Valley Hall effect in disordered monolayer MoS$_2$ from first principles

Thomas Olsen[1,2,*] and Ivo Souza[1,3,†]

[1]*Centro de Física de Materiales, Universidad del País Vasco, 20018 San Sebastián, Spain*
[2]*Center for Atomic-Scale Materials Design, Department of Physics, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark*
[3]*Ikerbasque Foundation, 48013 Bilbao, Spain*

Electrons in certain two-dimensional crystals possess a pseudospin degree of freedom associated with the existence of two inequivalent valleys in the Brillouin zone. If, as in monolayer MoS$_2$, inversion symmetry is broken and time-reversal symmetry is present, equal and opposite amounts of k-space Berry curvature accumulate in each of the two valleys. This is conveniently quantified by the integral of the Berry curvature over a single valley—the valley Hall conductivity. We generalize this definition to include contributions from disorder described with the supercell approach, by mapping ("unfolding") the Berry curvature from the folded Brillouin zone of the disordered supercell onto the normal Brillouin zone of the pristine crystal, and then averaging over several realizations of disorder. We use this scheme to study from first principles the effect of sulfur vacancies on the valley Hall conductivity of monolayer MoS$_2$. In dirty samples the intrinsic valley Hall conductivity receives gating-dependent corrections that are only weakly dependent on the impurity concentration, consistent with side-jump scattering and the unfolded Berry curvature can be interpreted as a k-space resolved side jump. At low impurity concentrations skew scattering dominates, leading to a divergent valley Hall conductivity in the clean limit. The implications for the recently observed photoinduced anomalous Hall effect are discussed.

I. INTRODUCTION

Monolayers of MoS$_2$ and related transition-metal dichalcogenides (TMDs) have recently become the subject of intense investigation, due in part to the possibility of manipulating the so-called "valley" degree of freedom [1]. These materials have the symmetry of a honeycomb structure with a staggered sublattice, thus lacking an inversion center. The band structure exhibits a direct gap at the two inequivalent valleys centered at the high-symmetry points K and $K' = -K$ in the Brillouin zone (see Fig. 1), where the topmost valence bands are primarily composed of transition-metal d states [2]. Time-reversal symmetry, which takes **k** into −**k** and therefore maps one valley onto the other, dictates that states in a given band at K and K' carry antiparallel angular momenta. This inspired Xiao *et al.* to propose using circularly polarized light as a means of selectively exciting carriers from a particular valley [3,4]. The effect was rapidly confirmed experimentally, by demonstrating that excitation with circularly polarized light results in polarized fluorescence [5–7].

The broken inversion symmetry in monolayer MoS$_2$ induces a nonzero Berry curvature on the Bloch bands (in contrast, the Berry curvature vanishes identically for bilayer and bulk MoS$_2$, both of which are centrosymmetric). The Berry curvature is defined in terms of the cell-periodic Bloch states as

$$\Omega_{n,xy}(\mathbf{k}) = -2\,\mathrm{Im}\sum_n \langle \partial_{k_x} u_{n\mathbf{k}} | \partial_{k_y} u_{n\mathbf{k}} \rangle, \tag{1}$$

and it modifies the current response to an applied electric field by adding an "anomalous velocity" term to the semiclassical equations of motion [8]. A well-known consequence is the anomalous Hall effect (AHE) in magnetic materials, where the

Berry curvature is induced by broken time-reversal symmetry. The intrinsic part of the anomalous Hall conductivity (AHC), denoted below by the superscript "0", is given by the Brillouin zone (BZ) integral of the Berry curvature summed over the occupied states [8,9].

$$\Omega_{xy}(\mathbf{k}) = \sum_n f_{n\mathbf{k}} \Omega_{n,xy}(\mathbf{k}), \tag{2}$$

$$\sigma_{xy}^0 = -\frac{e^2}{h}\int_{\mathrm{BZ}} \frac{d^d k}{(2\pi)^{d-1}} \Omega_{xy}(\mathbf{k}), \tag{3}$$

where $f_{n\mathbf{k}}$ is the occupation factor and d is the dimensionality. For $d = 2$ the AHC has units of conductance (S), and for $d = 3$ it has units of conductivity (S/cm).

Monolayer MoS$_2$ is nonmagnetic, and the presence of time-reversal symmetry implies the relation [8]

$$\Omega_{xy}(-\mathbf{k}) = -\Omega_{xy}(\mathbf{k}). \tag{4}$$

Thus equal and opposite amounts of Berry curvature accumulate in the two valleys, resulting in a cancellation of the valley Hall currents and a vanishing AHC. Time-reversal symmetry can however be broken by illuminating the sample with circularly polarized light, leading to a *photoinduced* AHE. The valley-selective photoexcitation creates a carrier imbalance which in turn removes the exact cancellation between the Hall currents in the two valleys. This so-called *valley Hall effect* was first discussed for graphene systems with broken inversion symmetry [3], and later for monolayer MoS$_2$ [4]. The effect was subsequently measured by Mak *at al.* in transistors of MoS$_2$ monolayers [10].

