\beta}}{m^2}\hat{k}_{\alpha}\hat{k}_{\beta} \tag{F14}$$

is formally the same as the standard $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian [31], but with the $\hat{k}_{\alpha}$ operators defined in curvilinear coordinates, and $(\hat{\mathcal{H}}_{\text{EF}}^{(1)})_{n,n'}$ is the strain-dependent term defined in Eq. (18).

APPENDIX G: NUMERICAL ESTIMATES

1. Strain model in a MOSFET

Here we provide quantitative estimates for the new Hamiltonian terms that depend on the inhomogeneous strain tensor and its derivatives. For simplicity, we refer to the approximate Hamiltonian of Eq. (19) and we apply it to describe a model silicon nanostructure. We consider a system similar to the one described in Ref. [49], consisting of a MOSFET with two SiGe stressors placed, respectively, above the source and the drain gates. An analytical model that approximately describes the stress tensor in such a system can be derived by solving the elasticity problem of a localized force on a semi-infinite plate [48,49].

From the stress tensor, one should then obtain the strain tensor $\varepsilon_{\beta}^{\alpha}(\mathbf{r})$ by applying the compliance relations [4], and the displacement $u^{\alpha}(\mathbf{r})$ by integrating the strain tensor according to Eq. (A3), taking, e.g., $\mathbf{r}_0 = \mathbf{0}$ and $\mathbf{u}(\mathbf{0}) = \mathbf{0}$ (which is equivalent to fixing the origin of the laboratory reference frame). However, the expression of the stress tensor given in Ref. [49] is notoriously an approximate one, which does not satisfy the compatibility equations exactly [48]. A consequence of this

fact is that there exists no function $u^{\alpha}(\mathbf{r})$ such that the strain tensor derived from the given stress (let us call it $\varepsilon_{\text{approx.}}$) satisfies Eq. (2).

An equivalent alternative formulation of this statement is that the line integral in Eq. (A3), when evaluated on $\varepsilon_{\text{approx.}}$, depends on the path connecting $\mathbf{r}_0$ to $\mathbf{r}$. This can be verified by comparing the results obtained with the following paths [with $\mathbf{r} \equiv (x, y, z)$]:

$$\text{Path 1}: \quad (0,0,0) \to (x,0,0) \to (x,y,0) \to (x,y,z),$$
$$\text{Path 2}: \quad (0,0,0) \to (0,0,z) \to (0,y,z) \to (x,y,z). \tag{G1}$$

We notice, however, that the displacements obtained by using these two paths are nearly identical almost everywhere. The discrepancies occur on small domains close to the stressors, where the simplifying assumptions beyond the analytical solution for the stress tensor apparently imply some nonphysical behavior if Path 1 is used.

Therefore, we assume that the displacement obtained with Path 2 is the physically correct one, and we adopt its expression for the function $\mathbf{u}(\mathbf{r})$. Inverting our perspective, we now take this as our starting point, and we obtain the strain tensor $\varepsilon_{\beta}^{\alpha}(\mathbf{r})$ analytically by applying Eq. (2). The strain tensor obtained in this way, by construction, is uniquely defined and obviously satisfies the compatibility conditions.

The displacement field that we adopt is then given by

$$u^x(\mathbf{r}) = \frac{\sigma_0 t}{2\pi} \sum_{i=1}^{4} s_i \left\{\left(-s_{11} + s_{12} + \frac{s_{44}}{2}\right)\frac{z^2}{x_i^2 + z^2}\left(s_{11} - s_{12} + \frac{s_{44}}{2}\right)\frac{z^2}{(x-x_i)^2 + z^2} + \left(s_{11} + s_{12} + \frac{s_{44}}{2}\right)\ln\left(\frac{(x-x_i)^2 + z^2}{x_i^2 + z^2}\right)\right\},$$

$$u^y(\mathbf{r}) = 0,$$

$$u^z(\mathbf{r}) = \frac{\sigma_0 t}{\pi} \sum_{i=1}^{4} s_i \left\{\left(s_{11} - s_{12} - \frac{s_{44}}{2}\right)\frac{x_i z}{x_i^2 + z^2} - \frac{s_{44}}{2}\frac{(x-x_i)z}{(x-x_i)^2 + z^2} - (s_{11} + s_{12})\arctan\left(\frac{z}{x_i}\right)\right.$$

$$\left. + \frac{s_{44}}{2}\arctan\left(\frac{x_i}{z}\right) - \frac{s_{44}}{2}\arctan\left(\frac{x_i - x}{z}\right)\right\}. \tag{G2}$$

