Modulated Rashba interaction in a quantum wire: Spin and charge dynamics

Mariana Malard
Faculdade UnB Planaltina, Universidade de Brasilia, 73300-000 Planaltina-DF, Brazil

Inna Grusha and G. I. Japaridze
Andronikashvili Institute of Physics, Tamarashvili 6, 0177 Tbilisi, Georgia and College of Engineering, Ilia State University, Cholokasvili Ave. 3-5, 0162 Tbilisi, Georgia

Henrik Johannesson
Department of Physics, University of Gothenburg, SE-412 96 Gothenburg, Sweden
(Received 19 May 2011; revised manuscript received 29 June 2011; published 12 August 2011)

It was recently shown that a spatially modulated Rashba spin-orbit coupling in a quantum wire drives a transition from a metallic to an insulating state when the wave number of the modulation becomes commensurate with the Fermi wavelength of the electrons in the wire [Japaridze *et al.*, Phys. Rev. B **80**, 041308(R) (2009)]. On the basis of experimental data from a gated InAs heterostructure it was suggested that the effect may be put to practical use in a future spin transistor design. In the present article we revisit the problem and present a detailed analysis of the underlying physics. First, we explore how the buildup of charge density wave correlations in the quantum wire due to the periodic gate configuration that produces the Rashba modulation influences the transition to the insulating state. The interplay between the modulations of the charge density and that of the spin-orbit coupling turns out to be quite subtle: Depending on the relative phase between the two modulations, the joint action of the Rashba interaction and charge density wave correlations may either enhance or reduce the Rashba current blockade effect. Second, we inquire about the role of the Dresselhaus spin-orbit coupling that is generically present in a quantum wire embedded in semiconductor heterostructure. While the Dresselhaus coupling is found to work against the current blockade of the insulating state, the effect is small in most materials. Using an effective field theory approach, we also carry out an analysis of effects from electron-electron interactions and show how the single-particle gap in the insulating state can be extracted from the more easily accessible collective charge and spin excitation thresholds. The smallness of the single-particle gap together with the antiphase relation between the Rashba and chemical potential modulations pose serious difficulties for realizing a Rashba-controlled current switch in an InAs-based device. Some alternative designs are discussed.

DOI: 10.1103/PhysRevB.84.075466 PACS number(s): 71.30.+h, 71.70.Ej, 85.35.Be

I. INTRODUCTION

The ability to control and manipulate electron spins in semiconductors via an external electric field forms the basis of the emerging spintronics technology.[1] In what has become a paradigm for the next-generation spintronics device—the Datta-Das spin field effect transistor[2]—spin-polarized electrons are injected from a ferromagnetic emitter into a quantum wire patterned in a semiconductor heterostructure. The Rashba spin-orbit interaction[3] intrinsic to a quantum well patterned in a semiconductor heterostructure causes spin flips of the injected electrons with a rate tunable by an electrical gate, and by contacting a ferromagnetic collector to the other end of the wire, electrons are either accepted or rejected depending on their spin directions. However, present techniques for injecting spin-polarized electrons from a ferromagnetic metal into a semiconductor are quite inefficient. This, among other difficulties, has obstructed the actual fabrication of a Datta-Das transistor. The best efficiency rates to date, using a Schottky contact for spin injection, are still far below what is required for a working device.[4] While other designs for spin transistors have been proposed, these suffer from similar technical difficulties as the original Datta-Das proposal. Alternative blueprints for spin transistors that do not rely on spin-polarized electron injection are thus very much wanted.

In a recent work it was shown that a smoothly modulated Rashba spin-orbit coupling in a quantum wire drives a transition from a metallic to an insulating state when the wave number of the modulation becomes commensurate with the Fermi wavelength of the electrons in the wire.[5] It was suggested that this effect may be put to practical use in a device where a configuration of equally spaced nanosized gates are placed on top of a biased quantum wire. When charged, the gate configuration produces a periodic modulation of the Rashba interaction, thus blocking the current when the electron density is tuned to commensurability by an additional backgate. By decharging the gate, the current is free to flow again. This would realize an "on-off" current switch, controllable by the backgate. The advantage of this proposal is precisely that it dispenses with the need to inject spin-polarized electrons into the current-carrying channel of the device.

The proposal in Ref. 5 was inspired by earlier work by Wang[6] and Gong and Yang,[7] showing that a current in a quantum wire where segments with a uniform Rashba coupling alternates with segments with no coupling gets blocked when the number of segments becomes sufficiently large.[8] However, the Peierls-type mechanism of the spin-based current switch identified in Ref. 5 differs from that in Refs. 6 and 7, where the current blockade is simply caused by electron scattering at the artificially sharp boundaries between the wire segments

(similar to the scattering off the boundary between the wire and the ferromagnetic collector in the Datta-Das transistor). Importantly, by instead modeling the Rashba interaction as *smoothly* modulated—thus faithfully taking into account the fact that the top gates that produce the effective Rashba field are of finite extent—yields the extra bonus of allowing for a well-controlled analysis of effects from electron-electron interactions.[5] It was found that in the experimentally relevant parameter range, the electron-electron interactions *enhance* the current blockade effect, thus assisting the use of a gate-controlled modulated Rashba interaction as a current switch.

In the present article we revisit the problem to obtain a more detailed picture of the underlying physics. First, we shall explore how the buildup of charge-density-wave (CDW) correlations in the quantum wire due to the presence of the periodic gate configuration influences the current blockade caused by the modulated Rashba interaction. While one would maybe expect the concurrent modulation of the charge density to always assist the current blockade, the interplay between the two effects turns out to be more subtle: When the two modulations are in phase they do work in tandem, but, when antiphased, a *crossover* regime is observed where the two modulations compete with each other and, as a result, the Rashba current blockade effect is reduced by the joint action of the Rashba interaction and CDW correlations. While at first surprising, we shall be able to provide a simple explanation of this crossover effect. Second, we shall inquire about the role of the *Dresselhaus spin-orbit interaction* present in any semiconductor heterostructure that supports a quantum wire (as most heterostructures used in experiments are made out of compounds with broken lattice inversion symmetry, thus implying the presence of a Dresselhaus interaction).[9] The Dresselhaus interaction is found to oppose the current-blockade effect, but as long as the Rashba interaction dominates that of Dresselhaus, the effect is small and does not detract from the viability of using a modulated Rashba interactions as the *modus operandi* for a novel type of spin transistor.

The rest of the paper is organized as follows: In Sec. II we lay the groundwork and construct the minimal model that captures the effect of a modulated Rashba spin-orbit interaction in a quantum wire. In Sec. III we show that a stripped-down version of the model—describing noninteracting electrons—can be mapped onto two independent sine-Gordon models using bosonization. We perform a renormalization-group (RG) analysis of the relevant low-energy limit of the theory and extract the condition for an opening of a mass gap in the spin and charge sectors. In Sec. IV we extend the analysis to the realistic case of interacting electrons. This analysis is patterned on that in the previous section, albeit with some added technical subtleties. By carrying it out with Sec. III as a template, we believe that our results will gain in transparency and ease of interpretation. Again, we extract the condition under which an insulating gap opens, allowing us to assess the effectiveness of using a gate-controlled modulated Rashba interaction as a current switch. In Sec. V we then carry out a case study, using our results to predict the size of the gap for a quantum wire patterned in a gated InAs-based heterostructure for which good experimental data are available. While we find that for this particular structure the gap will be too small to be usable for a current switch, our analysis points the way to more effective designs. Finally, in Sec. VI we summarize our results. Throughout the paper we try to provide enough detail to make it essentially self-contained to a reader with some acquaintance with bosonization and perturbative RG methods.

II. THE MODEL

In the following we consider a setup with a 1D quantum wire formed in a gated 2D quantum well supported by a semiconductor heterostructure. We assume that the electrons in the wire are ballistic, restricting us to wire lengths on the micronscale for most materials. Moreover, by modeling the wire as an *ideal* 1D wire that carries only one conduction channel, we will neglect effects from the transverse confining potential. This simplification greatly facilitates our analysis, but, as we shall argue, has little or no effects on our results. In the standard tight-binding formalism,[10] the kinetic energy and the chemical potential as well as the interaction energy between the electrons in the wire are described by the lattice Hamiltonians H_0 and $H_{\text{e-e}}$, respectively, with

$$H_0 = -t\sum_{n,\zeta}(c^\dagger_{n,\zeta}c_{n+1,\zeta} + \text{H.c.}) - \mu\sum_{n,\zeta} c^\dagger_{n,\zeta}c_{n,\zeta}, \tag{1}$$

$$H_{\text{e-e}} = \frac{1}{2}\sum_{n,n',\zeta,\zeta'} V(n-n')c^\dagger_{n,\zeta}c^\dagger_{n',\zeta'}c_{n',\zeta'}c_{n,\zeta}. \tag{2}$$

Here $c^\dagger_{n,\zeta}$ ($c_{n,\zeta}$) is the creation (annihilation) operator for an electron with spin $\zeta = \uparrow, \downarrow$ on site n, t is the electron hopping amplitude, and μ a uniform chemical potential controllable by an electrical backgate. The Coulomb interaction $V(n-n')$ between electrons at sites n and n' is screened by the metallic gates in the device, with the screening length set by the distance to the nearest gate.[11]

The electrons in a 2D quantum well are subject to two types of spin-orbit interactions, the *Dresselhaus*[9] and *Rashba*[3] interactions, both originating from the inversion asymmetry of the potential $V(\boldsymbol{r}) = V_{\text{cr}}(\boldsymbol{r}) + V_{\text{ext}}(\boldsymbol{r})$, where $V_{\text{cr}}(\boldsymbol{r})$ is the periodic crystal potential, and $V_{\text{ext}}(\boldsymbol{r})$ is the aperiodic part containing effects from other sources (quantum-well confinement, impurities, electrical gates, etc.). The potential gradient $\nabla V(\boldsymbol{r})$ produces a Pauli spin-orbit interaction that can be written as

$$H_{\text{SO}} = \lambda_{\text{cr}}[\boldsymbol{k}\times\nabla V_{\text{ext}}(\boldsymbol{r})]\cdot\boldsymbol{\sigma} - \boldsymbol{b}(\boldsymbol{k})\cdot\boldsymbol{\sigma}, \tag{3}$$

where, in the first term, the contribution from V_{cr} has been absorbed in the effective constant λ_{cr}, while in the second term, $\boldsymbol{b}(\boldsymbol{k})$ is an intrinsic spin-orbit field produced by V_{cr} only. Here $\boldsymbol{k}$ is the wave number of an electron, with $\boldsymbol{\sigma}$ the vector of Pauli matrices representing its spin. In semiconductors where the crystal potential lacks inversion symmetry, i.e., $V_{\text{cr}}(-\boldsymbol{r}) \neq V_{\text{cr}}(\boldsymbol{r})$ (including zinc-blende lattice structures, to which the often used GaAs and InAs quantum wells belong), the internal spin-orbit field $\boldsymbol{b}(\boldsymbol{k})$ in Eq. (3) fails to average to zero in a unit cell, resulting in a spin splitting encoded by the effective Dresselhaus interaction.[9] For a heterostructure grown along [001], with the electrons confined to the quantum well in the

xy plane, the leading term in the Dresselhaus interaction takes the simple form

$$H_\beta = \beta(k_x\sigma_x - k_y\sigma_y), \tag{4}$$

with β a material- and structure-dependent parameter.[12]

The spin degeneracy in a quantum well can be lifted also because of the structure inversion asymmetry of the confining potential contained in $V_{\rm ext}(\boldsymbol{r})$. More precisely, the spatial asymmetry of the edge of the conduction band along the growth direction of the quantum well (i.e., in the z direction perpendicular to the symmetry plane of the well) mimics an electric field in that same direction, and one obtains from Eq. (3) the Rashba interaction[3]

$$H_\alpha = \alpha(k_x\sigma_y - k_y\sigma_x). \tag{5}$$

The Rashba coupling α has a complex dependence on several distinct features of the quantum well, including the ion distribution in the nearby doping layers,[13] the relative asymmetry of the electron density at the two quantum well interfaces,[14] and, importantly, the applied gate electric field.[15] The latter feature allows for a gate control of the Rashba coupling α, with a variation of more than a factor of 2 from its base value reported for InAs quantum wells.[16] One must realize that α in Eq. (5) is a spatial average of a microscopic randomly fluctuating Rashba coupling. In a zinc-blende lattice structure the fluctuations can be quite large, with a root-mean-square deviation roughly of the same size as the average α.[13] As discussed in Ref. 17, for quantum wells with an anomalously large Rashba coupling—as in the HgTe quantum wells which support quantum spin Hall states—this large disordering effect may cause an Anderson transition to an insulating state when the electron-electron interaction is weakly screened. In other zinc-blende lattice structures, like GaAs or InAs favored in most spintronics applications, the disordering effect is weaker, with a Rashba-induced localization length that is expected to be much longer than the mean-free path due to impurity scattering. Having already assumed that the wire has a length that is smaller than the mean free path, we can therefore ignore the random fluctuations in the Rashba coupling in what follows.

Projecting the Dresselhaus and Rashba interactions in Eqs. (4) and (5) along the direction $\hat{x}$ of the quantum wire and using the same tight-binding lattice formalism as in Eqs. (1) and (2), one obtains

$$H_{DR} = -i\sum_{n,\zeta,\zeta'} c^\dagger_{n,\zeta}\left[\gamma_D\,\sigma^x_{\zeta\zeta'} + \gamma_R\,\sigma^y_{\zeta\zeta'}\right]c_{n+1,\zeta'} + \text{H.c.}, \tag{6}$$

where $\gamma_D = \beta a^{-1}$, $\gamma_R = \alpha a^{-1}$, with a the lattice spacing. The relative sign and magnitude of γ_D and γ_R depends on the material as well as on the particular design of the heterostructure, with $|\gamma_D| \approx |\gamma_R| \approx 5\times 10^{-2}$ meV in a typical GaAs-based quantum well, while in a HgTe quantum well the Rashba coupling is orders of magnitude larger than that of Dresselhaus, with $|\gamma_R| \gtrapprox 10^2 \times |\gamma_D| \approx 10$ meV.[18] Let us mention in passing that the effect of a *uniform* spin-orbit interaction on the electron dynamics in a quantum wire has been theoretically investigated for both noninteracting[19] and interacting electrons[20–22] and is by now well understood.

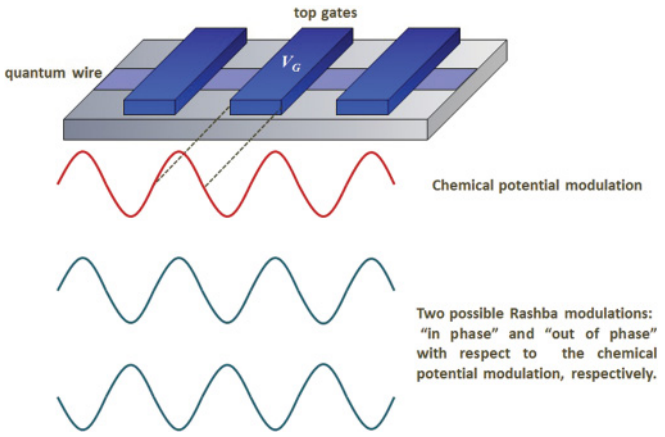

FIG. 1. (Color online) Schematic figure of the device studied in the paper. The modulated chemical potential and Rashba interaction are shown for both "in phase" and "out of phase" modulations. The amplitudes of the chemical potential and Rashba modulations are proportional to the gate voltage V_G.