Compared to the conventional AHE in ferromagnetic metals, the theoretical modeling of the photoinduced AHE in TMDs poses the additional challenge that the AHC should in principle be calculated for a *nonequilibrium* photoexcited state, but to our knowledge such a calculation has not yet been attempted. Instead, an approximate but more tractable

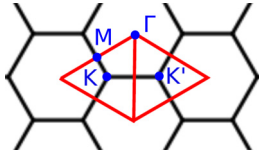*tolsen@fysik.dtu.dk
†ivo_souza@ehu.es

FIG. 1. (Color online) Brillouin zone of monolayer MoS$_2$, partitioned into two triangular valleys delimited by lines of vanishing Berry curvature, drawn in red. The two valleys are centered at the high-symmetry points K and K'.

approach is often used [3,4,10]. The idea is to introduce an auxiliary quantity σ_{xy}^V, which we will call the *valley Hall conductivity* (VHC), defined as the integral of the Berry curvature over a single valley domain in the BZ. For example, the intrinsic VHC of the valley centered at K in Fig. 1 is

$$\sigma_{xy}^{0,K}(\varepsilon_F) = -\frac{e^2}{h} \int_K \frac{d^2k}{2\pi} \Omega_{xy}(\mathbf{k}), \tag{5}$$

and similarly for the valley centered at K'. [The demarcation of the two valley domains will be discussed further in Sec. II A. Note also that Eq. (5) depends on the Fermi level ε_F through the occupation factors in Eq. (2).] The photoinduced AHC $\delta\sigma_{xy}$ is then approximated by the sum of the VHCs in the two valleys, positing a Fermi-level shift $\delta\varepsilon$ between them to mimic the effect of the valley-selective photoexcitation,

$$\delta\sigma_{xy}(\varepsilon_F, \delta\varepsilon) = \sigma_{xy}^K(\varepsilon_F + \delta\varepsilon) + \sigma_{xy}^{K'}(\varepsilon_F). \tag{6}$$

When $\delta\varepsilon = 0$ the AHC vanishes, and a nonzero $\delta\sigma_{xy}$ appears when $\delta\varepsilon \neq 0$. This approach also allows for a direct comparison with model calculations, without having to consider the details of how the carrier imbalance between the valleys is generated.

This will be the basic approach taken in the present work. We have suppressed the superscript 0 from this last equation to emphasize that it remains valid when the nonintrinsic contributions which we will now discuss are taken into account.

Impurities are always present in real samples, and their *extrinsic* contributions to the photoinduced AHE in TMDs should be taken into account alongside the intrinsic response described by Eqs. (5) and (6). This is well known in the context of the AHE in ferromagnetic metals, where historically two types of extrinsic contributions have been considered—*side jump* and *skew scattering* [9]. In a simplified picture, the sidejump effect originates in the anomalous velocity that a wave packet may acquire as it moves through an impurity potential, while skew scattering arises from the chiral part of a standard transition-rate expression. With some effort, both contributions can be incorporated into the semiclassical Boltzmann-transport framework [11].

The correspondence between the semiclassical treatment of the AHC and a fully quantum-mechanical (Kubo-Streda) calculation based on a perturbative expansion in powers of the disorder strength was carefully worked out in Ref. [12]. It became clear from that analysis that not all terms fall distinctly into either of the above physical interpretations of extrinsic contributions to the AHC, and for many purposes it is more practical to base the distinction on the scaling with impurity

concentration [9]. According to this viewpoint skew scattering is defined as the part of the AHC that scales inversely with the impurity concentration, while the part that is independent of the impurity concentration has both intrinsic and side-jump components. Although the intrinsic contribution is sharply defined theoretically in terms of the electronic structure of the pristine crystal by Eq. (3), experimentally it is not known how to separate it from the side-jump part. Note that the anomalous Hall response of pristine samples at low temperatures is dominated by skew scattering, with the intrinsic contribution only becoming significant in moderately resistive samples (where it competes with side-jump scattering). This analysis, originally developed for the AHC in ferromagnetic metals, carries over to the VHC and photoinduced AHC in TMDs.

It is well established that sulfur vacancies constitute the main source of disorder in MoS$_2$ [13–18]. The formation energies and thermodynamics of these defects have been thoroughly studied [19], but their influence on transport and optical properties remains largely unexplored. Modeling the effects of disorder from first principles is a challenging task in general, but there are noteworthy examples where the AHC in ferromagnetic materials has been calculated using the coherent potential approximation [20–25]; also, an *ab initio* implementation of the side-jump contribution to the AHC has been carried out assuming scattering centers with δ-function potentials [26].

In this work, we develop a computational scheme that allows us to include in a realistic manner the effect of impurities in the calculation of the VHC and of the photoinduced AHE in TMDs. In a first step, we perform several supercell calculations at the desired impurity concentration, corresponding to different realizations of disorder. In order to carry out the calculations efficiently while maintaining first-principles-like accuracy, we construct effective Hamiltonians in a Wannier-function basis, starting from density functional theory calculations on smaller cells [27]. Recall that the definition of the VHC in Eq. (5) requires identifying the individual valley domains in the BZ where the Berry curvature is to be integrated. It is not clear *a priori* how to do so in the context of a supercell calculation, since the electronic states cannot be labeled by wave vectors in the normal BZ of Fig. 1. In order to overcome this difficulty, in a second step we use a "BZ unfolding" technique [28] to map the results of the supercell calculation onto the normal BZ of the pristine crystal. More precisely, we express the AHC of each disordered supercell configuration as an integral of the *supercell* Berry curvature over the folded BZ, and then unfold the Berry curvature onto the normal BZ according to the prescription of Ref. [29]. Having done that, the VHC (including the contributions from disorder) can then be obtained by integrating the *unfolded* Berry curvature over a single valley domain in Fig. 1, and averaging the result over several realizations of disorder.

We have used the above first-principles-based methodology to study the influence of sulfur vacancies on the VHC of MoS$_2$, as well as the photoinduced AHC, which is the quantity measured in experiments. The calculated VHC as a function of defect concentration was compared with model calculations where the valence and conduction-band edges in each valley are described by a massive Dirac Hamiltonian with a random distribution of δ-function scatterers [12].