Here: s_{11}, s_{12} and s_{44} are the compliance constants of Si [4]; $\{x_i; i = 1, 2, 3, 4\}$ is the set of sidewall positions defining the extension of the stressors; generalizing Ref. [49], we take

$$x_1 = -w - \frac{d}{2}, \quad x_2 = -\frac{d}{2},$$

$$x_3 = \frac{d}{2}, \quad x_4 = w + \frac{d}{2}, \tag{G3}$$

where w is the stressors' width (along the x direction) and d is their separation (d is larger than the MOSFET channel length); s_i are signs that account for the orientations of the forces applied by the stressors (taking $s_1 = s_3 = +1$ and $s_2 = s_4 = -1$, one reproduces the configuration of Ref. [49]); t is the thickness of the stressors in the z direction, and σ_0 is the biaxial stress parameter in the SiGe stressors.

2. Main new terms due to the inhomogeneity of strain

We focus on the approximate envelope-function Hamiltonian given in Eq. (19) and discuss the three main new terms resulting from our formulation, all of which are diagonal in the band index:

$$\hat{T}_1 \equiv \frac{\hbar^2}{4m} \big[\nabla^2 \varepsilon_\mu^\mu(\boldsymbol{r}) \big], \quad \hat{T}_2 \equiv -\frac{\hbar^2}{m} \varepsilon_{\alpha\beta}^{\text{sym}}(\boldsymbol{r}) \hat{k}_\alpha \hat{k}_\beta,$$

$$\hat{T}_3 \equiv \frac{\hbar^2}{m} \big[\partial_\alpha \varepsilon_{\alpha\beta}^{\text{sym}}(\boldsymbol{r}) \big] \mathrm{i}\, \hat{k}_\beta. \tag{G4}$$

We start from the first term, $\hat{T}_1$. The trace of the strain tensor is proportional to the trace of the stress tensor, and it can be shown [48] that the Laplacian of the latter vanishes in an *isotropic* system. Therefore, in such systems $\hat{T}_1$ vanishes identically. However, crystals are not isotropic, and assuming that they are is an approximation that might fail on small spatial scales such as those pertaining to nanostructures (the anisotropy of crystals is reflected in the values of the compliance parameters, which connect stress and strain tensors). Overall, $\hat{T}_1$ is a function of coordinates only, which appears on the diagonal of the $\boldsymbol{k} \cdot \boldsymbol{p}$ matrix. Therefore, it represents a correction to the slow-varying confinement potential.

The term $\hat{T}_2$ is a quadratic function of the momentum operator acting on the envelope functions; it can be viewed as a spatially dependent correction to the effective-mass terms, where the spatial dependence is due to the inhomogeneity of the strain tensor. This is the only term among those in Eq. (G4) that survives if the strain is homogeneous; however, it has been ignored so far.

The term $\hat{T}_3$ has a completely new form, being linear in the momentum operator. It is not analogous to a spin-orbit term, since it does not couple different bands.