We shall assume that the wire is patterned in a heterostructure on top of which are placed a periodic sequence of equally sized nanoscale gates, positively charged and pairwise separated by the same distance as their extensions along the direction of the wire. The gates may be realized by a series of ultrasmall capacitively coupled metallic electrodes deposited on the top of the heterostructure, as illustrated in Fig. 1. By charging the gates, one produces a periodic modulation of the Rashba coupling, together with a concurrent modulation of the local chemical potential in the wire, with amplitudes depending on the associated voltage drop across the well (proportional to the gate voltage V_G). The modulation will be smoothly varying along the wire, reflecting the finite extent of the gates in addition to effects from distortions and stray electric fields. To a good approximation, the modulation can be represented by a simple harmonic and thus we may write

$$H_{\rm mod} = i\sum_{n,\zeta,\zeta'}\left[\gamma_{R\rm mod}\cos(qna)c^\dagger_{n,\zeta}\sigma^y_{\zeta\zeta'}c_{n+1,\zeta'} - \text{H.c.}\right] - \sum_{n,\zeta}\mu_{\rm mod}\cos(qna)c^\dagger_{n,\zeta}c_{n,\zeta}. \tag{7}$$

Here $|\gamma_{R\rm mod}|$ ($\mu_{\rm mod}$) is the amplitude of the Rashba field (local chemical potential) modulation, both of wave number q. Note that the Rashba coupling and the chemical potential modulations are "in phase" when $\gamma_{R\rm mod} > 0$ (and hence has the same sign as $\mu_{\rm mod}$, which is always positive), while for $\gamma_{R\rm mod} < 0$ the two modulations are "out of phase" by π. The two possible phase relations between the Rashba and the chemical potential modulations are illustrated in Fig. 1. Assuming that the gate electrodes that produce the chemical potential modulation are positively charged, the segment of the wire below a gate has an enhanced magnitude of the local chemical potential (see Fig. 1) but with a negative sign. Note that the negative sign has been taken out of the sum in Eq. (7) [as well as in Eq. (1), which contains the *uniform* chemical potential].[23]

The second-quantized expression for the lattice Hamiltonian $H_{\rm mod}$ in Eq. (7) provides a microscopic

definition of the Rashba and chemical potential modulations and is manifestly Hermitian by virtue of the subtraction of the H.c. term. This procedure replaces the symmetrization $\alpha(x)(-i\partial/\partial x) \to \{\alpha(x), -i\partial/\partial x\}/2$ for a spatially varying Rashba interaction $\alpha(x)(-i\partial/\partial x)\sigma_y$ often employed in the literature as a means to ensure Hermiticity in a first-quantized continuum formalism.[24]

It is important to stress that the relation between a spatially modulated gate bias and a Rashba interaction may be more complex than that which transpires from our simple model. Already when tuning a gate voltage that is *uniform* along a quantum wire patterned in a heterostructure, the Rashba interaction has been shown to sometimes respond in a surprising way, even reversing its sign without a reversal of the gate bias.[25,26] In fact, the details of how the various effects mentioned above (including the external gate voltage) influence the magnitude and the sign of the Rashba parameter in a gated heterostructure have proven notoriously difficult to sort out and remain a somewhat contagious issue.[27] We shall not attempt to add to this discussion but instead focus on the physics implied by the idealized situation described by $H_{\rm mod}$ in Eq. (7).

Having defined our model by the Hamiltonian

$$H = H_0 + H_{\text{e-e}} + H_{DR} + H_{\rm mod}, \tag{8}$$

with $H_0, H_{\text{e-e}}, H_{DR}$, and $H_{\rm mod}$ given by Eqs. (1), (2), (6), and (7), respectively, it is now convenient to pass to a basis which diagonalizes the uniform spin-orbit interaction H_{DR}. For this purpose we first perform a rotation of the coordinate system by an angle $2\theta = \arctan(\gamma_D/\gamma_R)$ around the $\hat{z}$ axis to select the direction of the combined uniform Rashba and Dresselhaus field, $\sim \gamma_D\,\hat{x} + \gamma_R\,\hat{y}$, as our new $\hat{y}'$ axis:

$$e^{-i\theta\sigma_z}[\gamma_D\sigma_x + \gamma_R\sigma_y]e^{i\theta\sigma_z} = \gamma_{\rm eff}\,\sigma_{y'}, \tag{9}$$

where $\gamma_{\rm eff} = \sqrt{\gamma_R^2 + \gamma_D^2}$. We then introduce a spinor basis which diagonalizes $\sigma_{y'}$,

$$\begin{pmatrix} d_{n,+} \\ d_{n,-} \end{pmatrix} \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\theta}c_{n,\uparrow} - ie^{i\theta}c_{n,\downarrow} \\ -ie^{-i\theta}c_{n,\uparrow} + e^{i\theta}c_{n,\downarrow} \end{pmatrix}, \tag{10}$$

where the spinor components $\tau = \pm$ of the operator $d_{n,\tau}$ label the new quantized spin projections along $\hat{y}'$, with y' defining the orientation of the axis around which the expectation value of the spin will now be precessing. With this, we write the transformed Hamiltonian as

$$H' = H_0' + H_{\text{e-e}}' + H_{DR}' + H_{\rm mod}', \tag{11}$$

with

$$H_0' = -t\sum_{n,\tau}(d_{n,\tau}^\dagger d_{n+1,\tau} + \text{H.c.}) - \mu\sum_{n,\tau} d_{n,\tau}^\dagger d_{n,\tau} \tag{12}$$

$$H_{\text{e-e}}' = \frac{1}{2}\sum_{n,n',\tau,\tau'} V(n-n')d_{n,\tau}^\dagger d_{n',\tau'}^\dagger d_{n',\tau'} d_{n,\tau} \tag{13}$$

$$H_{DR}' = -i\,\gamma_{\rm eff}\sum_{n,\tau}\tau\, d_{n,\tau}^\dagger d_{n+1,\tau} + \text{H.c.}, \tag{14}$$

$$\begin{aligned} H_{\rm mod}' = &-i\sum_{n,\tau}\gamma_R(n)\tau\cos(2\theta)d_{n,\tau}^\dagger d_{n+1,\tau} \\ &+ i\sum_{n,\tau}\gamma_R(n)\sin(2\theta)d_{n,\tau}^\dagger d_{n+1,-\tau} \\ &- \frac{1}{2}\sum_{n,\tau}\mu(n)d_{n,\tau}^\dagger d_{n,\tau} + \text{H.c.}, \end{aligned} \tag{15}$$

with $\gamma_R(n) \equiv \gamma_{R\rm mod}\cos(qna)$ and $\mu(n) \equiv \mu_{\rm mod}\cos(qna)$.

Let us add a comment that our procedure leading up to Eqs. (11)–(15) is not to be confounded with the gauge transformation approach to two-dimensional spin-orbit interactions recently suggested by Tokatly and Sherman[28] (see also Ref. 29). Whereas our transformation is simply a global spinor rotation, the gauge transformation in Ref. 28 is by construction a local rotation, yielding a manifest spin-charge duality. It would be interesting to explore whether the approach by Tokatly and Sherman[28] can be adapted to the case also of a modulated spin-orbit interaction, but for now we leave this for the future.

While the theory defined by Eqs. (11)–(15) may look forbiddingly complex, we find that a bosonization approach yields a well-controlled analytical solution in the physically relevant limit of low energies. In the next section we study the case with no electron-electron interaction, i.e., with $V(n-n') = 0$ for all n,n' in Eq. (13). This simplification allows us to focus on the key elements of our solution approach, paving the ground for the more elaborate analysis of the full theory in Sec. IV.

III. NONINTERACTING ELECTRONS

A. Effective Hamiltonian

Neglecting the electron-electron interaction $H_{\text{e-e}}$, and taking $\gamma_R(n) = \mu(n) = 0$ (assuming that there is no modulated electric field present), the remaining piece of the Hamiltonian in Eq. (11), $H_0' + H_{DR}'$, is easily diagonalized by a Fourier transform,

$$H_0' + H_{DR}' = \sum_{k,\tau=\pm} E_\tau^{(0)}(k)d_{k,\tau}^\dagger d_{k,\tau}, \tag{16}$$

where

$$E_\tau^{(0)}(k) = -2\tilde{t}\cos[(k+\tau q_0)a] - \mu, \tag{17}$$

with $\tilde{t} = \sqrt{t^2+\gamma_{\rm eff}^2}$ and $q_0 a = \arctan(\gamma_{\rm eff}/t)$ and where a is the lattice constant. At band-filling $\nu = N_e/2N_0$, with N_e [N_0] the number of electrons [lattice sites], the system is characterized by the four Fermi points $k_{F,R}^\tau = k_F + \tau q_0, k_{F,L}^\tau = -k_F + \tau q_0$ ($\tau = \pm$), where $k_F = \pi\nu/a$, reflecting the band splitting caused by the uniform spin-orbit interaction H_{DR}' in Eq. (14).

To analyze the effect of adding the modulated term $H_{\rm mod}'$ in Eq. (15) to $H_0' + H_{DR}'$, it is convenient to linearize the spectrum around these Fermi points and then pass to a continuum limit with $na \to x$. By decomposing the lattice operators $d_{n,\tau}$ into right- and left-moving fields $R_\tau(x)$ and $L_\tau(x)$,

$$d_{n,\tau} \to \sqrt{a}\left[e^{ik_{F,R}^\tau x}R_\tau(x) + e^{ik_{F,L}^\tau x}L_\tau(x)\right],$$

we find that in this limit $H_0' + H_{DR}' + H_{\rm mod}' = \sum_\tau \int dx\, (\mathcal{H}_\tau + \text{H.c.})$, with

$$\begin{aligned}\mathcal{H}_\tau =& -i(v_F/2)(:R_\tau^\dagger(x)\partial_x R_\tau(x): - :L_\tau^\dagger(x)\partial_x L_\tau(x):)\\ &-(\lambda_R e^{-i\pi\nu} + \mu_{\rm mod})\cos(qx)e^{-2ik_F x} R_\tau^\dagger(x)L_\tau(x)\\ &+ i\lambda_D \sin(\pi\nu)\cos(qx)e^{-iq_0\tau(2x+a)}\\ &\times [R_\tau^\dagger(x)R_{-\tau}(x) - L_\tau^\dagger(x)L_{-\tau}(x)],\end{aligned} \tag{18}$$

where $v_F = 2a\tilde{t}\sin(\pi\nu)$, $\lambda_R = 2\tilde{\gamma}_R \sin(q_0 a)$, and $\lambda_D = \tilde{\gamma}_D$, with $\tilde{\gamma}_j = \gamma_{R\rm mod}\gamma_j(\gamma_R^2 + \gamma_D^2)^{-1/2}$, $j = R, D$. The normal ordering $: \ldots :$ is carried out with respect to the filled Dirac sea. Note that in deriving Eq. (18) we have omitted all rapidly oscillating terms that vanish on integration.

The Hamiltonian in Eq. (18) supports four distinct limiting cases, depending on the difference between the modulation wave number q and the parameters k_F, q_0:

(i) $|q \pm 2k_F| \simeq \mathcal{O}(1/a)$, $|q \pm 2q_0| \simeq \mathcal{O}(1/a)$;
(ii) $|q \pm 2k_F| \simeq \mathcal{O}(1/a)$, $|q - 2q_0| \ll \mathcal{O}(1/a)$;
(iii) $|q - 2k_F| \ll \mathcal{O}(1/a)$, $|q \pm 2q_0| \simeq \mathcal{O}(1/a)$;
(iv) $|q - 2k_F| \ll \mathcal{O}(1/a)$, $|q - 2q_0| \ll \mathcal{O}(1/a)$.

In the first case (i), all terms in Eq. (18) proportional to λ_R or $\mu_{\rm mod}$ or λ_D rapidly oscillate and thus average to zero when integrated. It follows that in this limit the model describes a two-component free Fermi gas, i.e., a metallic phase with gapless excitations. In contrast, in case (ii), when $|q - 2q_0| \ll \mathcal{O}(1/a)$, the corresponding terms proportional to λ_D become slowly varying and contribute to the dynamics. These terms emulate the presence of a transverse effective field, causing electrons to flip their spins along the direction of the combined uniform Rashba and Dresselhaus fields. Turning to case (iii), with $|q - 2k_F| \ll \mathcal{O}(1/a)$ but with $|q \pm 2q_0| \simeq \mathcal{O}(1/a)$, one now finds that the terms proportional to λ_D are washed away on integration, while the terms proportional to λ_R or to $\mu_{\rm mod}$ survive. This implies that backscattering and CDW correlations come into play, dramatically changing the physics: A band gap opens at all four Fermi points, causing a *transition to a nonmagnetic insulating state.* Finally, in case (iv), *all* terms in Eq. (16) contribute to the integrated Hamiltonian, leading to a rather complex theory. This case, however, where k_F and q_0 both approach q, requires a fine tuning of both the electron density (on which k_F depends) *and* the uniform Rashba interaction in Eq. (5) (on which q_0 depends). This case is expected to be difficult to realize in an experiment, and in the following we shall focus on the more accessible case (iii).