The paper is organized as follows. We start Sec. II by reviewing some basic features of the electronic structure of MoS$_2$. We then evaluate the intrinsic VHC [Eq. (5)] and photoinduced AHC [Eq. (6)] for the massive Dirac Hamiltonian without disorder, and carry out the corresponding *ab initio* calculations for pristine MoS$_2$. Our main results are presented in Sec. III, where disorder effects are included in the calculation of the VHC, both for the massive Dirac Hamiltonian and for MoS$_2$ with sulfur vacancies. The two types of calculations are found to be in reasonable agreement, and we then proceed to calculate the photoinduced AHC for the disordered massive Dirac model as a function of gating voltage and Fermi-level shift, finding good agreement with the experimental measurements. Our conclusions are summarized in Sec. IV, and the Appendixes present the details of the *ab initio* calculations, the BZ unfolding method, and the effective-Hamiltonian methodology.

II. PRISTINE MoS$_2$

A. Energy bands and Berry curvature

Ab initio density-functional theory calculations were carried out for monolayer MoS$_2$ as described in Appendix A 1. The calculated Kohn-Sham energy bands are shown in the upper panel of Fig. 2, color coded by the spin component $\langle S_z \rangle$ orthogonal to the layer. The minimum direct gap is situated at K and $K' = -K$, with a value of $\sim$1.7 eV. Away from the time-reversal-invariant points Γ and M the spin degeneracy is split by the combination of broken inversion symmetry and spin-orbit coupling (the degeneracy is actually protected along the entire Γ-M line by mirror symmetry). The two topmost valence bands exhibit a maximum spin-orbit splitting of $\sim$0.15 eV at K and K', where S_z is a good quantum number and Kramers-degenerate partners have opposite spin character: $\varepsilon_{K\uparrow} = \varepsilon_{K'\downarrow}$.

The lower panel of Fig. 2 shows the Berry curvature summed over the valence bands, Eq. (2). In agreement with Eq. (4), Ω_{xy} is an odd function of $\mathbf{k}$. Its magnitude peaks at
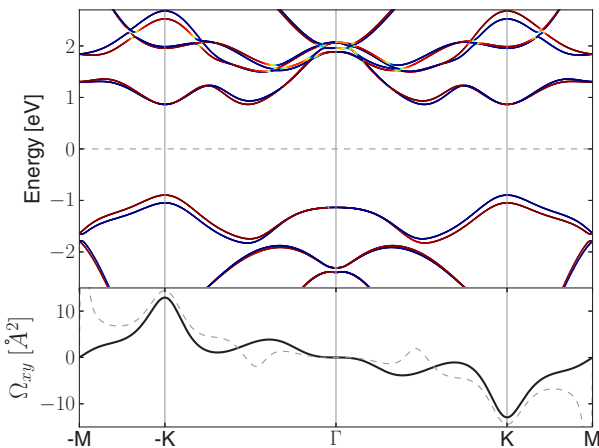

FIG. 2. (Color online) Top: calculated band structure of monolayer MoS$_2$. Energies are measured from the Fermi level, and the bands are color coded by the spin expectation value $\langle S_z \rangle$, with red corresponding to spin up and blue to spin down. Bottom: Berry curvature summed over the valence bands [Eq. (2)]. The dashed line is the Berry curvature evaluated from the two topmost valence bands.

the minimum-gap points K and K', where the sign is dictated by the dominant contributions coming from the topmost valence bands. At K those bands are mainly composed of molybdenum d states with $m_l = 2$; according to the optical selection rules [7] those states can be excited using left-handed polarized light. The Berry curvature has a secondary peak between Γ and K; there, a pair of lower-lying valence bands approaches the topmost ones, and also contributes significantly to the Berry curvature.

It should be noted that because the Berry curvature is induced by the broken spatial inversion, it is not directly related to the spin-orbit splitting evident in Fig. 2; in fact, Ω_{xy} is practically unaltered if the spin-orbit interaction is switched off. (This is in sharp contrast to the Berry curvature induced by broken time-reversal symmetry in ferromagnetic metals, which vanishes in the absence of spin-orbit coupling [8,9].) Likewise, the extrinsic scattering contributions do not rely on spin-orbit coupling. Thus one can use a spinless model such as the massive Dirac Hamiltonian of Sec. II B 1 to describe the band edges and valley Berry curvature in MoS$_2$; spin is then accounted for by inserting a factor of two in the calculated Ω_{xy}.

The Hamiltonian of monolayer MoS$_2$ is invariant under reflection across the vertical planes containing the lines that connect a Mo atom to the neighboring S atoms. The corresponding symmetry elements in reciprocal space are the Γ-M mirror lines. The Berry curvature transforms like a magnetic field in reciprocal space [8]. In particular, the component $\Omega_{xy} = \Omega_z$ is odd under reflection across the Γ-M lines, and hence it vanishes along those lines, which form the boundaries between the two valleys in Fig. 1. This allows us to uniquely define the intrinsic VHC according to Eq. (5).

B. Intrinsic valley Hall conductivity

1. Massive Dirac model

The valence and conduction-band edges of a single valley of MoS$_2$ and related materials are often modeled by the massive Dirac Hamiltonian [1,3,4]. The Hamiltonian for the K valley in Fig. 2 reads

$$H(\mathbf{k}) = \hbar v(-k_x \sigma_x + k_y \sigma_y) + \Delta \sigma_z, \qquad (7)$$

where $\mathbf{k} = (k_x, k_y)$ is measured from the valley center K, σ_i are the Pauli matrices, and Δ is the mass parameter. The energy eigenvalues and Berry curvature are

$$\varepsilon_\pm(\mathbf{k}) = \pm\sqrt{\Delta^2 + v^2 \hbar^2 k^2}, \qquad (8)$$

$$\Omega_{\pm,xy}(\mathbf{k}) = \pm \frac{\Delta v^2 \hbar^2}{2(\Delta^2 + v^2 \hbar^2 k^2)^{3/2}}, \qquad (9)$$

where $k^2 = k_x^2 + k_y^2$. In the case of the K' valley Eq. (8) remains unchanged, while Eq. (9) flips sign.