3. Numerical values

We here present numerical estimates related to the quantities $\hat{T}_1$, $\hat{T}_2$, and $\hat{T}_3$. For illustrative purposes, we use the experimental values given in Ref. [49] to define the stressors ($\sigma_0 = -1.9$ GPa, $w = 180$ nm, $t = 50$ nm), but we assume that the distance d between them can be changed. In particular, we apply this analysis to a hypothetical device based on the 22-nm-FDSOI technology [50], and we take $d = 30$ nm. The resulting displacement field is displayed in Fig. 1.
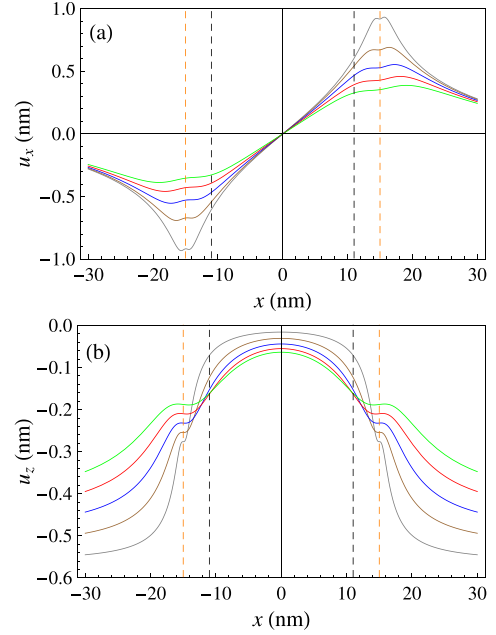

FIG. 1. Components (a) u^x and (b) u^z of the displacement from Eq. (G2), as functions of x and for selected values of $z = 1$ nm (gray), 2 nm (brown), 3 nm (blue), 4 nm (red), 5 nm (green). Parameters ($\sigma_0 = -1.9$ GPa, $t = 50$ nm, $w = 180$ nm) are taken from Ref. [49], except for $d = 30$ nm. Vertical dashed lines correspond to the extremities of the channel at $x = \pm 11$ nm (black) and to the internal borders of the stressors at $x = \pm 15$ nm (orange).

The following discussion is not meant to be exhaustive of all possible cases; it rather illustrates the main physical features of the new terms. Their actual values and the extent of their impact on the physical properties of the confined hole states are strongly device dependent, and a comprehensive analysis of the parameter space is beyond the scope of the present work.

In Fig. 2 we plot $T_1(\boldsymbol{r})$ as a function of position. As discussed in the main text, this quantity should be thought of as an effective correction to the confinement potential. As can be seen from Fig. 2(a), for values of z close to the top of the channel, this correction develops attractive (repulsive) pockets for electrons (holes), close to the ends of the channel along the x direction. In certain geometries, this might be an unwanted source of perturbation for the confinement, especially in the electron case, where the lateral attractive pockets compete with the confinement at the center of the channel. In the particular device discussed here, however, such an effect seems to be minor, due to the small values achieved by T_1 over the channel and its near-uniformity close to the center of the channel.

The operator $\hat{T}_2$ is determined by the coefficients of the quadratic forms $\hat{k}_\alpha \hat{k}_\beta$, which, apart from a multiplicative constant, are the symmetric components of the strain tensor. They are shown in Fig. 3. For the considered device, the diagonal terms (xx) and (zz) are even functions of x, and they reach values of the order of several meV nm^2 close to the center of the channel, while the (zx) term is an odd function of x, and its values in the relevant region are smaller. For comparison, the coefficients of $\hat{k}_x^2$ and $\hat{k}_z^2$ in the $\boldsymbol{k} \cdot \boldsymbol{p}$ Hamiltonian for heavy
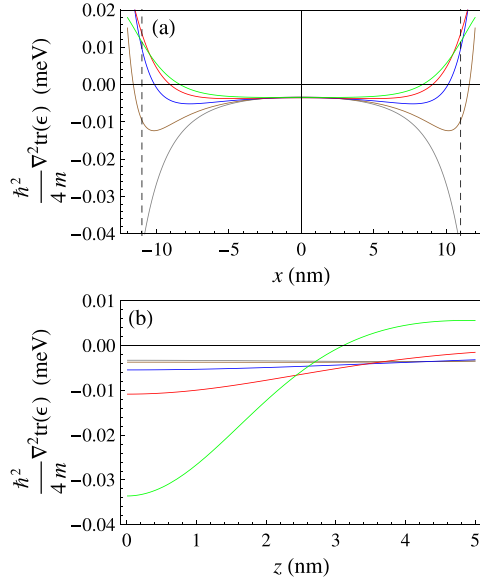
FIG. 2. Quantity $T_1(\mathbf{r})$ (a) as a function of x for selected values of z (same colors and conventions as in Fig. 1), and (b) as a function of z for selected values of $x = 0$ nm (gray), 2.5 nm (brown), 5 nm (blue), 7.5 nm (red), and 10 nm (green).
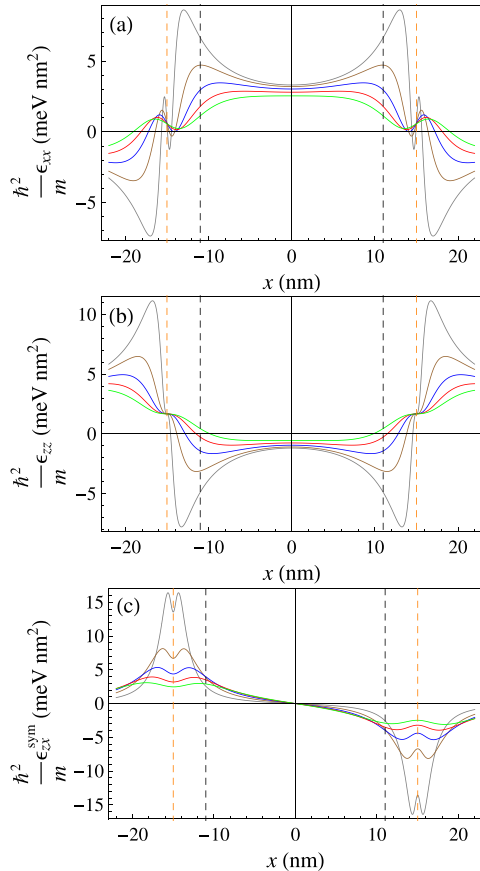