B. Bosonization picture: Band insulator from modulated Rashba interaction

To see how the spectacular effect driven by the modulated Rashba interaction comes about [case (iii) in the previous section], it is useful to bosonize the theory. Using standard bosonization, we write the right- and left-moving fermionic fields as

$$R_\tau(x) = \frac{\eta_\tau}{\sqrt{2\pi a}} e^{i\sqrt{\pi}[\varphi_\tau(x) + \vartheta_\tau(x)]}, \tag{19}$$

$$L_\tau(x) = \frac{\bar{\eta}_\tau}{\sqrt{2\pi a}} e^{i\sqrt{\pi}[\varphi_\tau(x) - \vartheta_\tau(x)]}, \tag{20}$$

where $\varphi_\tau(x)$ and $\vartheta_\tau(x)$ are dual bosonic fields satisfying $\partial_t\varphi_\tau = v_F\partial_x\vartheta_\tau$ and where η_τ and $\bar{\eta}_\tau$ are Klein factors which keep track of the fermion statistics for electrons in different branches.[30]

Inserting the bosonized forms of $R_\tau(x)$ and $L_\tau(x)$ into Eq. (18) and carrying out some simple algebra, one obtains the bosonized Hamiltonian

$$\begin{aligned}&H_0' + H_{DR}' + H_{\rm mod}'\\ &= \sum_\tau \int dx \left(\frac{v_F}{2}[(\partial_x\vartheta_\tau)^2 + (\partial_x\varphi_\tau)^2] \right.\\ &\quad + \sum_{j=\pm1} \left\{ \frac{\lambda_R}{\pi a} \sin[(q + 2jk_F)x + \pi\nu + \sqrt{4\pi}\varphi_\tau] \right.\\ &\quad \left.\left. - \frac{\mu_{\rm mod}}{2\pi a} \sin[(q + 2jk_F)x + \sqrt{4\pi}\varphi_\tau] \right\} \right).\end{aligned} \tag{21}$$

It is useful to cast the Hamiltonian in Eq. (21) on the more compact form

$$\begin{aligned}&H_0' + H_{DR}' + H_{\rm mod}'\\ &= \sum_\tau \int dx \left\{ \frac{v_F}{2}[(\partial_x\vartheta_\tau)^2 + (\partial_x\varphi_\tau)^2] \right.\\ &\quad \left. + \frac{M_R}{\pi a} \sum_{j=\pm1} \cos[(q + 2jk_F)x + \phi_0 + \sqrt{4\pi}\varphi_\tau] \right\},\end{aligned} \tag{22}$$

where

$$M_R = \sqrt{\lambda_R^2 + \mu_{\rm mod}\lambda_R\cos(\pi\nu) + \mu_{\rm mod}^2/4}, \tag{23}$$

$$\phi_0 = -\arctan\left[\frac{\mu_{\rm mod} + 2\lambda_R\cos(\pi\nu)}{2\lambda_R\sin(\pi\nu)} \right]. \tag{24}$$

For the case that we are interested in, i.e., with $|q - 2k_F| \ll \mathcal{O}(1/a)$, the $j = -1$ component of the modulated term in Eq. (22) comes into play.[31] For this case we can gauge out the small term $\propto x$ from the argument of the cosine by the transformation $(q - 2k_F)x + \phi_0 + \sqrt{4\pi}\varphi_\tau \to \sqrt{4\pi}\varphi_\tau$ and rewrite the Hamiltonian as $H_0' + H_{DR}' + H_{\rm mod}' = \sum_\tau \int dx\, \mathcal{H}_{{\rm bos},\tau}(x)$, with Hamiltonian densities

$$\begin{aligned}\mathcal{H}_{{\rm bos},\tau} =& \frac{v_F}{2}[(\partial_x\varphi_\tau)^2 + (\partial_x\vartheta_\tau)^2] - \frac{\mu_{\rm eff}}{\sqrt{\pi}}\partial_x\varphi_\tau\\ &+ \frac{M_R}{\pi a}\cos(\sqrt{4\pi}\varphi_\tau),\end{aligned} \tag{25}$$

where

$$\mu_{\rm eff} = v_F(2k_F - q)/2 \tag{26}$$

serves as an effective chemical potential. By tuning the density of electrons so $\mu_{\rm eff} = 0$, the system is seen to be governed by two commuting sine-Gordon models[32] with interaction terms $\cos(\beta\varphi_+)$ and $\cos(\beta\varphi_-)$, respectively, where $\beta^2 = 4\pi$. As follows from the exact solution of the sine-Gordon model,[33] in this case the excitation spectrum is gapped and consists of *solitons* and *antisolitons* with masses $M_+ = M_- = M_R$ (together with soliton-antisoliton bound states, so-called *breathers*, with masses $\geqslant M_R$). A soliton (or antisoliton) corresponds to a configuration of the field φ_τ, for

a given component τ, that connects two neighboring minima $\sqrt{4\pi}\varphi_\tau^0 = \pi + 2\pi n$ ($n \in \mathcal{Z}$) of the functional potential $V[\varphi_\tau] = M_R \cos(\sqrt{4\pi}\varphi_\tau)$. The previous field configurations define the set of possible ground states of φ_τ with vacuum expectation values $\langle\varphi_\tau^0\rangle = \sqrt{\pi}(1/2 + n)$. For example, a field configuration where $\varphi_\tau(-\infty) = \sqrt{\pi}/2$ [$3\sqrt{\pi}/2$] and $\varphi_\tau(\infty) = 3\sqrt{\pi}/2$ [$\sqrt{\pi}/2$] supports a soliton [antisoliton] with fermion number $N_\tau = 1$ [-1], defined by

$$N_\tau = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} dx\, \partial_x \varphi_\tau(x). \tag{27}$$

The charge and spin quantum numbers of the single-particle excitation are given by

$$Q = N_+ + N_-, \quad S_z = \tfrac{1}{2}(N_+ - N_-). \tag{28}$$

The simplest single-particle excitation is obtained by considering a soliton or antisoliton in the spin $\tau = +$ component, keeping the ground state unperturbed for the spin $\tau = -$ component: $N_+ = \pm 1$, $N_- = 0$. Such an excitation has charge $Q = \pm 1$ and spin $S_z = \pm 1/2$ (with spin projections $\tau = \pm$ along the direction of the momentum-dependent combined uniform Rashba and Dresselhaus fields). Thus, *the elementary excitations of the system are free massive fermions with mass M_R, each carrying unit charge and spin 1/2.* It follows that the joint action of the modulated Rashba coupling and the chemical potential, with the electron density tuned to satisfy the commensurability condition $\mu_{\rm eff} = 0$, turns the electron gas into an effective band insulator. The corresponding band gap is equal to the doubled mass of the single-particle excitation, $\Delta = 2M_R$, since conservation of charge and spin requires the simultaneous excitation of a soliton and an antisoliton. Note from Eq. (23) and the definition of λ_R after Eq. (18) that the effect of the Dresselhaus interaction is to reduce the gap. Fortunately, as we shall show in Sec. V, this unwanted effect (from the point of view of spintronics applications) is negligible when compared to the stronger Rashba interaction.

C. Bosonization picture in the spin-charge basis

The nature of the metal-insulator transition becomes more transparent if we treat the model in a basis with charge (c) and spin (s) bosons—the standard basis in which to include effects of electron-electron interactions.[30] Thus, introducing the dual *charge fields*

$$\varphi_c = \tfrac{1}{\sqrt{2}}(\varphi_+ + \varphi_-), \quad \vartheta_c = \tfrac{1}{\sqrt{2}}(\vartheta_+ + \vartheta_-) \tag{29}$$

and *spin fields*

$$\varphi_s = \tfrac{1}{\sqrt{2}}(\varphi_+ - \varphi_-), \quad \vartheta_s = \tfrac{1}{\sqrt{2}}(\vartheta_+ - \vartheta_-), \tag{30}$$

some simple algebra on Eq. (25) yields that $H_0' + H_{DR}' + H_{\rm mod}' = \int dx\, [\mathcal{H}_{0c} + \mathcal{H}_{0s} + \mathcal{H}_{cs}]$, with

$$\mathcal{H}_{0c} = \frac{v_F}{2}[(\partial_x\varphi_c)^2 + (\partial_x\vartheta_c)^2] - \sqrt{\frac{2}{\pi}}\mu_{\rm eff}\partial_x\varphi_c, \tag{31}$$

$$\mathcal{H}_{0s} = \frac{v_F}{2}[(\partial_x\varphi_s)^2 + (\partial_x\vartheta_s)^2], \tag{32}$$

$$\mathcal{H}_{cs} = \frac{2M_R}{\pi a}\cos(\sqrt{2\pi}\varphi_c)\cos(\sqrt{2\pi}\varphi_s). \tag{33}$$

At $\mu_{\rm eff} = 0$ Eqs. (31)–(33) describe two bosonic charge and spin fields coupled by the strongly (renormalization-group) relevant operator $\mathcal{H}_{cs}$ (of scaling dimension 1). This operator drives the system to a strong-coupling regime where the charge and spin fields are pinned at their ground-state expectation values. Therefore, at $\mu_{\rm eff} = 0$, both charge and spin excitations develop a gap (let us call them M_c and M_s, respectively) and the system becomes a nonmagnetic insulator, consistent with the finding in the previous section where the system develops a gap also in the bosonic $\tau = \pm$ basis. As $\mu_{\rm eff}$ is tuned away from zero the influence of the operator $\mathcal{H}_{cs}$ in Eq. (33) gets weaker and eventually averages to zero on integration when $\mu_{\rm eff}$ exceeds the insulator band gap given by the mass M_c of charge excitations. At this point, the system then turns metallic. Therefore, the competition between the chemical potential term and the commensurability energy drives a continuous insulator-to-metal transition from a gapped phase at $\mu_{\rm eff} < \mu_{\rm eff}^c$ to a gapless phase at $\mu_{\rm eff} > \mu_{\rm eff}^c = M_c$.[34] The (de-)tuning of $\mu_{\rm eff}$ in our Eq. (26) can be achieved by changing either k_F or q, i.e., the band filling ν ($k_F = \pi\nu/a$) or the wavelength λ of the gate modulation ($q = 2\pi/\lambda$). Either alternative poses its own experimental difficulties, although we expect that the band filling is more easily controllable, using a back gate with a variable voltage. Therefore, we shall hereafter assume that the tuning mechanism is provided by an adjustable band filling. Thus (rephrasing this in a language closer to experiment) by detuning the voltage of the backgate of the device so the electron density n_s of the quantum well is shifted from the value $\pi/2\lambda^2$, one will observe a transition from a nonmagnetic insulating state into a metallic phase.[35] In this phase the electrons in the wire exhibit ordinary Fermi liquid behavior with gapless quasiparticle excitations. This kind of transition belongs to the universality class of commensurate-to-incommensurate transitions:[30] The conductivity σ close to the transition scales as $\sigma \sim (\mu - \mu_c)^{1/2}$, with the compressibility κ diverging as $\kappa \sim (\mu - \mu_c)^{-1/2}$, before dropping to zero on the insulating side.

The insulator-to-metal transition just discussed corresponds to the picture put forward by Schulz in Ref. 36 where a Hamiltonian similar to that defined by our Eqs. (31)–(33) is refermionized into a two-band model [cf. Eq. (4) in Ref. 36]. The two bands are separated by a gap Δ, with a chemical potential $\mu_0 = 0$ corresponding to a completely filled lower band. In other words, in this state the system is a band insulator with a gap Δ. For μ_0 smaller than the critical value $-\Delta/2$, holes are introduced at the top of the lower band, whereas for μ_0 larger than $\Delta/2$, electrons are added to the bottom of the upper band; in both cases the system becomes metallic. This refermionized picture thus makes it clear that it takes a *finite* critical μ_0 for the transition to occur: by tuning the chemical potential to zero the system not only develops a gap but also a rigidity that sustains the gap when the system is shifted away from commensurability.

Going back to Eqs. (31)–(33), it is instructive to see how the single-fermion excitations obtained in the previous section can be reconstructed in the present spin-charge basis. Here we follow a route developed in Ref. 37 in studies of a similar problem in the case of the ionic Hubbard model. First, note that only the relative sign between $\langle\cos(\sqrt{2\pi}\varphi_c)\rangle$ and $\langle\cos(\sqrt{2\pi}\varphi_s)\rangle$ in Eq. (33) is fixed to be negative

(so as to minimize $\langle \mathcal{H}_{cs} \rangle$). Thus, there are two possibilities for the ground-state charge and spin field expectation values:

$$\text{I:} \quad \langle \varphi_c \rangle = \sqrt{2\pi}\, m \quad \text{and} \quad \langle \varphi_s \rangle = \sqrt{\frac{\pi}{2}}(2n+1), \tag{34}$$

$$\text{II:} \quad \langle \varphi_c \rangle = \sqrt{\frac{\pi}{2}}(2m+1) \quad \text{and} \quad \langle \varphi_s \rangle = \sqrt{2\pi}\, n, \tag{35}$$

with $m,n \in \mathcal{Z}$. To obtain the single-fermion excitations one has to consider field configurations that connect two ground states that belong to *distinct* sets I Eq. (34) and II Eq. (35). As an example, a field configuration that connects $\varphi_c = 0$ with $\varphi_c = \sqrt{\pi/2}$ in the charge sector and $\varphi_s = \sqrt{\pi/2}$ with $\varphi_s = \sqrt{2\pi}$ in the spin sector corresponds to an excitation with charge and spin quantum numbers[39]

$$Q = \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} dx\, \partial_x \varphi_c(x) = 1, \tag{36}$$

$$S_z = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\, \partial_x \varphi_s(x) = 1/2, \tag{37}$$

i.e., a massive fermion (of mass M_R), which is the elementary excitation in the band insulator. To obtain a pure charge or spin excitation one must consider field configurations that connect ground states *within* the sets I and II. For example, given set I in Eq. (34), we can lock the charge at $\varphi_c = 0$ and consider a *spin soliton* connecting the ground states at $\varphi_s = \sqrt{\pi/2}$ and $\varphi_s = 3\sqrt{\pi/2}$. Such an excitation carries charge $Q = 0$ and spin

$$S_z = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\, \partial_x \varphi_s(x) = 1. \tag{38}$$

In the noninteracting case considered here it is clear that that this excitation is built from two massive fermions with opposite charge and the same spin. Similarly, a *charge soliton* can be obtained by locking the spin at one of the possible ground states and consider a charge field configuration that connects, say, $\varphi_c = 0$ and $\varphi_c = \sqrt{2\pi}$. This excitation carries charge

$$Q = \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} dx\, \partial_x \varphi_c(x) = 2 \tag{39}$$

while having zero spin, being built from two massive fermions with the same charge and opposite spin.

Following this logic, a derivation of M_c and M_s should give $M_c = M_s = 2M_R$. As we shall see, the mean-field approach used in the next section to evaluate M_c and M_s gives a slightly overestimated value. We shall return to this issue below and show how it can be resolved by a proper regularization procedure.

The opening of a gap for both charge and spin excitations[38] at commensurability, $\mu_{\text{eff}} = 0$, reflects the fact that the system has turned into a *nonmagnetic band insulator*. Using the standard bosonized expression for the charge density,[39]

$$\rho_c(x) \simeq \frac{1}{\sqrt{2\pi}} \partial_x \varphi_c + A \sin(\sqrt{2\pi}\varphi_c + 2k_F x) \cos(\sqrt{2\pi}\varphi_s) \tag{40}$$

with A a constant, one verifies from Eqs. (34) and (35) (which apply at $\mu_{\text{eff}} = 0$), that the ground state of the system corresponds to a *CDW-type band insulator*, with long-range charge-density modulation

$$\rho_c(x) \simeq \rho_c^m \sin(2k_F x), \tag{41}$$

where

$$\rho_c^m \sim \langle \cos(\sqrt{2\pi}\varphi_c) \rangle \langle \cos(\sqrt{2\pi}\varphi_s) \rangle\,. \tag{42}$$

As should be clear from the nonconservation of spin in the presence of the spin-orbit interactions, the massiveness of the spin excitations does not correspond to the formation of a spin density wave (SDW). Indeed, by writing down the bosonized expression[39] for a SDW with spin projection along the direction of the combined uniform Rashba and Dresselhaus fields, $\rho_s(x) \simeq (1/\sqrt{2\pi})\partial_x \varphi_s + B \cos(\sqrt{2\pi}\varphi_c + 2k_F x) \sin(\sqrt{2\pi}\varphi_s)$, with B a constant, one immediately verifies from Eqs. (34) and (35) that it has no amplitude for a long-range $2k_F$ modulation.