Experimental realizations of the valley Hall effect usually involves a gate, whose effect we model as a shift in the Fermi level. When ε_F lies in the conduction band ($\varepsilon_F > \Delta$) the VHC becomes

$$\sigma_{xy}^{0,K} = -\frac{2e^2}{h} \left[\int \frac{d^2 k}{2\pi} \Omega_-(k) + \int_{|k| < k_F} \frac{d^2 k}{2\pi} \Omega_+(k) \right]$$

$$= \frac{e^2}{2h} - \frac{e^2 \Delta}{2h} \left(\frac{1}{\Delta} - \frac{1}{\varepsilon_F} \right). \qquad (10)$$

The valley and spin degrees of freedom can be included by coupling four copies of this model. Referring to the *ab initio* band structure of MoS$_2$ we find that the massive Dirac model provides a good fit around the valleys if we take $\Delta = 0.86$ eV, corresponding to half of the gap at K and K' in Fig. 2. The velocity parameter v is calculated as $v^2 = \Delta/m_e^*$, where $m_e^* = 0.4m_e$ is the effective mass of the conduction-band valley.

For the ungated case where ε_F lies in the gap, the model gives an intrinsic VHC of exactly $e^2/2h$. The deviation from this result in the *ab initio* calculation of Fig. 2 measures the contributions from lower-lying valence bands and from the crystal potential, which gives rise to nonhyperbolic bands away from the top of the K and K' valleys. [It should be noted, though, that a Chern insulating system with an inversion center will retain the exact value $\sigma_{xy}^{0,K} = e^2/2h$ when the crystal potential is included, because symmetry implies $\Omega(\mathbf{k}) = \Omega(-\mathbf{k})$ and the topology implies $\sigma_{xy}^0 = 2\sigma_{xy}^{0,K} = e^2/h$ [30,31].] From the *ab initio* calculation of the Berry curvature in MoS$_2$ we find, using Eq. (5), $\sigma_{xy}^{0,K} = 0.71e^2/h$; since both spin channels contribute equal amounts of Berry curvature, this corresponds to 71% of the result obtained for the massive Dirac model.

Measuring the valley Hall effect requires generating a carrier imbalance between the two valleys, which we describe as a relative shift $\delta\varepsilon$ between the Fermi levels in the two valleys. According to Eq. (6), a positive value of $\delta\varepsilon$ means that the Fermi level is higher at K than at K', a situation that can be produced by illuminating the sample with left-handed polarized light propagating along $+\hat{\mathbf{z}}$ (see Sec. II A). Thus the sign of $\delta\varepsilon$ equals the helicity of the incident light, and the magnitude depends on the intensity. Combining Eqs. (6) and (10) we obtain, for $\varepsilon_F > \Delta$,

$$\delta\sigma_{xy}^0(\varepsilon_F, \delta\varepsilon) = -\frac{e^2\Delta}{2h}\left(\frac{1}{\varepsilon_F} - \frac{1}{\varepsilon_F + \delta\varepsilon}\right). \qquad (11)$$

We can also express this quantity in terms of carrier imbalance δn_c between the conduction-band edges in the two valleys. For a single valley

$$n_c(\varepsilon_F) = \frac{\pi k_F^2}{(2\pi)^2} = \frac{\varepsilon_F^2 - \Delta^2}{4\pi v^2\hbar^2}, \qquad (12)$$

so that

$$\delta n_c = \frac{2\varepsilon_F\delta\varepsilon + \delta\varepsilon^2}{4\pi v^2\hbar^2}. \qquad (13)$$

For $\delta\varepsilon \ll \varepsilon_F$ we then get

$$\delta\sigma_{xy}^0 \approx -\frac{e^2}{h}\frac{\Delta}{\varepsilon_F^2}\delta\varepsilon \approx -\frac{2\pi e^2 v^2\hbar\Delta}{\varepsilon_F^3}\delta n_c. \qquad (14)$$

In this limit the photoinduced AHC becomes linear in both the small energy shift $\delta\varepsilon$ and in the small carrier imbalance δn_c. The case where ε_F lies in the gap ($|\varepsilon_F| < \Delta$) was considered in Ref. [10], and the corresponding expression can be obtained by setting $\varepsilon_F = \Delta$ in Eq. (14). As will be shown in Sec. III E, the position of the Fermi level has a minute effect on the intrinsic and side-jump contributions, but it can have a large effect on the skew-scattering contribution, which becomes dominant for clean samples.
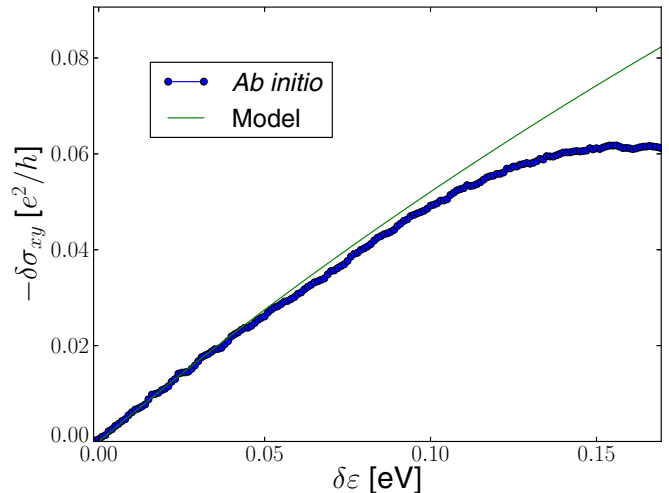

FIG. 3. (Color online) Intrinsic photoinduced anomalous Hall conductivity of a single layer of MoS$_2$ without gating ($\varepsilon_F = 0$), plotted as a function of the Fermi-level shift $\delta\varepsilon$ between the K and K' valleys. The thin green line corresponds to Eq. (14) with $\varepsilon_f = \Delta$ with the parameters given in the text, and the blue dotted line was obtained by evaluating Eqs. (5) and (6) from first principles.