FIG. 3. Prefactors of the quadratic forms of the momentum components appearing in $\hat{T}_2$ as functions of x, for selected values of z (same colors and conventions as in Fig. 1). (a) coefficient of $\hat{k}_x^2$, (b) coefficient of $\hat{k}_z^2$, (c) coefficient of $\hat{k}_x\hat{k}_z$.
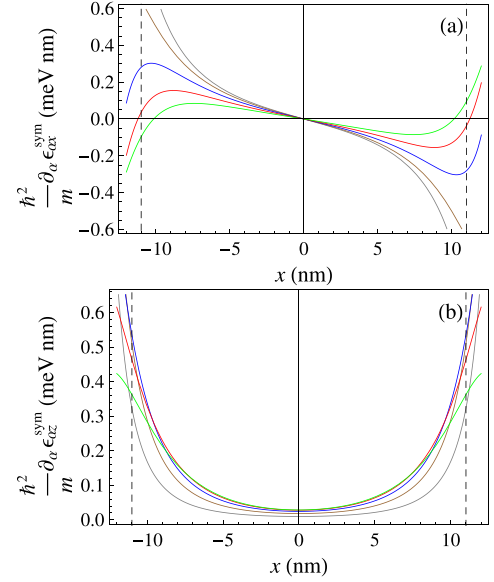

FIG. 4. Prefactors of the linear forms of the momentum components appearing in $\hat{T}_3$ as functions of x, for selected values of z (same colors and conventions as in Fig. 1). (a) coefficient of $\hat{k}_x$, (b) coefficient of $\hat{k}_z$.

and light holes in silicon are

$$\frac{\hbar^2}{2m_{x,\mathrm{H}}} = 176.174 \text{ meV nm}^2,$$

$$\frac{\hbar^2}{2m_{x,\mathrm{L}}} = 150.342 \text{ meV nm}^2,$$

$$\frac{\hbar^2}{2m_{z,\mathrm{H}}} = 137.426 \text{ meV nm}^2,$$

$$\frac{\hbar^2}{2m_{z,\mathrm{L}}} = 189.089 \text{ meV nm}^2, \tag{G5}$$

where $m_{\alpha,h}$ is the effective mass along direction α, and $h = H/L$ denotes heavy/light holes. Therefore, the $\hat{T}_2$ term produces a sizable, spatially dependent correction to the effective masses, which for the considered device can be of the order of 1% to 2%, depending on the hole species and on the direction.