D. Bosonic mean-field theory in the spin-charge basis

To pave the ground for including electron-electron interactions into the problem, we next decouple the interaction term $\mathcal{H}_{cs}$ in Eq. (33) in a mean-field manner by introducing

$$m_c = 2M_R |\langle \cos(\sqrt{2\pi}\varphi_s) \rangle|, \tag{43}$$

$$m_s = 2M_R |\langle \cos(\sqrt{2\pi}\varphi_c) \rangle|\,. \tag{44}$$

Note that the mean-field decoupling is well controlled since, at the strong-coupling fixed point, fluctuations are strongly suppressed by the pinning of the charge and spin bosons. Using Eqs. (43) and (44), we find that the mean-field version of the bosonized Hamiltonian $H_0' + H_{DR}' + H_{\text{mod}}' = \int dx\, [\mathcal{H}_{0c} + \mathcal{H}_{0s} + \mathcal{H}_{cs}]$, defined in Eqs. (31)–(33), can be written as $H_{\text{mean}} = \int dx\, [\mathcal{H}_c + \mathcal{H}_s]$ with

$$\mathcal{H}_c = \frac{v_F}{2}[(\partial_x \varphi_c)^2 + (\partial_x \vartheta_c)^2] + \frac{m_c}{\pi a} \cos(\sqrt{2\pi}\varphi_c) - \sqrt{\frac{2}{\pi}} \mu_{\text{eff}} \partial_x \varphi_c, \tag{45}$$

$$\mathcal{H}_s = \frac{v_F}{2}[(\partial_x \varphi_s)^2 + (\partial_x \vartheta_s)^2] + \frac{m_s}{\pi a} \cos(\sqrt{2\pi}\varphi_s). \tag{46}$$

When $\mu_{\text{eff}} = 0$, the Hamiltonian defined by Eqs. (45) and (46) is again given by a sum of two decoupled sine-Gordon models [cf. Eq. (25)]. However, the dimensionalities of the $\cos(\beta\varphi)$ operators at $\beta^2 = 2\pi$ [spin-charge basis, Eqs. (45) and (46)] and $\beta^2 = 4\pi$ [$\tau = \pm$ basis, Eq. (25)] differ.

By exploiting the exact solution of the sine-Gordon model, we can easily estimate the size of the insulating gap in the spin-charge basis.[40] The excitation spectra of Eq. (45) at $\mu_{\text{eff}} = 0$ and Eq. (46) consist of solitons and antisolitons with masses M_c and M_s, respectively (in addition to the charge and spin breathers with masses bounded below by M_c and M_s, respectively). These charge and spin soliton masses are related to the "bare" masses m_c and m_s in Eqs. (45) and (46) by[41]

$$M_\kappa/\Lambda = \mathcal{C}_0\, (m_\kappa/\Lambda)^{2/3}\,, \quad \kappa = c,s. \tag{47}$$

with Λ an energy cutoff that blocks excitations into the second conduction channel of the quantum wire (for details, see Sec. V).

The ground-state expectation values of $\cos(\sqrt{2\pi}\varphi_\kappa)$ are, in turn, given by[42]

$$|\langle\cos\sqrt{2\pi}\varphi_\kappa\rangle| = \mathcal{C}_1\,(M_\kappa/\Lambda)^{1/2}\,, \quad \kappa = c,s, \tag{48}$$

with $\mathcal{C}_0 \approx 1.4$ and $\mathcal{C}_1 \approx 1.0$. (For details, see Appendix.) By combining Eqs. (47) and (48) with (43) and (44) one reads off that

$$M_c = M_s = 2\gamma M_R, \tag{49}$$

with $\gamma = \mathcal{C}_0^{3/2}\mathcal{C}_1 \simeq 1.7$.

Note that charge and spin solitons, though formally decoupled, move with the same velocity v_F [cf. Eqs. (45) and (46)], a record of their composite nature since, as demonstrated in Sec. III C, charge and spin excitations are built from single fermions of mass M_R, unit charge and spin $S = 1/2$, the latter being the elementary excitations in the $\tau = \pm$ basis of Sec. III B. Thus, the mean-field treatment in the spin-charge basis faithfully captures the character of the excitations. However, the size of the two-particle gap, M_c or M_s as given in Eq. (49), gets overestimated by a factor of 1.7 when compared to the result $\Delta = 2M_R$, obtained in Sec. III B. As shown in the next section, the factor of 1.7 can be removed by introducing a regularized form of the gap.

Recall that tuning the effective chemical potential away from zero "closes" the band gap and thus drives an insulator-to-metal transition by depinning the charge field from its ground-state expectation value. The combined Eqs. (44), (47), and (48) reveal that, in the process, the spin sector becomes gapless as well, as $M_s \sim \langle\cos(\sqrt{2\pi}\varphi_c)\rangle = 0$.

E. Functional behavior of the effective band gap

Having established that the masses of the charge and spin excitations in the noninteracting theory, M_c and M_s, respectively, are determined by the single-fermion mass M_R, let us return to Eq. (23) to analyze its dependence on the relative phase between the two modulations and their amplitudes λ_R and $\mu_{\rm mod}$. As emphasized in the previous sections, the mass M_R is a key parameter of our theory, encoding the effective band gap $\Delta = 2M_R$ in the insulating state of noninteracting electrons.

In Sec. V, when analyzing a generic gated heterostructure, we see that both λ_R and $\mu_{\rm mod}$ depend linearly on the voltage V_G of the top gates (cf. Fig. 1). We may thus write $\lambda_R = c_1 V_G$ and $\mu_{\rm mod} = c_2 V_G$, where c_1 and c_2 are constants depending on the details of the setup and of the sample.

To analyze the gap behavior it is important to distinguish the two ways in which the parameters λ_R and $\mu_{\rm mod}$ can be varied: One possibility is to consider (i) a *fixed system* (i.e., keeping c_1 and c_2 fixed) and vary the gate voltage V_G; alternatively, one may consider (ii) *different systems* but keep the gate voltage fixed, e.g., by testing different samples from an ensemble of properly gated heterostructures [all of which satisfy the commensurability condition $|q - 2k_F| \ll \mathcal{O}(1/a)$].

Let us start by investigating possibility (i). In this case, we can rewrite $\mu_{\rm mod} = (c_2/c_1)\lambda_R$ and Eq. (23) as

$$M_R = c(\nu)|\lambda_R|, \tag{50}$$

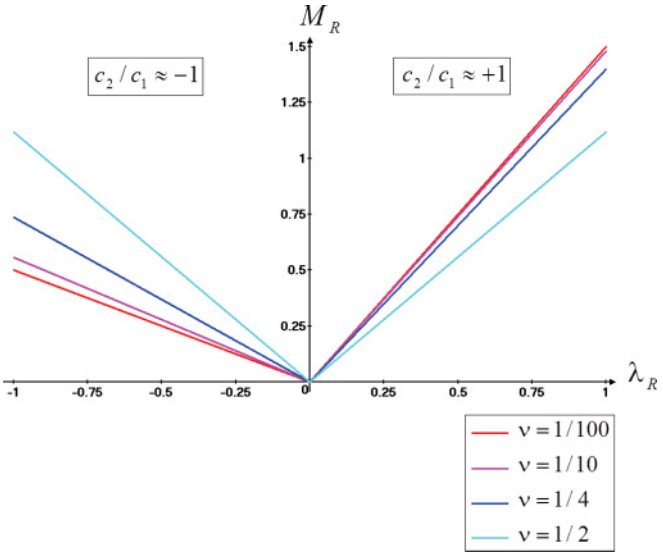

FIG. 2. (Color online) M_R as a function of λ_R (common arbitrary units) for given values of the band-filling ν. "In phase" ["Out of phase"] modulations correspond to $c_2/c_1 \simeq 1$ [$c_2/c_1 \simeq -1$].

where $c(\nu) = \sqrt{1 + (c_2/c_1)\cos(\pi\nu) + (c_2/c_1)^2/4}$ is a system specific parameter adjustable by the band filling ν (which, in turn, can be varied by a back gate with a variable voltage).

Figure 2 shows M_R as a function of λ_R for band fillings $\nu = 1/100,\ 1/10,\ 1/4,\ 1/2$. The reason for considering systems only up to half-filling is the following: Due to the commensurability condition $|q - 2k_F| \ll \mathcal{O}(1/a)$, the values of the filling ν in Fig. 2 correspond to modulation wavelengths $\lambda = 100a, 10a, 4a,$ and $2a$, respectively (as seen from the relations $q = 2\pi/\lambda$ and $k_F = \pi\nu/a$). Since the ultrasmall gates that we propose to be used for producing the modulation each has a spatial extension $\lambda/2$ along the quantum wire (cf. Fig. 1), it follows that $\nu = 1/2$ sets an upper (and, in practice, unattainable) physical limit for possible band fillings: If $\nu > 1/2$, the dimension of a gate would have to become subatomic. In fact, as can be gleaned from the experimental data cited in Sec. V, our theory would likely break down already for band fillings around $\nu \approx 1/3$ since at larger fillings higher sub-bands will come into play, causing sub-band mixing. As we shall also see in Sec. V, the 1D band filling with present-day semiconductor heterostructures is typically around $\nu \approx 1/10$, implying a gate extension of a few nanometers. This presents a challenge to the experimentalist.

The plots in Fig. 2 are shown for λ_R running from -1 to $+1$, thus accounting for the two possible phase relations between the Rashba and the chemical potential modulations. For "in phase" ["out of phase"] modulations, the plots are shown for test systems where $c_2/c_1 \simeq 1$ [$c_2/c_1 \simeq -1$]. We see that for both "in phase" and "out of phase" modulations, the gap is an increasing linear function of $|\lambda_R|$ with slope depending on the band filling. The increase of the gap is consistent with the phenomenological expectation that the insulating state gets more stable as the pinning Rashba interaction goes stronger and is in agreement also with the corresponding result in Ref. 5. There, however, the modulation of the charge density was not taken into account and, thus, the formalism did not capture the gap dependence on the band filling. The split of

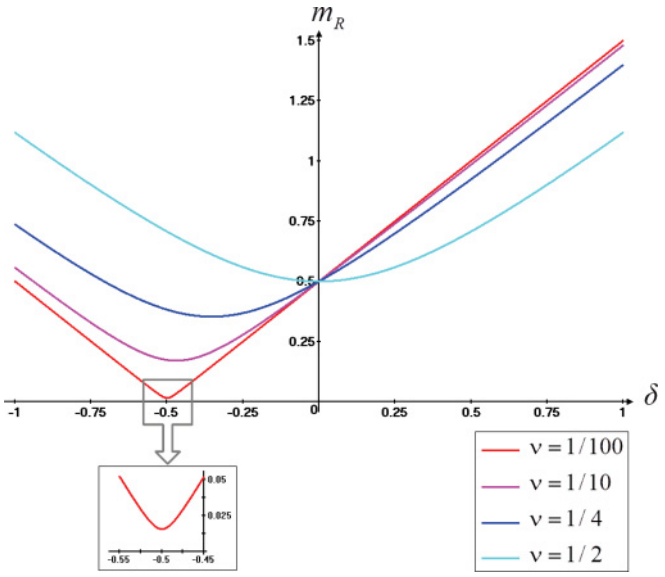

FIG. 3. (Color online) m_R as a function of the ratio $\delta \equiv \lambda_R/\mu_{\rm mod}$ for given values of the band-filling ν.

a single gap line for different values of ν, as manifest in Fig. 2, is an interesting feature of the system resulting from the combination of the modulated Rashba interaction and CDW correlations.

Another interesting aspect of the gap behavior is that, given a certain band filling ν and a value for $|\lambda_R|$, the gap for "in phase" modulations is larger than for "out of phase" modulations, implying a stronger localization effect when the the Rashba interaction and the chemical potential act in "unison." The difference $M_R(\lambda_R) - M_R(-\lambda_R)$ goes to zero as ν approaches $1/2$ (half-filled band).

Turning now to case (ii), we can define a new variable $\delta \equiv \lambda_R/\mu_{\rm mod} = c_1/c_2$ that characterizes a particular setup, material, or design and is independent of the value of the applied gate voltage. With that we can rewrite Eq. (23) as

$$m_R = \sqrt{\delta^2 + \cos(\pi\nu)\delta + 1/4}, \tag{51}$$

where $m_R = M_R/\mu_{\rm mod}$.

Figure 3 shows m_R as a function of δ for the same band fillings $\nu = 1/100, 1/10, 1/4, 1/2$ used in case (i). The plots are shown for δ running from -1 to $+1$, accounting for "in phase" and "out of phase" Rashba and chemical potential modulations.

To understand the behavior revealed by Fig. 3, let us first look at the right-half of the graph where the Rashba and chemical potential modulations are "in phase." The plots here show a monotonic increase of m_R with δ, implying that with devices where the top gate voltages V_G have been tuned to produce the same fixed value of $\mu_{\rm mod}$, the gap will be larger in the device with the larger value of λ_R. This behavior is equivalent to that obtained in case (i).

Now turn to the left half of the graph where the Rashba and chemical potential modulations are "out of phase." Again, comparing devices having the same band filling and for which the magnitude V_G of the gate voltages have been tuned to give the same $\mu_{\rm mod}$, we observe an unexpected feature. Let us follow what happens when going through different devices by moving along a curve with a fixed band filling (smaller than $1/2$): the gap first decreases with the strength of the Rashba interaction until it reaches a minimum at $\delta^\star = \cos(\pi\nu)/2$; past this value the "normal" increasing behavior is recovered. This *crossover* behavior gets more pronounced for smaller values of ν, i.e., as the system goes more diluted. For $\nu = 1/100$, for example, the crossover almost annihilates the gap at $\delta \simeq 0.5$ (but not completely as can be seen by zooming in around that point).

To understand how this phenomenon comes about, consider a *Gedanken* experiment with a single device where λ_R is allowed to vary while $\mu_{\rm mod}$ is kept fixed. In the case with antiphased modulations, the chemical potential (Rashba potential) will have a maximum (minimum) in the middle of the segment, call it "A," below one of the small positively charged gates [and vice versa for a neighboring gate-free segment, call it "B" (cf. Fig. 1)]. Consider first the case with $\lambda_R = 0$. Here all electrons will reside in the A regions since this configuration is energetically more advantageous. This causes a localization of electrons, with a gap $M_R = \mu_{\rm mod}/2$ as seen in Fig. 3 for $\delta = 0$. Now turn on λ_R. To take advantage of the spin-flip Rashba hopping, some electrons will start migrating into the B regions. This weakens the localization effect of the chemical potential modulation, with the result that the gap *decreases*. By successively increasing λ_R, more and more electrons will migrate into the B regions, and for a sufficiently large value of λ_R, equal to $\mu_{\rm mod}\cos(\pi\nu)/2$, *all* electrons will reside in the B regions. From that point on the gap will increase monotonically as λ_R (or γ_R) is further increased, as displayed in Fig. 3. For fillings close to the upper physical limit, that is $1/2$, a small λ_R is enough for a complete electronic migration between the chemical potential A regions to the spin-orbit coupling B regions while, for a more dilute system, the process is "slower," demanding a larger λ_R. The two extreme cases are $\nu = 1/2$ for which the A regions are emptied right away for any nonzero λ_R and $\nu \to 0$ for which the necessary λ_R is (still just) half the amplitude $\mu_{\rm mod}$ of the original pinning chemical potential.

In the next section we shall investigate how the electron-electron interaction influences the results thus obtained.