2. First-principles calculations

In Fig. 3 we show the intrinsic photoinduced AHC $\delta\sigma_{xy}^0$ calculated for a single spin channel as a function of the single-valley energy shift $\delta\varepsilon$, for $\varepsilon_F = 0$. It should be noted that due to the spin-orbit split valence bands in MoS$_2$, it is possible to selectively excite a single spin channel by using an optical frequency tuned to the transition between the topmost valence band and the conduction band [4,10]. The *ab initio* calculations are seen to agree very well with the model results at the bottom of the valley. When the energy shift approaches $\sim$0.1 eV the secondary valley located between Γ and K starts to contribute, and the model results become unreliable.

The measurements of the valley Hall effect reported in Ref. [10] involved photoexcitation of states in a single valley. In that work, the carrier density was estimated from photoconductivity measurements, and the photoinduced AHC was displayed as a function of carrier density. The typical density of photoexcited carriers reported in Ref. [10] is of the order of 10^{-11} cm^{-2}, which corresponds to an energy shift $\delta\varepsilon \sim 1$ meV; this is far below the point where the linear model (14) breaks down.

III. DISORDERED MoS$_2$

Recent experimental as well as theoretical studies have demonstrated that sulfur vacancies are the dominant source of disorder in MoS$_2$ [13–18]. In the following we will therefore focus exclusively on sulfur vacancies, and calculate their effect on the VHC. Before we delve into the *ab initio* calculations, let us briefly review the theoretical results for the AHC in a massive Dirac model with impurity scattering included, which we shall use as a simple model for the VHC of disordered MoS$_2$.

A. Massive Dirac model

The extrinsic contributions to the AHC of the massive Dirac model of Sec. II B 1 have been calculated in Ref. [12] in the limit of weak and dilute scattering. The scattering centers were assumed to be randomly distributed, and the impurity potential for a given configuration was chosen as $V(\mathbf{r}) = \sum_i V_i \delta(\mathbf{r} - \mathbf{R_i})$, where $\mathbf{R_i}$ are impurity sites. The intrinsic (0), side-jump (SJ), and skew-scattering (SS) contributions to the VHC of the K valley are given by [12]

$$\sigma_{xy}^{0,K} = \frac{e^2 \Delta}{2h \sqrt{(v \hbar k_F)^2 + \Delta^2}}, \tag{15}$$

$$\sigma_{xy}^{\mathrm{SJ},K} = \frac{e^2 \Delta}{2h \sqrt{(v \hbar k_F)^2 + \Delta^2}}$$
$$\times \left[\frac{4(v \hbar k_F)^2}{4 \Delta^2 + (v \hbar k_F)^2} + \frac{3(v \hbar k_F)^4}{(4 \Delta^2 + (v \hbar k_F)^2)^2} \right], \tag{16}$$

$$\sigma_{xy}^{\mathrm{SS},K} = \frac{e^2 \langle V_i^3 \rangle_c \Delta}{h x \langle V_i^2 \rangle_c^2} \frac{(v \hbar k_F)^4}{(4 \Delta^2 + (v \hbar k_F)^2)^2}. \tag{17}$$

Equation (15) is just Eq. (10) recast in a different form, and in Eq. (17) $\langle V_i^n \rangle_c$ denotes the nth cumulant moment of the distribution of impurity strengths. As mentioned in the Introduction, scattering contributions which are independent of the impurity concentration x are classified as side jump, and those which scale inversely with x are classified as skew scattering. Note that the second term in Eq. (16) originates from a fourth-order expansion of the scattering rate and could, from a different point of view, be regarded as a skew-scattering contribution. Likewise, the side-jump contribution as defined here contains contributions that cannot directly by ascribed to a coordinate shift [12].

In Fig. 4 we show the VHC calculated from Eqs. (15)–(17) as a function of the impurity concentration x, for different
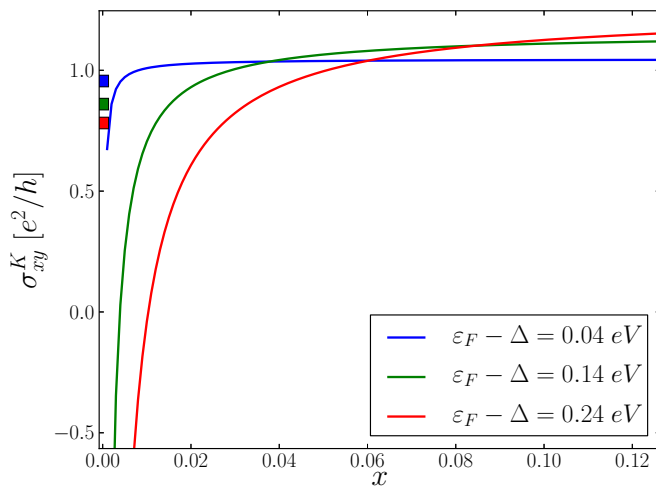