Finally, the term $\hat{T}_3$ is determined by the spatially dependent coefficients multiplying the components of the momentum operator. For the case at hand, these are displayed in Fig. 4. The considered geometry produces coefficients of $\hat{k}_x$ and $\hat{k}_z$ that are, respectively, an odd and an even function of x. Therefore, in this particular case, the $\hat{k}_x$ term likely has a minor impact on the states of the confined holes; the $\hat{k}_z$ term, instead, is associated with characteristic energies of the order of $\langle \frac{\hbar^2}{m}\partial_\alpha\varepsilon_{\alpha z}^{\mathrm{sym}}\rangle\hat{k}_z \approx \overline{\frac{\hbar^2}{m}\partial_\alpha\varepsilon_{\alpha z}^{\mathrm{sym}}}\frac{1}{L_z}$, where L_z is a typical confinement length along the z direction, and $\overline{\frac{\hbar^2}{m}\partial_\alpha\varepsilon_{\alpha z}^{\mathrm{sym}}}$ is an average value of the coefficient depicted in Fig. 4(b) close to the center of the channel. Taking for the latter a representative value of 0.1 meV nm, and $L_z = 10$ nm or 5 nm (as in currently available or downscaled MOSFETs, respectively [22]), one obtains energy scales of the order of $0.01 - 0.02$ meV, corresponding to effective frequencies of $2.4 - 4.8$ GHz.

[1] J. J. Pla, A. Bienfait, G. Pica, J. Mansir, F. A. Mohiyaddin, Z. Zeng, Y. M. Niquet, A. Morello, T. Schenkel, J. J. L. Morton, and P. Bertet, Strain-induced spin-resonance shifts in silicon devices, Phys. Rev. Appl. **9**, 044014 (2018).

[2] C. C. Lo, M. Urdampilleta, P. Ross, M. F. Gonzalez-Zalba, J. Mansir, S. A. Lyon, M. L. W. Thewalt, and J. J. L. Morton, Hybrid optical–electrical detection of donor electron spins with bound excitons in silicon, Nat. Mater. **14**, 490 (2015).

[3] T. Thorbeck and N. M. Zimmerman, Formation of strain-induced quantum dots in gated semiconductor nanostructures, AIP Adv. **5**, 087107 (2015).

[4] V. Sverdlov, *Strain-Induced Effects in Advanced MOSFETs* (Springer-Verlag, Wien, 2011).

[5] Y. Sun, S. E. Thompson, and T. Nishida, Physics of strain effects in semiconductors and metal-oxide-semiconductor field-effect transistors, J. Appl. Phys. **101**, 104503 (2007).

[6] Y.-M. Niquet, C. Delerue, and C. Krzeminski, Effects of strain on the carrier mobility in silicon nanowires, Nano Lett. **12**, 3545 (2012).

[7] 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).

[8] J. Yoneda, K. Takeda, T. Otsuka, T. Nakajima, M. R. Delbecq, G. Allison, T. Honda, T. Kodera, S. Oda, Y. Hoshi, N. Usami, K. M. Itoh, and S. Tarucha, A quantum-dot spin qubit with coherence limited by charge noise and fidelity higher than 99.9%, Nat. Nanotechnol. **13**, 102 (2018).

[9] G. Scappucci, C. Kloeffel, F. A. Zwanenburg, D. Loss, M. Myronov, J.-J. Zhang, S. De Franceschi, G. Katsaros, and M. Veldhorst, The germanium quantum information route, Nat. Rev. Mater. **6**, 926 (2021).

[10] A. Chatterjee, P. Stevenson, S. De Franceschi, A. Morello, N. P. de Leon, and F. Kuemmeth, Semiconductor qubits in practice, Nat. Rev. Phys. **3**, 157 (2021).

[11] P. Stano and D. Loss, Review of performance metrics of spin qubits in gated semiconducting nanostructures, Nat. Rev. Phys. **4**, 672 (2022).

[12] A. M. J. Zwerver, T. Krähenmann, T. F. Watson, L. Lampert, H. C. George, R. Pillarisetty, S. A. Bojarski, P. Amin, S. V. Amitonov, J. M. Boter, R. Caudillo, D. Correas-Serrano, J. P. Dehollain, G. Droulers, E. M. Henry, R. Kotlyar, M. Lodari, F. Lüthi, D. J. Michalak, B. K. Mueller *et al.*, Qubits made by advanced semiconductor manufacturing, Nat. Electron. **5**, 184 (2022).