IV. INTERACTING ELECTRONS

A. Adding interactions: Bosonization picture in the spin-charge basis

To incorporate the electron-electron interaction $H'_{\rm e\text{-}e}$ in Eq. (13) into the bosonic theory we perform the same steps as in Sec. III B, first linearizing the spectrum around the four Fermi points and taking a continuum limit. This yields $H'_{\rm e\text{-}e} = \int dx\, \mathcal{H}_{\rm e\text{-}e}(x)$, with

$$\mathcal{H}_{\rm e\text{-}e} = g_1^{\tau\tau'} :R_\tau^\dagger L_\tau L_{\tau'}^\dagger R_{\tau'}: + g_2^{\tau\tau'} :R_\tau^\dagger R_\tau L_{\tau'}^\dagger L_{\tau'}: + \frac{g_4^{\tau\tau'}}{2}(:L_\tau^\dagger L_\tau L_{\tau'}^\dagger L_{\tau'}: + R \leftrightarrow L), \tag{52}$$

with $\tau,\tau' = \pm$ summed over and where $g_1^{\tau\tau'}$ and $g_2^{\tau\tau'}$ are the amplitudes, respectively, for back and forward ("dispersive") scattering between electrons of different chiralities and $g_4^{\tau\tau'}$ is the amplitude for forward scattering between electrons of equal chirality.[30] Whereas the $g_2^{\tau\tau'}$ and $g_4^{\tau\tau'}$ processes

correspond to scattering with small momentum transfer, the $g_1^{\tau\tau'}$ process transfers momentum $k \sim 2k_F$. For a screened Coulomb interaction with a nonzero screening length, the $g_1^{\tau\tau'}$ amplitude is therefore quite small and can usually be neglected. This is certainly so in the present case since in a semiconductor structure the Coulomb interaction is much smaller than the bandwidth. It follows that in this limit the $k \sim 2k_F$ scattering becomes marginally irrelevant and renormalizes to zero at low energies. Importantly, this conclusion is not invalidated by the presence of the spin-orbit couplings.[21] From now on we shall therefore consider the simpler theory where the backscattering has been renormalized away, i.e., with $g_1^{\tau\tau'} \simeq 0$.

For a system at commensurate band-filling $\nu = 1/2n$, with $n \geqslant 1$ an integer, the Hamiltonian density in Eq. (52) should be supplemented by an umklapp term that describes the transfer of $2n$ electrons of equal chirality to the opposite Fermi point through exchange of momentum with the lattice. As is well known, these processes drive a transition to an insulating state at a critical value of the Coulomb interaction determined by the number $2n$ of electrons participating in the process.[43] However, the screened Coulomb interaction in a gated semiconductor structure is too weak to support such a transition except at a half-filled ($n = 1$) or possibly a quarter-filled ($n = 2$) band.[11] This should be contrasted with the commensurability condition $q = 2k_F$ for driving a metal-to-insulator transition via a modulated Rashba interaction, as derived in Sec. III B. Since $q = 2\pi/\lambda$, with λ the wavelength of the Rashba modulation, and $k_F = \nu\pi/a$, this condition translates to $\lambda = 2na$ when $\nu = 1/2n$. Thus, with a Rashba modulation tuned to commensurability with $2k_F$, umklapp processes at $n = 1,2$ could come into play only for a sequence of electrical gates of near-atomic dimensions, $\lambda/2 \sim a$. For this reason we shall neglect umklapp processes when studying the novel physics coming from a Rashba modulation.

Having disposed of backscattering and umklapp processes, the remaining electron-electron interaction in Eq. (52) is now easily bosonized using Eqs. (19), (20), (29), and (30). The resulting expression for the bosonized mean-field theory representing the full H' in Eq. (11) then takes the form $H'_{\text{mean}} = \int dx\,[\mathcal{H}'_c + \mathcal{H}'_s]$, with

$$\mathcal{H}'_c = \int dx \left\{ \frac{v_c}{2}[(\partial_x\vartheta_c)^2 + (\partial_x\varphi_c)^2] - \mu_{\text{eff}}\sqrt{\frac{2K_c}{\pi}}\partial_x\varphi_c + \frac{m_c}{\pi a}\cos(\sqrt{2\pi K_c}\varphi_c)\right\}, \tag{53}$$

$$\mathcal{H}'_s = \int dx \left\{ \frac{v_s}{2}[(\partial_x\vartheta_s)^2 + (\partial_x\varphi_s)^2] + \frac{m_s}{\pi a}\cos(\sqrt{2\pi K_s}\varphi_s)\right\}, \tag{54}$$

where we have performed the field transformations $\varphi_i \to \sqrt{K_i}\varphi_i$ and $\vartheta_i \to \vartheta_i/\sqrt{K_i}$, $i = c,s$. A comparison with Eqs. (45) and (46) shows that H'_{mean} has the same structure as the mean-field theory for noninteracting electrons and is given by two decoupled sine-Gordon models when the commensurability condition $\mu_{\text{eff}} = 0$ is satisfied. The electron-electron interaction is encoded by the new parameters v_i and K_i, $i = c,s$, as well as by the reparameterization of the bare masses m_c and m_s due to the transformation $\varphi_i \to \sqrt{K_i}\varphi_i$ [cf. Eqs. (43) and (44)]. In the weak-interaction limit considered here, v_i and K_i can be given explicit representations in terms of the scattering amplitudes in Eq. (52). Introducing the conventional "g-ology" notation[30] $g_\parallel \equiv g^{\tau\tau}$ for parallel spins and $g_\perp \equiv g^{\tau-\tau}$ for opposite spins, one has that

$$v_i = v_F[(1 + y_{4i}/2)^2 - (y_i/2)^2]^{1/2}, \tag{55}$$

$$K_i = \left[\frac{1 + y_{4i}/2 + y_i/2}{1 + y_{4i}/2 - y_i/2}\right]^{1/2}, \tag{56}$$

for $i = c,s$, where

$$y_i = \frac{g_i}{\pi v_F}, \qquad y_{4i} = \frac{g_{4i}}{\pi v_F}, \tag{57}$$

$$g_i = -g_{2\parallel} \mp g_{2\perp}, \qquad g_{4i} = g_{4\parallel} \pm g_{4\perp}, \tag{58}$$

with the upper and lower signs in Eqs. (58) referring to c and s, respectively. If backscattering processes were to be included in the theory, $g_{2\parallel} \to g_{2\parallel} - g_{1\parallel} \equiv \tilde{g}_{2\parallel}$ in Eq. (58). In addition, the K_s parameter in the spin sector would become subject to a RG flow, coupled to the marginally irrelevant flow of $g_{1\perp}$, the amplitude for backscattering of electrons with opposite spins.[30] The breaking of spin-rotational invariance by the presence of spin-orbit interactions implies that the RG fixed-point value of K_s, call it K_s^*, is not slaved to unity but can take larger values. However, with the backscattering processes being weak the resulting renormalization would be small. We will return to this issue in Sec. V.

B. Charge-, spin-, and single-particle gaps

Given the bosonized mean-field theory defined by Eqs. (53) and (54) we shall now address the question of how electron-electron interactions influence the Rashba-induced single-particle gap established in Sec. III B. As anticipated in Sec. III C, this task gets complicated by the fact that already for noninteracting electrons the excitation gap in the spin-charge basis is nontrivially related to the single-particle gap, being in effect a composite two-particle gap. Moreover, as seen in Eq. (49), the mean-field theory in the spin-charge basis overestimates the actual size of this two-particle gap. The situation for interacting electrons gets further confounded by the fact that the spin and charge gaps are no longer identical but take on separate values, reflecting the collective nature of the excitations in the presence of electron-electron interactions.

Taking off from Sec. III D where we calculated the mean-field charge soliton mass M_c and spin soliton mass M_s for the case of noninteracting electrons, we perform a similar procedure, now with electron-electron interactions included, starting with the reparametrized sine-Gordon models in Eqs. (53) and (54). Note that by construction, and in exact analogy with the noninteracting case discussed in Sec. III D, M_c and M_s are the mean-field approximations of the spin and charge gaps of the fully interacting theory, Δ_c and Δ_s, respectively. Using Eqs. (43), (44), (53), and (54), we get the following relations among M_c, M_s, and M_R:

$$\eta_c^{-1}M_c = \eta_s^{-1}M_s = \Lambda(2M_R/\Lambda)^{2/(4-K_c-K_s)}, \tag{59}$$

where $\eta_c = \eta_c(K_c, K_s)$ satisfies

$$\eta_c^{16-4K_c-4K_s} \equiv C_c^{(4-K_c)(4-K_s)} \times B_c^{2K_s} C_s^{(4-K_s)K_s} B_s^{8-2K_s}, \tag{60}$$

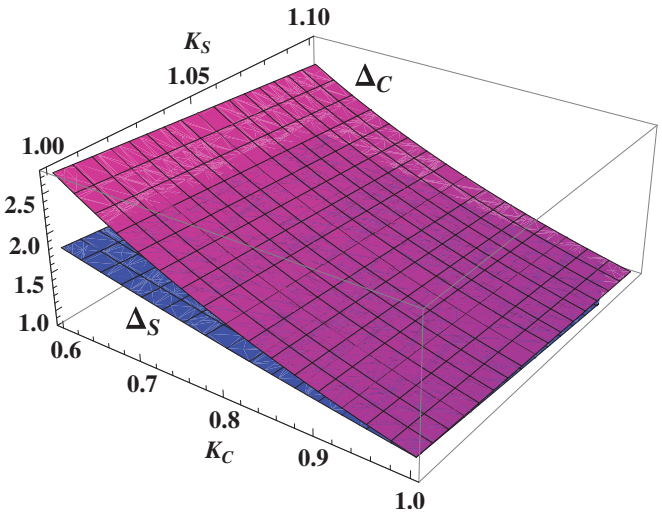

FIG. 4. (Color online) The mean-field regularized charge (Δ_c) and spin (Δ_s) gaps measured in units of the bare gap M_R, as a function of the parameters K_c and K_s in the experimentally relevant parameter range $0.6 \leqslant K_c \leqslant 1.0$ and $1.0 \leqslant K_s \leqslant 1.1$.

with η_s given by the same expression, but with $c \leftrightarrow s$, and where $B_i \equiv B(K_i)$, $C_i \equiv C(K_i), i = c,s$, are defined in Appendix.

The mean-field version of the noninteracting theory is recovered by choosing $K_c = K_s = 1$, for which

$$\begin{aligned}\eta_c(1,1) &= C_c^{9/8} B_c^{2/8} C_s^{3/8} B_s^{6/8}\,|_{K_c=K_s=1} \\ &= C(1)^{3/2} B(1) \\ &\approx 1.7,\end{aligned} \tag{61}$$

with the identical number for $\eta_s(1,1)$, the result which we arrived at already in Eq. (49) via a slightly different route. Thus, to repeat, while the mean-field theory correctly reproduces the identity $M_c = M_s$ for noninteracting electrons, the size of the corresponding two-particle gap $\Delta_c = \Delta_s$ gets overestimated by a factor of 1.7.

We can improve on the situation by dividing away this number *for all* K_c and K_s, thus in effect defining a regularized version of the mean-field spin and charge gaps,

$$\Delta_i \equiv \eta_i^{-1}(1,1) M_i, \tag{62}$$

with $M_i, i = c,s$, given in Eq. (59). By construction, this produces gives the correct noninteracting limit.

Figure 4 shows Δ_c and Δ_s for the experimentally relevant parameter range $0.6 \leqslant K_c \leqslant 1.0$ and $1.0 \leqslant K_s \leqslant 1.1$. (As an example, to be elaborated on in Sec. V, a generic quantum wire obtained by gating an InAs heterostructure is well described by taking $K_c \approx 0.7$ and $K_s \approx 1.1$.) The fundamental features of the influence of electronic interactions on the charge and spin gaps can be gleaned from Fig. 5 that shows a projection of the previous surfaces on the $K_s = 1.0$ plane.

As is manifest by these curves, the important property that $\Delta_s \leqslant \Delta_c$, expected on physical grounds,[30] is respected by the regularized mean-field gaps. The equality is valid at $K_c = K_s = 1.0$, that is, in the absence of electronic interactions, reproducing the result of Sec. III D.

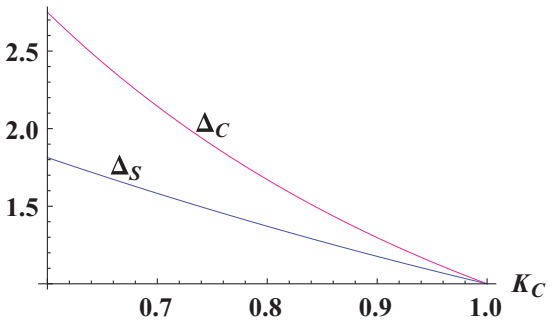

FIG. 5. (Color online) Projections of the mean-field regularized charge (Δ_c) and spin (Δ_s) gaps on the $K_s = 1.0$ plane.

Figure 5 also shows that both charge and spin gaps are decreasing functions of the parameter K_c. Since K_c decreases with increasing g couplings [cf. Eqs. (56)–(58)], our results show that the gaps increase with the g couplings, i.e., are robust against electronic interactions. Note that the impact of electronic interactions is particularly strong on the charge gap, which, in the considered range of parameters, more than doubles by increasing the strength of electronic interactions.

Let us proceed to the calculation of the single-particle gap which, in the experimentally relevant case of electron (hole) transport through a quantum wire, defines the characteristic energy scale of the system. In particular, this is the gap that determines the current blockade effect, the key element in a spin transistor design based on a modulated Rashba interaction.