FIG. 4. (Color online) Valley Hall conductivity of the disordered massive Dirac model, Eqs. (15)–(17), evaluated as a function of impurity concentration x for three different gating levels, with $\langle V_i^3 \rangle_c / \langle V_i^2 \rangle_c^2 = -0.4$ eV and $\Delta = 0.86$ eV. The squares at $x = 0$ correspond to the intrinsic contribution of Eq. (15).

values of $\epsilon_F - \Delta$ (the Fermi level relative to the bottom of the conduction band). We have set the remaining model parameters to $\Delta = 0.86$ eV (same as in Sec. II B 1) and $\langle V_i^3 \rangle_c / \langle V_i^2 \rangle_c^2 = -0.4$ eV^{-1}. The latter value was chosen to fit the *ab initio* results for the VHC of MoS$_2$, to be presented in Sec. III D. The VHC of the disordered massive Dirac model converges to the side-jump value at large impurity concentrations, but one should keep in mind that Eqs. (16) and (17) were derived under the assumption of dilute disorder and weak scattering. The side-jump correction to the intrinsic part becomes larger for higher density of states at ε_F, and becomes negligible at very low carrier concentrations. Furthermore, at low carrier concentrations skew-scattering only becomes significant at very low impurity concentrations. It thus appears that for low carrier concentrations the intrinsic contribution gives a good account of the VHC—even at rather low impurity concentrations.

B. Unfolded band structure and Berry curvature

In pristine systems, the Berry curvature provides a useful k-resolved measure of the AHC and VHC. For example, from Fig. 2 it is clear that states near the valleys at K and K' have the largest potential for contributing to the VHC. However, if we would like to know how a given distribution of impurities affects the VHC, this picture immediately breaks down since the pristine Brillouin zone is no longer relevant once translational symmetry is broken. On the other hand, if a given impurity distribution is represented in a supercell, the VHC will still be given as a k-space integral of the Berry curvature, but now the domain will be the Brillouin zone corresponding to the supercell (SBZ), which is not directly comparable to the normal Brillouin zone (NBZ). Nevertheless, we can expand the supercell states in terms of states in the pristine system and thus unfold the supercell curvature to the pristine Brillouin zone. For a general band quantity $a_N(\mathbf{K})$ defined in SBZ we can thus define the unfolded quantity $a_n^{(u)}(\mathbf{k})$ in the NBZ as

$$a_n^{(u)}(\mathbf{k}) = \sum_N |\langle N\mathbf{K}|n\mathbf{k} \rangle|^2 a_N(\mathbf{K}), \tag{18}$$

where $\mathbf{K}$ is the SBZ crystal momentum, which is related to k by translation of a supercell reciprocal lattice vector. Brillouin zone integrals can be written in terms of the unfolded function since

$$A = \int_{\mathrm{SBZ}} d\mathbf{K} \sum_N a_N(\mathbf{K})$$
$$= \int_{\mathrm{SBZ}} d\mathbf{K} \int_{\mathrm{NBZ}} d\mathbf{k} \sum_{nN} |\langle N\mathbf{K}|n\mathbf{k} \rangle|^2 a_N(\mathbf{K})$$
$$= \int_{\mathrm{NBZ}} d\mathbf{k} \sum_n a_n^{(u)}(\mathbf{k}), \tag{19}$$

where a complete set of states in the NBZ was inserted in the second line, and we invoked the fact that $\langle N\mathbf{K}|n\mathbf{k} \rangle$ is nonvanishing only if $\mathbf{k}$ downfolds onto $\mathbf{K}$.

The above method for unfolding k-space quantities has previously been applied to band structures of disordered systems [27,28] and more recently to the Berry curvature [29].
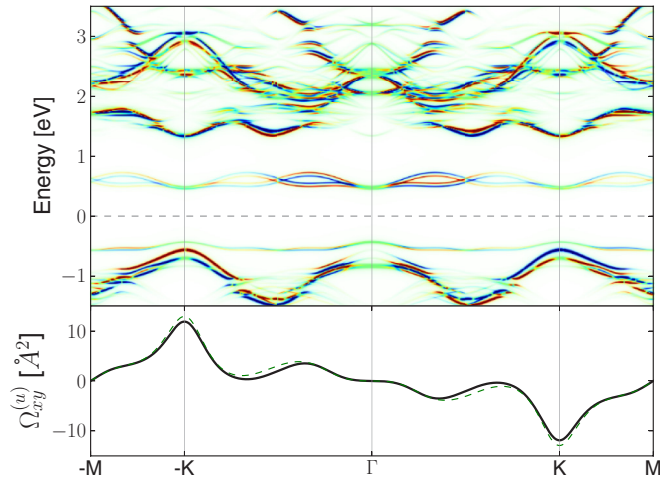
FIG. 5. (Color online) Unfolded spectral function (top) and Berry curvature (bottom) of a 3×3 supercell of MoS_2 with a single sulfur vacancy. The Fermi level $\varepsilon_F = 0$ is in the gap, and the Berry curvature is largely unaffected by the presence of the impurity (the Berry curvature for pristine MoS_2 ,taken from Fig. 2, is drawn as a dashed green line). The bands in the top panel are color coded by the unfolded spin expectation value $\langle S_z \rangle$ with red corresponding to spin up and blue to spin down.

In the case of band structures the object of interest is the spectral function

$$A^{(u)}(\omega,\mathbf{k}) = \sum_{nN} |\langle N\mathbf{K}|n\mathbf{k}\rangle|^2 \delta(\omega - \varepsilon_{N\mathbf{K}}). \qquad (20)$$

The treatement of the Berry curvature is more subtle, since a naive application of Eq. (18) would lead to

$$\Omega^{(u)}(\mathbf{k}) = \sum_{nN} |\langle N\mathbf{K}|n\mathbf{k}\rangle|^2 f_{N\mathbf{K}} \Omega_N(\mathbf{K}), \qquad (21)$$

which is gauge dependent. In Ref. [29] a gauge-invariant expression for $\Omega^{(u)}(\mathbf{k})$ was obtained which is different from Eq. (21), and in the present work we have used the gauge-invariant expression in all calculations.