[13] A. Saraiva, W. H. Lim, C. H. Yang, C. C. Escott, A. Laucht, and A. S. Dzurak, Materials for silicon quantum dots and their impact on electron spin qubits, Adv. Funct. Mater. **32**, 2105488 (2022).

[14] G. Burkard, T. D. Ladd, A. Pan, J. M. Nichol, and J. R. Petta, Semiconductor spin qubits, Rev. Mod. Phys. **95**, 025003 (2023).

[15] B. Venitucci, L. Bourdet, D. Pouzada, and Y.-M. Niquet, Electrical manipulation of semiconductor spin qubits within the g-matrix formalism, Phys. Rev. B **98**, 155319 (2018).

[16] B. Venitucci and Y.-M. Niquet, Simple model for electrical hole spin manipulation in semiconductor quantum dots: Impact of dot material and orientation, Phys. Rev. B **99**, 115317 (2019).

[17] B. Hetényi, C. Kloeffel, and D. Loss, Exchange interaction of hole-spin qubits in double quantum dots in highly anisotropic semiconductors, Phys. Rev. Res. **2**, 033036 (2020).

[18] S. Bosco, B. Hetényi, and D. Loss, Hole spin qubits in Si FinFETs with fully tunable spin-orbit coupling and sweet spots for charge noise, PRX Quantum **2**, 010348 (2021).

[19] S. Bosco, M. Benito, C. Adelsberger, and D. Loss, Squeezed hole spin qubits in Ge quantum dots with ultrafast gates at low power, Phys. Rev. B **104**, 115425 (2021).

[20] S. Bosco and D. Loss, Fully tunable hyperfine interactions of hole spin qubits in Si and Ge quantum dots, Phys. Rev. Lett. **127**, 190501 (2021).

[21] A. Secchi, L. Bellentani, A. Bertoni, and F. Troiani, Interacting holes in Si and Ge double quantum dots: From a multiband approach to an effective-spin picture, Phys. Rev. B **104**, 035302 (2021).

[22] L. Bellentani, M. Bina, S. Bonen, A. Secchi, A. Bertoni, S. P. Voinigescu, A. Padovani, L. Larcher, and F. Troiani, Toward hole-spin qubits in Si*p*-MOSFETs within a planar CMOS foundry technology, Phys. Rev. Appl. **16**, 054034 (2021).

[23] A. Secchi, L. Bellentani, A. Bertoni, and F. Troiani, Inter- and intraband Coulomb interactions between holes in silicon nanostructures, Phys. Rev. B **104**, 205409 (2021).

[24] A. Secchi and F. Troiani, Theory of multidimensional quantum capacitance and its application to spin and charge discrimination in quantum dot arrays, Phys. Rev. B **107**, 155411 (2023).

[25] A. Secchi and F. Troiani, Multi-dimensional quantum capacitance of the two-site Hubbard model: The role of tunable interdot tunneling, Entropy **25**, 82 (2023).

[26] G. Forghieri, A. Secchi, A. Bertoni, P. Bordone, and F. Troiani, Quantum estimation and remote charge sensing with a hole-spin qubit in silicon, Phys. Rev. Res. **5**, 043159 (2023).

[27] D. Fernández-Fernández, Y. Ban, and G. Platero, Quantum control of hole spin qubits in double quantum dots, Phys. Rev. Appl. **18**, 054090 (2022).

[28] J. C. Abadillo-Uriel, E. A. Rodríguez-Mena, B. Martinez, and Y.-M. Niquet, Hole-spin driving by strain-induced spin-orbit interactions, Phys. Rev. Lett. **131**, 097002 (2023).

[29] S. D. Liles, F. Martins, D. S. Miserev, A. A. Kiselev, I. D. Thorvaldson, M. J. Rendell, I. K. Jin, F. E. Hudson, M. Veldhorst, K. M. Itoh, O. P. Sushkov, T. D. Ladd, A. S. Dzurak, and A. R. Hamilton, Electrical control of the g tensor of the first hole in a silicon MOS quantum dot, Phys. Rev. B **104**, 235303 (2021).