Let us first recapitulate the formal definitions for single-particle gaps following the classification used in Ref. 44. The single-particle gaps are defined as the energies necessary to add to the system one electron or one hole with spin projection $\sigma \equiv \pm 1/2$:

$$\Delta_\sigma^+ = [E_0(N+1,\sigma) - E_0(N, S=0)], \tag{63}$$

$$\Delta_\sigma^- = [E_0(N-1,\sigma) - E_0(N, S=0)], \tag{64}$$

In the case of noninteracting electrons, studied in Sec. III B, we identified the single-particle gaps with the mass M_R of the fermionic quasiparticles (massive sine-Gordon solitons and antisolitons in the $\tau = \pm$ basis): $\Delta_\tau = M_R$.[45] We are now equipped to take on the calculation of the single-particle gap in the presence of electronic interactions. As we will not be able to resolve the particle and hole gaps in Eqs. (63) and (64), we instead focus on the *average* single-particle gap

$$\bar{M} = \tfrac{1}{4}(\Delta_\sigma^+ + \Delta_{-\sigma}^+ + \Delta_{\sigma'}^+ + \Delta_{\sigma'}^-), \tag{65}$$

with $\Delta_\sigma^+ + \Delta_{-\sigma}^+$ corresponding to the energy required to add two particles with opposite spin, and $\Delta_{\sigma'}^+ + \Delta_{\sigma'}^-$ the energy to add a particle and a hole with the same spin. Now, as we saw in Sec. III C, a charge soliton (of mass Δ_c) is precisely built from a pair of fermions carrying opposite spin, with a spin soliton (of mass Δ_s) being composed of a particle-hole pair in a spin triplet state. While these properties were established

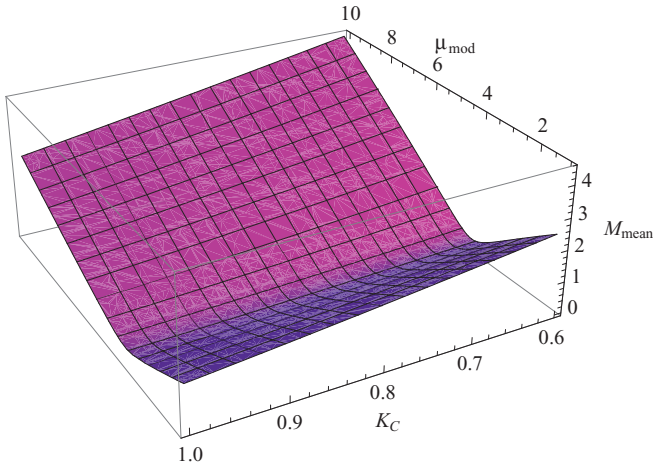

FIG. 6. (Color online) The mean-field value of $\bar{M}_{\rm mean}$ (meV) of the single-particle gap as a function of the parameters K_c and $\mu_{\rm mod}$ in the range $0.6 \leqslant K_c \leqslant 1.0$ for $K_s = 1.1$ and with $\lambda_R = -2$ meV, $\nu = 0.04$, $\Lambda = 100$ meV, and 1 meV $\leqslant \mu_{\rm mod} \leqslant 10$ meV.

for the case of noninteracting electrons, the generalizations of Eqs. (38) and (39) to the case of rescaled fields,

$$S_z = \sqrt{\frac{K_s}{2\pi}} \int_{-\infty}^{\infty} dx\, \partial_x \varphi_s(x), \tag{66}$$

$$Q = \sqrt{\frac{2K_c}{\pi}} \int_{-\infty}^{\infty} dx\, \partial_x \varphi_c(x), \tag{67}$$

show that the relation between the gaps as determined by the assignment of quantum numbers are unchanged by electron interactions. It follows from Eq. (65) that

$$\bar{M} = \tfrac{1}{4}(\Delta_c + \Delta_s), \tag{68}$$

from which we infer—with the help of Eqs. (59) and (62)—the mean-field (average) single-particle gap

$$\bar{M}_{\rm mean} = \kappa(K_c, K_s)\Lambda\,(2M_R/\Lambda)^{2/(4-K_c-K_s)}, \tag{69}$$

with

$$\kappa(K_c, K_s) \equiv \tfrac{1}{4}\eta_c^{-1}(1,1)[\eta_c(K_c, K_s) + \eta_s(K_c, K_s)], \tag{70}$$

and where M_R is given by Eq. (23). Equation (69) is the key result on which we shall build our analysis in the next section. In the limiting case of noninteracting electrons, where $K_c = K_s = 1$, $\eta_c = \eta_s$, and $\kappa(K_c, K_s) = 1/2$, we obtain, from Eq. (69), $\bar{M}_{\rm mean} = M_R$ and thus recover the result of Sec. III B.

In Fig. 6 we have plotted $\bar{M}_{\rm mean}$ in the range $0.6 \leqslant K_c \leqslant 1.0$, for $K_s = 1.1$ and with $\lambda_R = -2$ meV, $\nu = 0.04$, $\Lambda = 100$ meV, and 1 meV $\leqslant \mu_{\rm mod} \leqslant 10$ meV. (The previous values correspond to the case study carried out in Sec. V.) Again, note the significant effect of the electron-electron interactions on the size of the single-particle gap: decreasing the strength of electronic interactions, the gap also decreases.

V. APPLICATION: TOWARD A NEW TYPE OF CURRENT SWITCH?

In Ref. 5 it was argued that the insulating gap that opens when the Rashba modulation becomes commensurate with the band filling is sufficiently large to be exploited in a spin transistor design. When the modulation is turned off, the electrons are free to move and will carry a current when a drain-to-source voltage is applied. By charging the gates—thus turning on the modulation—the system becomes insulating and the current gets blocked. A rough estimate in Ref. 5, using data for a gated InAs heterostrucuture,[16] suggested a drain-to-source threshold voltage of the order of 100 mV. In this section we revisit the problem, now equipped with a more complete theory that sports a refined formula for the single-particle gap, Eq. (69), as well as a description of effects from the concurrent modulation of the chemical potential.

For a new type of current switch to become competitive it is essential that the ON-OFF switching time τ compares favorably with that of the ubiquitous metal-oxide-semiconductor field-effect transistors (MOSFETs) that are used in present-day electronics. Since τ grows with the applied gate voltage (with the power dissipation during switching being proportional to the square of the gate voltage), we shall consider the case where the device operates at or below a typical gate voltage of a MOSFET, in the 1-V range or lower. While the gate-controlled built-in electric fields in a doped heterostructure can be quite strong, the issue is whether the resulting Rashba effect—now combined with the chemical potential modulation—can become sufficiently large to be used as a switch with a gate voltage of this moderate size. An important constraint is here that leakage currents must be prevented in the OFF state. The present dominance of silicon complementary metal-oxide semiconductor (CMOS) devices for current switching is largely due to the fact that there is virtually no leakage current in the OFF state of a MOSFET, effectively protecting against unwanted signals as well as against standby power dissipation. Since an assessment of possible sources of leakage currents can only be made on basis of a specific technical design, we will not be able to fully address this constraint here. However, the minimal requirement that *thermal leakage* of charge must be prevented will be an important bench mark in our analysis. At room temperature, it translates into the requirement that the gap should be >25 meV. Clearly, this is a lower bound. Heating of the device, as well as the requirement that the source-to-drain voltage must not be too small, points to a minimum gap around, maybe, 100 meV. This is the number quoted in Ref. 5, but can it be reproduced within our more elaborate theory?

To find out, we must assign numbers to the parameters M_R, K_c, K_s, and Λ that enter the expression for the single-particle gap in Eq. (69). We shall use the same data source as in Ref. 5, obtained from the experimental work by Grundler on gate-controlled Rashba interaction in a square asymmetric InAs quantum well.[16] By adjusting the gate voltage appropriately after application of an LED pulse (which increases the 2D carrier density via the persistent photo effect), Grundler succeeded to tune the Rashba spin splitting without charging the 2D channel, thus allowing for a direct probe of the gate-voltage dependence of the Rashba parameter α. For our purpose, this is an important feature, since in our model we treat $\alpha_{\rm Rmod}$ as being independent of the electron density. Let us add that quantum wells based on InAs, realized in $In_{1-x}Ga_xAs/In_{1-x}Al_xAs$[16,46–48] or InAs/AlSb[49] heterostructures, are preferred choices in many proposals for spintronics applications due to their typical large Rashba

couplings $\alpha \approx 5 - 10 \times 10^{-12}$ eVm.[46] Moreover, unless the quantum well is designed with a very small valence band offset, it is also safe to assume that $\beta \ll \alpha$, with β the Dresselhaus coupling.[50] As we have found that the Dresselhaus interaction reduces the size of the insulating gap, this is an additional desired feature of the InAs quantum well probed in the experiment by Grundler.[16] In what follows we assume that the heterostructure studied in Ref. 16 has been gated to define a single-channel micron-range ballistic quantum wire.

A. Band gap for noninteracting electrons

Let us start by estimating the insulating band gap for the noninteracting theory. As already discussed, the band gap is twice the single-fermion mass M_R, defined in Eq. (23). Taking $\beta \ll \alpha$, we neglect the presence of the Dresselhaus interaction completely, for which case $\lambda_R \approx 2\gamma_{R\mathrm{mod}} \sin(q_0 a)$. By inspection of Eq. (23), an estimate of M_R then requires numbers for $\gamma_{R\mathrm{mod}}$, μ_{mod}, v, and $q_0 a$.

Beginning with $\gamma_{R\mathrm{mod}}$, we assume that this amplitude depends on the voltage of the small periodically spaced gates that produce the modulation (see Fig. 1) in the same way as γ_R depends on a uniform gate voltage. This is a reasonable assumption since—neglecting random fluctuations from dopant ions[17]—the internal electric field in the quantum well that supports the Rashba interaction is primarily determined by the slope of the band edge along the growth direction of the heterostructure (perpendicular to the 2D InAs quantum well interface), and hence its (extremal) value right below the center of one of the small periodically spaced gates should approach that for the case of a single large gate. Inspection of Fig. 2(b) in Ref. 16 reveals that the Rashba interaction α in the device considered *decreases* by roughly 2×10^{-11} eVm with an increase in gate voltage of 0.1 V, indicating that the Rashba and the chemical potential modulations are out of phase by π. As an aside, note that the data in Fig. 2(b) in Ref. 16 confirms the theoretical expectation that the change of the Rashba coupling with applied gate voltage is *linear*, a fact that we used in Sec. III E when analyzing the functional dependence of the effective band gap. Translating to our geometry, the decrease of the Rashba interaction with a positive increase of gate voltage implies that there is a minimum negative offset, call it $-\alpha_0$, from the uniform value $|\alpha_R|$ with no gates present. This is due to the fact that the transverse component of the net gate electric field (which controls the Rashba interaction) has a nonzero value at the midpoint between two gates. In other words, the maximum value of the *total* Rashba interaction (uniform + modulated) in the presence of the small gates is given by $\alpha_R - \alpha_0$ and is attained at the midpoint between two gates. Using the data quoted from Ref. 16, it follows for the amplitude of the Rashba modulation that $|\alpha_{R\mathrm{mod}}| = 1 \times 10^{-11} - \alpha_0/2$ eVm. The magnitude of the offset α_0 is small compared to $|\alpha_{R\mathrm{mod}}|$ and hence we take $|\alpha_{R\mathrm{mod}}| \approx 1 \times 10^{-11}$ eVm. We then have that $|\gamma_{R\mathrm{mod}}| = |\alpha_{R\mathrm{mod}}|/a \approx 20$ meV, where $a \approx 5$ Å is the lattice spacing in epitaxial InAs.[51] With $q_0 a \approx 0.1$, this in turn implies that $|\lambda_R| \approx 2$ meV.

Turning to the amplitude μ_{mod} of the modulated chemical potential, it is more difficult to obtain an accurate number here and we will have to do with a rough estimate. There are two types of contributions to μ_{mod} from the external gates; one coming from the transverse component of the applied gate electric field (enabling a local migration of charge from the quantum well into the dopant layers) and the other from its longitudinal component, inducing a rearrangement of charge inside the quantum wire. We expect the latter to dominate the modulation of the local chemical potential and neglect the small residual charge migration caused by the transverse electric field. As for the variation of the longitudinal electric field along the wire, an exact expression requires a precise description of the fringe fields and their superpositions from the periodic sequence of top gates. This goes far beyond the scope of our minimal approach, where we assume the simplest possible, but still physically meaningful, behavior: a longitudinal electric field that oscillates harmonically along the wire: $\boldsymbol{E}_{\parallel}(x) = E_{\parallel} \sin(qx)\hat{x}$, choosing $x = 0$ just below the center of one of the top gates. The variation of the chemical potential $\Delta\mu(x)$ along the wire is then related to the negative work required to bring an electron to x,

$$\Delta\mu(x) = -eE_{\parallel} \int_0^x \sin(qx')\,dx',$$
$$\mu(x) = -\mu_{\mathrm{mod}} \cos(qx), \tag{71}$$

where $\mu_{\mathrm{mod}} = -\mu(0) = -eE_{\parallel}/q > 0$. Note that the chemical potential variation of Eq. (71) is indeed the one assumed in the very construction of our model, Eq. (7).

With the assumption that the amplitudes of the longitudinal and transverse components of the net gate electric field are of the same order of magnitude, we make the approximation $E_{\parallel} \approx V_g/d$, where V_g is the voltage of a gate electrode and d is the perpendicular distance between the gate and the wire. Thus, $\mu_{\mathrm{mod}} = -eV_g/qd$. From the commensurability relation $q = 2k_F = 2\sqrt{2\pi n_s}$, with n_s the 2D electron density of the InAs quantum well, we arrive at the desired expression for μ_{mod} is terms of experimental parameters:

$$\mu_{\mathrm{mod}} = \frac{-eV_g}{2d\sqrt{2\pi n_s}}. \tag{72}$$

From Ref. 16 we have that $d \approx 60$ nm and $n_s \approx 0.9 \times 10^{16}$ m^{-2}. Using a gate voltage $V_g \approx +0.1$ V, we obtain that $\mu_{\mathrm{mod}} \approx 4$ meV. This number will get modified when including effects from Fermi statistics, Coulomb interaction, and the presence of the transverse component of the applied gate electric field. However, guided by experimentally inferred Fermi level variations with gate voltage in other InAs quantum wells,[46–48] we expect that our estimate of μ_{mod} for the geometry considered and with the type of heterostructure used in Ref. 16 lies within reasonable bounds. To have some margin for error, we must actually quote μ_{mod} only as taking possible values in a range including 4 meV. Going back to Fig. 6, we see that the single particle gap in the presence of electron-electron interactions is not a monotonic function of μ_{mod}, displaying a minimum exactly around 4 meV. So, restricting the error bar to lie inside the regime of interest (from the point of view of spintronics applications) of increasing gap, we consider μ_{mod} assuming values in the range from 4 to 10 meV.

In order to obtain a value for M_R in Eq. (23) (or, rather, an interval for possible values for M_R, considering the numerical

uncertainty in the $\mu_{\rm mod}$ parameter), the 1D band filling ν remains to be determined. With the assumption that the quantum wire has only a single conducting channel, this is straightforward. The data for the Rashba variation with top gate voltage in Ref. 16 were taken at a constant electron density $n_s \approx 0.9 \times 10^{16}$ m^{-2}. With the Fermi wave number k_F for the quantum wire expected to be roughly the same as for the 2D electron gas, this translates, via $k_F = \pi\nu/a = \sqrt{2\pi n_s}$, into:

$$\nu = a\sqrt{\frac{2n_s}{\pi}}. \tag{73}$$

Putting in the numbers, we get $\nu \approx 0.04$.

Inserting our estimates

$$\lambda_{\rm R} \approx -2\,{\rm meV}, \quad 4\,{\rm meV} < \mu_{\rm mod} < 10\,{\rm meV}, \quad \nu \approx 0.04,$$

into Eq. (23), we finally obtain that

$$0.3\,{\rm meV} \lesssim M_R \lesssim 3.0\,{\rm meV}, \tag{74}$$

with the upper bound corresponding to $\mu_{\rm mod} \approx 10$ meV and with the lower bound attained for $\mu_{\rm mod} \approx 4$ meV. Note the minus sign in $\lambda_R \approx -2$ meV, indicating that the two modulations in the type of device considered are antiphased, which is the reason for the nonmonotonic behavior of the gap as a function of $\mu_{\rm mod}$ (cf. discussion in Sec. III E).