C. Sampling impurity configurations

In Fig. 5 we show the unfolded spectral function and the unfolded Berry curvature of a periodic structure obtained as 3×3 supercell of MoS_2 with a single sulfur vacancy. The vacancy is seen to introduce both occupied and unoccupied states in the gap, but the Berry curvature is barely affected, since those states are rather localized. This system is an example of a $x = 1/18 = 0.56$ impurity concentration, but it is not necessarily representative of a typical disordered configuration at this concentration. As it turns out, the Berry curvature is largely insensitive to the impurity configuration as long as ε_F lies in the gap. However, the situation changes dramatically when ε_F is shifted to the conduction band. This situation is shown in Fig. 6. The blurred features in the unfolded spectral function are associated with scattering states, and they are accompanied by spiky features in the unfolded Berry curvature. From the spectral representation of the Berry
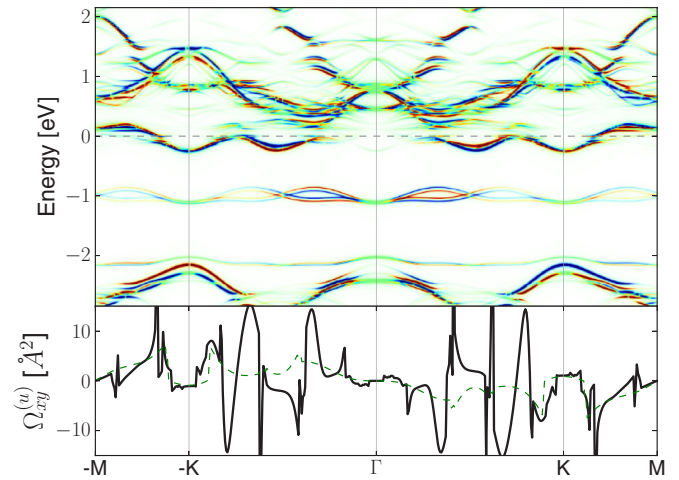

FIG. 6. (Color online) Same as Fig. 5, except that the Fermi level has been shifted to the conduction band, leading to large impurity-induced changes in the Berry curvature.

curvature in Eq. (A1) it is clear that such features arise whenever occupied and unoccupied states come close to ε_F.

As already mentioned, in experiments with MoS_2 transistors ε_F is typically controlled with a gate voltage [10]. Furthermore, exfoliated MoS_2 usually exhibits an intrinsic n doping, which has been attributed to Re impurities [19]. For these reasons, in the calculations reported below we have chosen to pin ε_F to the conduction band. This will facilitate the comparison with both experiments and model calculations.

A reliable calculation of the conductivity at a given impurity concentration requires averaging over a large number of configurations, each calculated using a large simulation cell, which is not feasible using standard *ab initio* methods. To proceed we construct an effective Hamiltonian based on *ab initio* DFT calculations and Wannier functions [27]. For each concentration of sulfur vacancies we randomly generate a large number ($\sim$1000) of disordered configurations in a 6×12 supercell, and for each of them we construct the effective Hamiltonian as described in Appendix B. For the lowest impurity concentration considered, $x = 1/216$, a larger 12×18 supercell was needed.

As a first check of the method, we calculate the optical conductivity of MoS_2 at various concentrations of sulfur vacancies. The results are shown in Fig. 7, and as expected the main effect of disorder is to broaden the spectrum. Note that since the Wannier functions used to construct the tight-binding Hamiltonian were obtained by disentangling bands up to 3.0 eV above the conduction-band minimum ε_c (see Appendix A 1), the calculated spectrum is reliable only up to 3.0 eV above the pristine absorption edge.

D. Valley Hall conductivity

Now we turn to the VHC in disordered samples, calculated by replacing the Berry curvature in Eq. (5) with the disorder-averaged unfolded Berry curvature $\langle \Omega_{xy}^{(u)}(\mathbf{k})\rangle_{\mathrm{dis}}$. The lowest impurity concentration we have considered is $x = 0.005$, for which $\langle \Omega_{xy}^{(u)}(\mathbf{k})\rangle_{\mathrm{dis}}$ becomes very spiky and is not particularly informative. In Fig. 8 we plot $\langle \Omega_{xy}^{(u)}(\mathbf{k})\rangle_{\mathrm{dis}}$ together with the
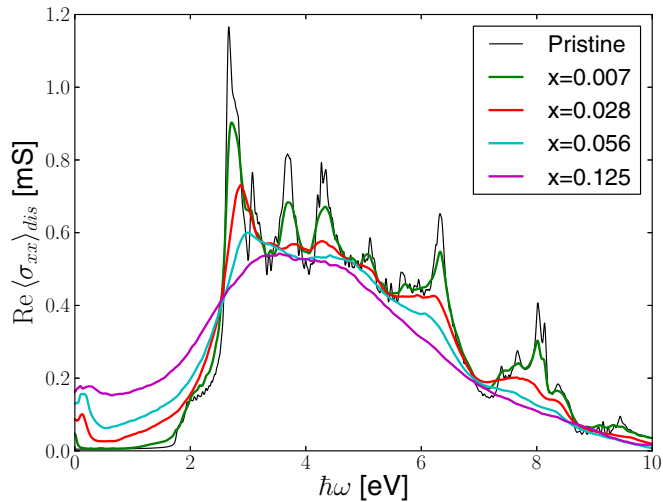
FIG. 7. (Color online) Configuration-averaged optical absorption spectrum of MoS_2 at various impurity concentration.