[30] J. M. Luttinger and W. Kohn, Motion of electrons and holes in perturbed periodic fields, Phys. Rev. **97**, 869 (1955).

[31] L. C. L. Y. Voon and M. Willatzen, *The k · p method* (Springer-Verlag, Berlin, 2009).

[32] R. Winkler, *Spin-Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems* (Springer-Verlag, Berlin, 2003).

[33] G. Bir and G. Pikus, *Symmetry and Strain-induced Effects in Semiconductors* (Wiley, New York, 1974).

[34] S. Richard, F. Aniel, G. Fishman, and N. Cavassilas, Energy-band structure in strained silicon: A 20-band k · p and Bir–Pikus Hamiltonian model, J. Appl. Phys. **94**, 1795 (2003).

[35] C. Y.-P. Chao and S. L. Chuang, Spin-orbit-coupling effects on the valence-band structure of strained semiconductor quantum wells, Phys. Rev. B **46**, 4110 (1992).

[36] T. B. Bahder, Eight-band k·p model of strained zinc-blende crystals, Phys. Rev. B **41**, 11992 (1990).

[37] T. B. Bahder, Analytic dispersion relations near the Γ point in strained zinc-blende crystals, Phys. Rev. B **45**, 1629 (1992).

[38] G. D. Whitfield, Theory of electron-phonon interactions, Phys. Rev. **121**, 720 (1961).

[39] Y. Zhang, Motion of electrons in semiconductors under inhomogeneous strain with application to laterally confined quantum wells, Phys. Rev. B **49**, 14352 (1994).

[40] W. Li, X. Qian, and J. Li, Envelope function method for electrons in slowly-varying inhomogeneously deformed crystals, J. Phys.: Condens. Matter **26**, 455801 (2014).

[41] B. S. DeWitt, Point transformations in quantum mechanics, Phys. Rev. **85**, 653 (1952).

[42] C. Gneiting, T. Fischer, and K. Hornberger, Quantum phase-space representation for curved configuration spaces, Phys. Rev. A **88**, 062117 (2013).

[43] A. I. Breev and A. V. Shapovalov, The Dirac equation in an external electromagnetic field: Symmetry algebra and exact integration, J. Phys.: Conf. Ser. **670**, 012015 (2016).

[44] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.110.045420 for the technical details related to (Sec. I) the derivation of the matrix elements of the Hamiltonian in curvilinear coordinates on the basis of the modified Luttinger-Kohn states, and to (Sec. II) the procedure for decoupling the manifold of interest from the remote bands.

[45] The reason for this is the fact that, for nonzero ε, both the ε-independent part and the ε-dependent part of H are not Hermitian operators. Their sum, instead, is Hermitian.

[46] P.-O. Löwdin, A note on the quantum-mechanical perturbation theory, J. Chem. Phys. **19**, 1396 (1951).

[47] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors* (Springer-Verlag, Berlin, 1996).

[48] S. Timoshenko and J. N. Goodier, *Theory of Elasticity* (McGraw-Hill, New York, 1951).

[49] E. Bonera, M. Bollani, D. Chrastina, F. Pezzoli, A. Picco, O. G. Schmidt, and D. Terziotti, Substrate strain manipulation by nanostructure perimeter forces, J. Appl. Phys. **113**, 164308 (2013).

[50] R. Carter, J. Mazurier, L. Pirro, J.-U. Sachse, P. Baars, J. Faul, C. Grass, G. Grasshoff, P. Javorka, T. Kammler, A. Preusse, S. Nielsen, T. Heller, J. Schmidt, H. Niebojewski, P.-Y. Chou, E. Smith, E. Erben, C. Metze, C. Bao *et al.*, 22-nm FDSOI technology for emerging mobile, internet-of-things, and RF applications, in *IEEE International Electron Devices Meeting (IEDM)* (IEEE, New York, 2016), pp. 2.2.1–2.2.4.