We should here point out that given the value of n_s in Ref. 16, our use of a *single* 1D conducting channel is not an unreasonable assumption. A first estimate, using an infinite-well confinement potential may suggest that a quantum wire with diameter $D \lesssim 25$ nm would satisfy the single-channel condition: $\lambda_F = 2\pi/k_F = \sqrt{2\pi/n_s} \approx 25$ nm. However, with the Fermi energy $E_F = \hbar^2\pi n_s/m^* \approx 40$ meV (with $m^* = 0.04m_e$ for a gated InAs quantum wire,[16] where m_e is the electron mass), self-consistency requires that the wire is not much wider than roughly $D/2$ since otherwise the Fermi level would cut through the first sub-band. In addition, for our modeling to make sense, our energy cutoff Λ, first introduced in Eq. (47), must be smaller than the distance ΔE from the Fermi level to the bottom of the first sub-band. Below we shall choose $\Lambda \approx 100$ meV, which—again assuming an infinite-well confinement with $\Delta E = \pi^2\hbar^2/2m^*D^2$—requires the wire to be at most 7–8 nm wide, still, however, within the realm of present-day technology.[52] One should note that a more realistic soft confinement potential could possibly open an additional conducting channel and lead to population of the first 1D sub-band, in which case our model would no longer apply. However, as the gate-controlled Rashba effect in Ref. 16 appears to be rather insensitive to a lowering of the value of n_s, this potential problem should, in principle, be easy to overcome [cf. Fig. 2(b) in Ref. 16].

B. Band gap for interacting electrons

To complete our analysis of the single-fermion mass, we need to include the effect of electron-electron interactions. These are encoded by the Luttinger liquid charge and spin parameters K_c and K_s. Considering that the forward-scattering amplitudes in Eq. (56) are all equal and given by $k \sim 0$ (zero momentum transfer) Fourier component $V(k \sim 0)$ of the screened Coulomb interaction,[30] $g_{2\parallel} = g_{2\perp} = g_{4\parallel} = g_{4\perp} \approx V(k \sim 0)/\hbar$, we obtain from Eq. (56), neglecting the small correction from backscattering processes with momentum transfer $k \sim 2k_F$,

$$K_c^{-1} = \sqrt{1 + \frac{2V(k \sim 0)}{\hbar\pi v_F}}. \tag{75}$$

The screening length of the interaction is roughly set by the perpendicular distance d between the quantum wire and the nearest metallic gate. A detailed analysis[53] leads to the expression

$$V(k \sim 0) \approx \frac{e^2}{\pi\epsilon_0\epsilon_r} \ln\left(\frac{2d'}{\xi}\right) + \mathcal{O}\left(\frac{\xi^2}{d^2}\right), \tag{76}$$

where ξ is the radius of the quantum wire and ϵ_r is the averaged relative permittivity of the dopant and capping layers between the quantum well and the nearest gate at a distance d' from the wire. As an interesting aside, note that the leading logarithmic term in Eq. (76) depends only on the permittivity of the environment and not on that of the wire, implying that electrons interact mainly with image charges and not with other electrons in the wire. With the backgate of the device in Ref. 16 being at a distance $d' \approx 15$ nm from the quantum well, and with an averaged permittivity $\epsilon_r \approx 12$ for the interjacent $In_{0.75}Al_{0.25}As$ and Si-doped $In_{0.75}Al_{0.25}As$ layers,[54] we obtain from Eqs. (75) and (76) that $K_c \approx 0.7$, taking $\xi \approx 5$ nm and considering that $v_F \approx 6 \times 10^5$ m/s.[54] We should alert the reader to the fact that the estimate for K_c also comes with some uncertainty, considering that it is obtained using the parameterization in Eq. (56), which is strictly valid only in the weak-coupling limit $K_c \approx 1$. Still, the *Bethe ansatz* and numerical results for this class of models have shown that the weak-coupling formula in Eq. (56) does surprisingly well in capturing effective K_c parameters also for intermediate strengths of the electron interaction when, as in the present case, the band filling is low, thus providing indirect support for our estimate.[55]

Turning to the spin parameter K_s, we already noted in Sec. IV [text after Eq. (58)] that its bare value predicted by Eq. (56) will renormalize to a value slightly larger than unity due to backscattering of electrons. Having ignored these scattering processes when writing down the spin Hamiltonian in Eq. (54) with the rationale that they are weak in a semiconductor device, we may compensate for their omission by adjusting the value of K_s by hand, setting it slightly larger than unity, say at $K_s \approx 1.1$. Guided by work on other models where K_s takes values that differ from unity we expect this to be a reasonable estimate.[21] In what follows K_s thus represents the expected RG fixed point value K_s^*, carrying an imprint of the marginally irrelevant backscattering term in the spin sector. It may be worth pointing out that the correction to K_c due to backscattering is smaller than that for K_s and is thus neglected.

Having put numbers on K_c and K_s we now go back to Eq. (60) and calculate, with the help of Eqs. (A4) and (A6) in the Appendix,

$$\eta_c(K_c = 0.7, K_s = 1.1) \approx 2.1 \tag{77}$$

$$\eta_s(K_c = 0.7, K_s = 1.1) \approx 1.4. \tag{78}$$

Combining this with Eqs. (61), (69), (70), and (74) and taking $\Lambda = \hbar v_F/\xi \approx 100$ meV, we finally obtain an estimate for the single-particle gap $\bar{M}_{\rm mean}$,

$$0.4 \text{ meV} \lesssim \bar{M}_{\rm mean} \lesssim 4 \text{ meV}. \tag{79}$$

Summarizing its specification, this result applies to a periodically gated 5-nm thin quantum wire embedded in a Rashba-active heterostructure of the type studied in Ref. 16, assuming that the Fermi wave number has been properly tuned to commensurability with the gate spacing and taking the gate voltage to be +0.1 V. Note that the length scale $\xi \sim \hbar v_F/\bar{M}_{\rm mean}$ at which the gap starts to open up lies within the interval $0.1\ \mu\text{m} \lesssim \xi \lesssim 1\ \mu$m, with the lower (upper) bound corresponding to $\bar{M}_{\rm mean} = 4$ meV (0.4 meV), thus fitting well within the ballistic regime of an InAs quantum wire.[51]

The estimate in Eq. (79) is strikingly lower than that in Ref. 5, where the same kind of system was analyzed using a simpler theory. One of the missing ingredients in that theory is the interplay between the modulation of the Rashba interaction with that of the local chemical potential, an effect that we have found to cause a significant reduction of the gap. Moreover, the gap in Ref. 16 was extracted from that of the collective charge excitations of the low-energy effective model, and not, as in the present approach, properly reconstructed as a *single-particle gap*, which is the one relevant for charge transport.[56]

One notes that the gap at the upper bound in Eq. (79) is far below the thermal threshold of ∼25 meV that is required to block thermal leakage of charge—a *sine qua non* for a functioning current switch. Thus, a usable device based on spin-orbit and charge modulation effects will clearly require a different type of heterostructure and/or input parameters than what we have assumed here. A significant improvement would be achieved if, by "band engineering,"[25,26] one could grow a heterostructure where the Rashba and chemical potential modulations are *in phase*, not out of phase as in the case study above. As evidenced by Eq. (23) and discussed in the context of Fig. 2, this will boost the resulting gap, especially for the case of low electronic density. Second, our analysis shows the importance of having a large electron-electron interaction. This, in principle, is obtainable by further reducing the electron density, with the added advantage of making the required gate spacing larger (as seen in the commensurability condition $q = 2k_F$ with $q = 2\pi/\lambda$ and $k_F = \pi\nu/a$), thus giving leave for larger and experimentally more tractable gates. Finally, and most obviously, by allowing for a larger gate bias than the 0.1 V used in the estimates above, the gap opening effect will be further boosted. Still, unless one allows for very large voltages, in the 5- to 10-V range, the gap—*as estimated within our formalism and for the system specifications used here*—will be too small for making a credible case for a working current switch at room temperature. At these large voltages, however, our proposed device would have no clear advantage compared to standard silicon CMOS designs. Moreover, since the growth of the modulated Rashba coupling will have saturated at much smaller voltages, any additional gap opening effect will primarily be due to the CDW correlations from the modulated chemical potential, not to the presence of a Rashba interaction.

Before closing the case, however, we wish to stress that our numerical estimates have been obtained by filtering experimental data through an effective low-energy field theory formalism based on a highly simplified lattice model. We have tried to be careful in processing the data; however, the approach we use is not optimally adapted for this task. As we have repeatedly pointed out, this makes our numbers marred with uncertainty. Whereas our theory does provide a "proof-of-concept" of using a periodically gated quantum wire for a low-bias current switch, a definite verdict about its practicability requires more work, based on a more sophisticated approach in modeling and analysis.

VI. SUMMARY

In conclusion, we have analyzed the spin and charge dynamics in a ballistic single-channel quantum wire in the presence of a gate-controlled harmonically modulated Rashba spin-orbit interaction, and with a concurrent harmonic modulation of the local chemical potential. To be able to model a quantum wire in a gated heterostructure with lattice inversion asymmetry, we have also allowed for a uniform Dresselhaus spin-orbit interaction.

Depending on the relations among the common wave number q of the two modulations, the Fermi momentum k_F, and a parameter q_0 that encodes the strength of the Dresselhaus and the uniform part of the Rashba interaction, the electrons in the wire may form a metallic or an insulating state. Specifically, and most interesting from the viewpoint of potential spintronics applications, when $|q - 2k_F| \ll \mathcal{O}(1/a)$ and $|q \pm 2q_0| \simeq \mathcal{O}(1/a)$ (with a the lattice spacing), a *nonmagnetic insulating state* is formed, with an effective band gap that depends on the amplitudes of the Rashba and chemical potential modulations as well as on the strengths of the uniform Dresselhaus and Rashba interactions. Whereas the Dresselhaus interaction *reduces* the band gap, the uniform part of the Rashba interaction *increases* its size. The gap also increases with the amplitude of the *modulated* part of the Rashba interaction, but only if the local chemical potential modulation is in phase with that of the Rashba interaction, *or* if the amplitude of the Rashba modulation is larger than some threshold value of the chemical potential modulation. Otherwise, the gap is a decreasing function of the Rashba modulation amplitude. The resulting crossover behavior of the effective band gap is controlled by the band filling, which sets the threshold value of the chemical potential modulation.

This gap-opening scenario, including the crossover behavior, is found to be robust against electron-electron interactions. To arrive at this conclusion we used a bosonization approach, mapping the interacting problem onto two mean-field decoupled sine-Gordon models. By a careful analysis of the structure of the ensuing charge and spin gaps, we devised a regularization scheme from which the size of the *single-particle gap* can be reconstructed and that allowed us to determine its dependence on the strength of the electron-electron interaction. Exploiting exact results for the sine-Gordon model, we found that the gap scales as $M_R^{2/(4-K_c-K_s)}$, where M_R is the sine-Gordon soliton (or antisoliton) mass for noninteracting electrons and where K_c and K_s are the Luttinger liquid charge and spin parameters, respectively. Whereas the scaling exponent agrees with that found in Ref. 5, our estimate

for the band gap (using data from the same experimental setup as that studied in Ref. 5) comes out dramatically smaller: As discussed in the previous section, the theory in Ref. 5 does not include the competition between Rashba and chemical potential modulations in the experimentally relevant parameter regime, and moreover, the band gap is not properly reconstructed as a single-particle gap from the collective spin and charge excitations in the bosonic spin-charge basis.[56]

While our analysis reveals shortcomings with proposals[5–7] to use present-day materials and designs for constructing a low-bias current switch from a gate-controlled modulated Rashba interaction, it also points to possible routes to overcome the problem. In addition to the obvious measure to search for materials with larger Rashba couplings, we have shown the importance to engineer heterostructures where the gate-controlled Rashba modulation is *in phase* with that of the local chemical potential produced by the gate configuration.[25,26] We have also shown that the size of the effective band gap can be significantly boosted by reducing the electron density in the quantum wire; this leads to a reduction of the screening of the electron-electron interaction and, with that, a larger gap-opening effect from the Rashba modulation.

From a more fundamental perspective, questions about how the gap-opening scenario is influenced by disorder,[57] magnetic field effects,[58] and sub-band mixing[19] are yet to be addressed. These become challenging problems in the context of a spatially varying spin-orbit interaction and may require a theoretical approach that goes beyond the effective field-theory approach that we have used here. Even with these questions unanswered, however, our prediction of an electrically driven commensurate-to-incommensurate phase transition is amenable to an experimental test. Of particular interest would be to test for the rigidity of the insulating state away from commensurability. As discussed in Sec. III, its robustness is determined by the size of the effective band gap and will thus be sensitive to the screening of the electron-electron interaction and hence to the density of electrons in the wire. By preparing setups with different modulation wave numbers and electron densities—but otherwise identical—an experiment should see an increase of the gap with lowered electron density, as predicted in our Eq. (69). Our prediction that the conductivity close to the transition scales with the chemical potential with a universal critical exponent 1/2 *independent of electron-electron interactions* is also open for experimental probes.

Considering the complexity of the spin and charge dynamics in a quantum wire when subject to a modulated Rashba spin-orbit interaction, further work—experimental as well as theoretical—may well uncover hidden features of this fascinating physical system.

ACKNOWLEDGMENTS

We thank Alvaro Ferraz for valuable discussions and U. Ekenberg for a helpful communication. This work was supported by the Brazilian CNPq and Ministry of Science and Technology (M.M.), Georgian NSF Grant No. ST09/09-447 and SCOPES Grant No. IZ73Z0-128058 (I.G. and G.I.J.), and Swedish Research Council Grant No. 621-2008-4358 (H.J.).

APPENDIX: MASS SCALES AND EXPECTATION VALUES IN THE SINE-GORDON MODEL

In this Appendix we show how to obtain Eqs. (47) and (48) from the results on the sine-Gordon model in Refs. 41 and 42. Following the convention in Ref. 42, we write the Euclidean action of the sine-Gordon model as

$$\mathcal{A}_{SG} = \int d^2x \left[\frac{1}{16\pi}(\partial_v \varphi)^2 - \frac{2\mu_0}{a^2}\cos(\beta\varphi) \right]. \quad \text{(A1)}$$

where μ_0 and β are dimensionless parameters, with $0 < \beta^2 < 1$. (Note that the presence of the prefactor $1/16\pi$ in the kinetic term of the action implies that the sine-Gordon coupling β differs by a factor of $\sqrt{8\pi}$ from the conventional one.[32] Also note that by defining $\mu \equiv \mu_0/a^2$, we have isolated the engineering dimension $1/(\text{length})^2$ of the bare mass μ in Ref. 42 in the square of the microscopic length a. Introducing a velocity parameter v via $d^2x \to v d\tau dx$ and rescaling the field, $\varphi \to \sqrt{8\pi}\varphi$, the corresponding Hamiltonian reads

$$H = \int dx \left\{ \frac{v}{2}[(\partial_x \vartheta)^2 + (\partial_x \varphi)^2] - \frac{2\mu_0 v}{a^2}\cos(\sqrt{8\pi}\beta\varphi) \right\} \quad \text{(A2)}$$

with $\partial_x \vartheta$ the conjugate momentum to φ. By the substitutions $\beta^2 = K_c/4$, $v = v_c$, and $2\mu_0 = m_c/\pi\Lambda_c$, with $\Lambda_c = v_c/a$ a UV cutoff, we recover the charge-sector mean-field Hamiltonian in Eq. (53) (when $\mu_{\text{eff}} = 0$). Similarly, the spin-sector mean-field Hamiltonian in Eq. (54) is obtained via the substitutions $\beta^2 = K_s/4$, $v = v_s$, and $2\mu_0 = m_s/\pi\Lambda_s$, with $\Lambda_s = v_s/a$, together with the phase shift $\varphi_s \to \varphi_s + \sqrt{\pi/8K_s}$ (which does not affect the renormalization of the theory).

The key result in Ref. 41 [encoded in Eqs. (2.12) and (4.1) in the same reference], which relates the sine-Gordon soliton mass to the bare mass parameter, can now be rephrased as

$$M_i/\Lambda = C(K_i)\left(\frac{m_i}{\Lambda}\right)^{2/(4-K_i)}, \quad i = c,s, \quad \text{(A3)}$$

where

$$C(K_i) = \frac{2}{\sqrt{\pi}}\frac{\Gamma\left(\frac{K_i}{8-2K_i}\right)}{\Gamma\left(\frac{2}{4-K_i}\right)}\left[\frac{\Gamma(1-K_i/4)}{2\Gamma(K_i/4)}\right]^{2/(4-K_i)}. \quad \text{(A4)}$$

Here Γ is the Γ function. To extract Eqs. (A3) and (A4) from Eqs. (2.12) and (4.1) in Ref. 41 we have used that $p = (2 - K_i)/2$ in these equations and also that the sine-Gordon action in Eq. (2.1) in Ref. 41 is the same as that in Eq. (A1) after having rescaled the field, $\varphi \to \sqrt{8\pi}\varphi$, and put $\mu_0 = \mu a^2$.

To obtain Eqs. (47) and (48), we also need to relate the soliton mass to the ground-state expectation value of the cosine field; recall from Eqs. (43) and (44) that the mean-field bare mass parameters m_c and m_s are defined in terms of $\langle\cos\sqrt{2\pi K_s}\varphi_s\rangle$ and $\langle\cos\sqrt{2\pi K_c}\varphi_c\rangle$, respectively. For this, we turn to Eq. (15) in Ref. 42, from which we infer

$$\langle\cos(\sqrt{2\pi K_i}\varphi_i)\rangle = B(K_i)(M/\Lambda)^{K_i/2}, \quad \text{(A5)}$$

where

$$B(K_i) = \pi^2[\Gamma(1/2+\xi_i/2)\Gamma(1-\xi_i/2)]^{(K_i-4)/2} \times \left[\frac{\sin(\pi\xi_i/2)}{2\sqrt{\pi}}\right]^{K_i/2} \frac{(1+\xi_i)\Gamma(1-K_i/4)}{\sin(\pi\xi_i)\Gamma(K_i/4)} \quad \text{(A6)}$$

with $\xi_i = K_i/(4-K_i)$. Combining Eqs. (A3)–(A6), we obtain the following expressions for the charge and spin soliton masses,

$$\frac{M_c}{\Lambda} = C_c\left[\frac{2M_R\langle\cos(\sqrt{2\pi K_s}\varphi_s)\rangle}{\Lambda}\right]^{2/\zeta_c} = C_c B_s^{2/\zeta_c}\left(\frac{2M_R}{\Lambda}\right)^{2/\zeta_c}\left(\frac{M_s}{\Lambda}\right)^{K_s/\zeta_c}, \quad \text{(A7)}$$

$$\frac{M_s}{\Lambda} = C_s\left[\frac{2M_R\langle\cos(\sqrt{2\pi K_c}\varphi_c)\rangle}{\Lambda}\right]^{2/\zeta_s} = C_s B_c^{2/\zeta_s}\left(\frac{2M_R}{\Lambda}\right)^{2/\zeta_s}\left(\frac{M_c}{\Lambda}\right)^{K_c/\zeta_s}, \quad \text{(A8)}$$

where $\zeta_i = 4-K_i, B_i = B(K_i)$, and $C_i = C(K_i), i = c,s$. Some straightforward algebra on Eqs. (A7) and (A8) finally yields Eqs. (59) and (60). In the important limiting case of noninteracting electrons, i.e., with $K_c = K_s = 1$, we have that $B(1) \equiv \mathcal{C}_1 \approx 1.0$ and $C(1) \equiv \mathcal{C}_0 \approx 1.4$ (cf. Sec. III D).

[1] I. Zutic, J. Fabian, and S. Das Sarma, Rev. Mod. Phys. **76**, 323 (2004).

[2] S. Datta and B. Das, Appl. Phys. Lett. **56**, 665 (1990).

[3] E. I. Rashba, Sov. Phys. Solid State **2**, 1109 (1960); Y. A. Bychkov and E. I. Rashba, J. Phys. C **17**, 6039 (1984).

[4] J. Fabian and I. Zutic, in *The Standard Model of Spin Injection, Lecture Notes of the 40th Spring School 2009: Spintronics–From GMR to Quantum Information*, edited by S. Blügel, D. Bürgler, M. Morgenstern, C. M. Schneider, and R. Waser (Institute of Solid State Research of the Forschungszentrum, Jülich, Germany, 2009).

[5] G. I. Japaridze, H. Johannesson, and A. Ferraz, Phys. Rev. B **80**, 041308(R) (2009).

[6] X. F. Wang, Phys. Rev. B **69**, 035302 (2004).

[7] S. J. Gong and Z. Q. Yang, J. Phys. Condens. Matter **19**, 446209 (2007).

[8] For other work on 1D charge/spin transport in the presence of a modulated Rashba interactions, see F. Mireles and G. Kirczenow, Phys. Rev. B **64**, 024426 (2001); X. F. Wang, P. Vasilopoulos, and F. M. Peeters, *ibid.* **71**, 125301 (2005); L. Zhang, P. Brusheim, and H. Q. Xu, *ibid.* **72**, 045347 (2005); L.-G. Wang, K. Chang, and K.-S. Chan, J. Appl. Phys. **99**, 043701 (2006); D. Sanchez and L. Serra, Phys. Rev. B **74**, 153313 (2006); D. Sanchez, L. Serra, and M.-S. Choi, *ibid.* **77**, 035315 (2008); X. B. Xiao and Y. G. Chen, Europhys. Lett. **90**, 47004 (2010).

[9] G. Dresselhaus, Phys. Rev. **100**, 580 (1955).

[10] For a review, see, e.g., P. Fulde, *Electron Correlations in Molecules and Solids* (Springer-Verlag, Berlin, 1991).

[11] W. Häusler, L. Kecke, and A. H. MacDonald, Phys. Rev. B **65**, 085104 (2002).

[12] M. I. Dyakonov and V. Y. Kachorovskii, Sov. Phys. Semicond. **20**, 110 (1986); E. I. Rashba and E. Ya. Sherman, Phys. Lett. A **129**, 175 (1988).

[13] E. Ya. Sherman, Phys. Rev. B **67**, 161303(R) (2003).

[14] L. E. Golub and E. L. Ivchenko, Phys. Rev. B **69**, 115333 (2004).

[15] For a review, see U. Ekenberg and D. M. Gvozdić, in *Quantum Wells: Theory, Fabrication and Applications*, edited by A. Ruyter and H. Mahoney (Nova Science, Hauppauge, 2010).

[16] D. Grundler, Phys. Rev. Lett. **84**, 6074 (2000).

[17] A. Ström, H. Johannesson, and G. I. Japaridze, Phys. Rev. Lett. **104**, 256804 (2010).

[18] For a tabulation of experimentally inferred Rashba and Dresselhaus couplings, see R. H. Silsbee, J. Phys. Condens. Matter **16**, R179 (2004).

[19] A. V. Moroz and C. H. W. Barnes, Phys. Rev. B **60**, 14272 (1999); **61**, R2464 (2000); E. A. de Andrada e Silva and G. C. La Rocca, *ibid.* **67**, 165318 (2003); S. L. Erlingsson, J. C. Egues, and D. Loss, Phys. Status Solidi C **3**, 4317 (2006); C. A. Perroni, D. Bercioux, V. Marigliano Ramaglia, and V. Cataudella, J. Phys. Condens. Matter **19**, 186227 (2007).

[20] W. Häusler, Phys. Rev. B **63**, 121310(R) (2001); M. Governale and U. Zülicke, *ibid.* **66**, 073311 (2002); Solid State Commun. **131**, 581 (2004); A. V. Moroz, K. V. Samokhin, and C. H. W. Barnes, Phys. Rev. B **62**, 16900 (2000); A. De Martino and R. Egger, Europhys. Lett. **56**, 570 (2001); A. Iucci, Phys. Rev. B **68**, 075107 (2003); W. Häusler, *ibid.* **70**, 115313 (2004).

[21] V. Gritsev, G. I. Japaridze, M. Pletuykhov, and D. Baeriswyl, Phys. Rev. Lett. **94**, 137207 (2005).

[22] A. Schulz, A. De Martino, P. Ingenhoven, and R. Egger, Phys. Rev. B **79**, 205432 (2009).

[23] Unless otherwise stated, when henceforth referring to "chemical potential modulation" or "modulated chemical potential" we mean the function $\mu_{\rm mod}\cos(qna)$ [or, in the continuum limit, $\mu_{\rm mod}\cos(qx)$], *without* the negative sign which multiplies the chemical potential term in Eq. (7).

[24] E. Ya. Sherman and J. Sinova, Phys. Rev. B **72**, 075318 (2005); A. Brataas, A. G. Mal'shukov, and Y. Tserkovnyak, New J. Phys. **9**, 345 (2007).

[25] V. I. Litvinov, Appl. Phys. Lett. **89**, 222108 (2006).

[26] T. Matsuda and K. Yoh, Physica E **42**, 979 (2010).

[27] See M. A. Toloza Sandoval, A. Ferreira da Silva, E. A. de Andrada e Silva, and G. C. la Rocca, J. Supercond. Nov. Magn. **23**, 171 (2010) and references therein.

[28] I. V. Tokatly and E. Ya. Sherman, Phys. Rev. B **82**, 161305(R) (2010).

[29] L. S. Levitov and E. I. Rashba, Phys. Rev. B **67**, 115324 (2003).

[30] T. Giamarchi, *Quantum Physics in One Dimension* (Oxford University Press, Oxford, 2004).

[31] Note that the $j = 1$ component in Eq. (22) is also present if both the band filling and q are extremely small, implying that also $q + 2k_F \ll \mathcal{O}(1/a)$. In what follows, however, we consider the more

generic situation that the commensurability condition $|q - 2k_F| \ll \mathcal{O}(1/a)$ does not hinge on the smallness of q and k_F and that, therefore, only the $j = -1$ term in Eq. (22) is present.

[32] S. Coleman, Phys. Rev. D **11**, 2088 (1975).

[33] R. F. Dashen, B. Hasslacher, and A. Neveu, Phys. Rev. D **10**, 3424 (1975).

[34] G. I. Japaridze and A. A. Nersesyan, JETP Lett. **27**, 334 (1978); V. L. Pokrovsky and A. L. Talapov, Phys. Rev. Lett. **42**, 65 (1979).

[35] Note that the commensurability condition $q = 2k_F$ translates precisely into $n_s = \pi/2\lambda^2$ via the relations $k_F = \sqrt{2\pi n_s}$ and $q = 2\pi/\lambda$.

[36] H. J. Schulz, Phys. Rev. B **22**, 5274 (1980).

[37] M. Fabrizio, A. O. Gogolin, and A. A. Nersesyan, Phys. Rev. Lett. **83**, 2014 (1999); Nucl. Phys. B **580**, 647 (2000).

[38] A. J. Heeger, S. Kivelson, J. R. Schrieffer, and W.-P. Su, Rev. Mod. Phys. **60**, 781 (1988).

[39] A. O. Gogolin, A. A. Nersesyan, and A. M. Tsvelik, *Bosonization and Strongly Correlated Systems* (Cambridge University Press, Cambridge, UK, 1998).

[40] G. I. Japaridze, R. Hayn, P. Lombardo, and E. Müller-Hartmann, Phys. Rev. B **75**, 245122 (2007).

[41] Al. B. Zamolodchikov, Int. J. Mod. Phys. A **10**, 1125 (1995).

[42] S. Lukyanov and A. Zamolodchikov, Nucl. Phys. B **493**, 571 (1997).

[43] T. Giamarchi and A. J. Millis, Phys. Rev. B **46**, 9325 (1992).

[44] S. R. Manmana, V. Meden, R. M. Noack, and K. Schönhammer, Phys. Rev. B **70**, 155115 (2004).

[45] Note that here, for consistency with the notation in Eqs. (63) and (64), we use the index σ instead of τ for the spin projections along the combined uniform Rashba and Dresselhaus fields.

[46] J. Nitta, T. Akazaki, H. Takayanagi, and T. Enoki, Phys. Rev. Lett. **78**, 1335 (1997).

[47] C.-M. Hu, J. Nitta, T. Akazaki, H. Takayanagai, J. Osaka, P. Pfeffer, and W. Zawadzki, Phys. Rev. B **60**, 7736 (1999).

[48] T. Koga, J. Nitta, T. Akazaki, and H. Takayanagi, Phys. Rev. Lett. **89**, 046801 (2002).

[49] J. P. Heida, B. J. van Wees, J. J. Kuipers, T. M. Klapwijk, and G. Borghs, Phys. Rev. B **57**, 11911 (1998).

[50] L. Meier, G. Salis, I. Shorubalko, E. Gini, S. Schön, and K. Ensslin, Nat. Phys. **3**, 650 (2007).

[51] J. Hornstra and W. J. Bartels, J. Cryst. Growth **44**, 513 (1978).

[52] See, e.g., D. K. Ferry, R. Akis, M. J. Gilbert, A. Cummings, and S. M. Ramey, Int. J. High Speed Electr. Syst. **17**, 445 (2007).

[53] K. Byczuk and T. Dietl, Phys. Rev. B **60**, 1507 (1999).

[54] Pallab Bhattacharaya, *Properties of lattice-matched and strained indium gallium arsenide*. (EMIS Datareviews Series No. 8., INSPEC, Michael Farday House, Six Hills Way, Stevenage, Herts SG1 2AY, United Kingtom, 1993).

[55] H. J. Schulz, Int. J. Mod. Phys. B **5**, 57 (1991).

[56] The overestimate of the effective band gap in Ref. 5 is in part also due to an incorrect parametrization of the argument of a Γ function that enters the expressions for the renormalized mass scales and expectation values in the sine-Gordon model. Specifically, the expression for $\mathcal{C}_1(K_i)$ given after Eq. (13) in Ref. 5 should be replaced by the right-hand side of Eq. (A4) in the present article.

[57] S. Kettemann, Phys. Rev. Lett. **98**, 176808 (2007); M. Scheid, M. Kohda, Y. Kunihashi, K. Richter, and J. Nitta, *ibid.* **101**, 266401 (2008).

[58] Y. Yu, Y. Wen, J. Li, Z. Su, and S. T. Chui, Phys. Rev. B **69**, 153307 (2004); P. Devillard, A. Crepieux, K. I. Imura, and T. Martin, *ibid.* **72**, 041309(R) (2005); H. C. Lee and S. R. Eric Yang, *ibid.* **72**, 245338 (2005); F. Cheng and G. Zhou, J. Phys. Condens. Matter **19**, 136215 (2007); J. Sun, S. Gangadharaiah, and O. A. Starykh, Phys. Rev. Lett. **98**, 126408 (2007); S. Gangadharaiah, J. Sun, and O. A. Starykh, Phys. Rev. B **78**, 054436 (2008).