spectral function $\langle A^{(u)}(\mathbf{k})\rangle_{dis}$ at four different intermediate impurity concentrations, with $\varepsilon_F - \Delta = 0.24$ eV. In general we observe that the peaks in the Berry curvature become enhanced and broadened by disorder, while retaining the qualitative features of the pristine Berry curvature. This tendency is maintained over a range of impurity concentrations (from $x = 0.02$ to $x = 0.1$) and gives rise to an impurity-

independent increase in the VHC, as will be discussed shortly. We will identify this with the side-jump corrected VHC, and so the difference between $\langle \Omega_{xy}^{(u)}(\mathbf{k})\rangle_{dis}$ and the pristine Berry curvature in Fig. 2 provides a measure of the k-space resolved side-jump scattering.

From the unfolded spectral functions in Fig. 8, it is clear that sulfur vacancies have the effect of lowering the overall potential, such that at high impurity concentrations the conduction bands are lowered with respect to the fixed Fermi level. However, for intermediate doping levels the exact position of ε_F does not have a large effect on the VHC. In fact, raising ε_F with respect to the conduction band tends to lower the VHC, since more Berry curvature from the conduction band will be included, which has the opposite sign compared to the dominant contribution from the valence bands. At larger impurity concentrations ($x = 0.125$) the Berry curvature becomes more smeared out, and it will eventually average to zero. At this point the system is so strongly perturbed that it cannot be analyzed in terms of scattering events.

In Fig. 9 we show the VHC as a function of impurity concentration at different Fermi-level shifts. At low impurity concentrations the VHC displays a divergent behavior, which on the basis of Eqs. (15) and (16) we attribute to skew-scattering processes. In this regime $\Omega_{xy}^{(u)}(\mathbf{k})$ becomes rather spiky, and its configurational average bears little resemblance to the pristine Berry curvature. We note that as the impurity concentration x decreases, it becomes progressively harder to converge the calculation. This is due to the fact that
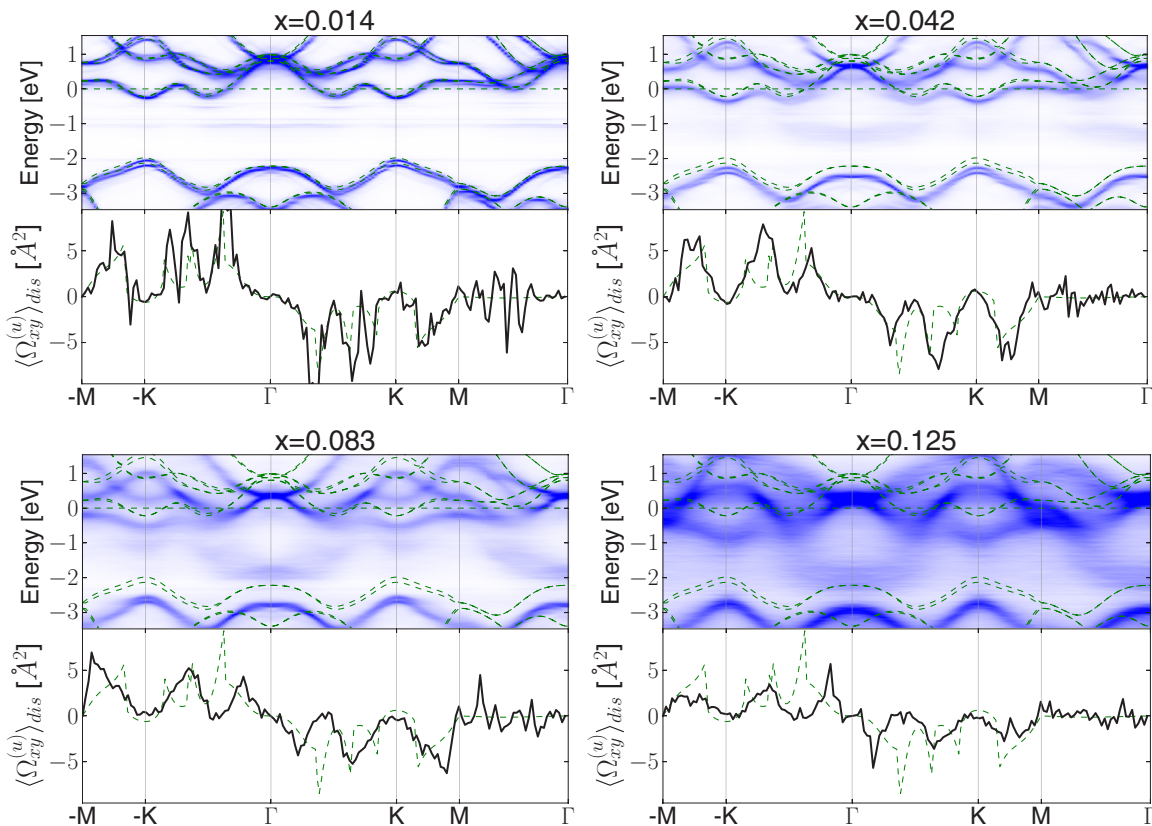

FIG. 8. (Color online) Configuration-averaged unfolded spectral function (top panels) and unfolded Berry curvature (bottom panels) of disordered MoS_2, for several impurity concentrations. Energies are measured from the Fermi level, which is fixed in the conduction band. The energy bands and Berry curvature of pristine MoS_2 are shown as dashed green lines for comparison.
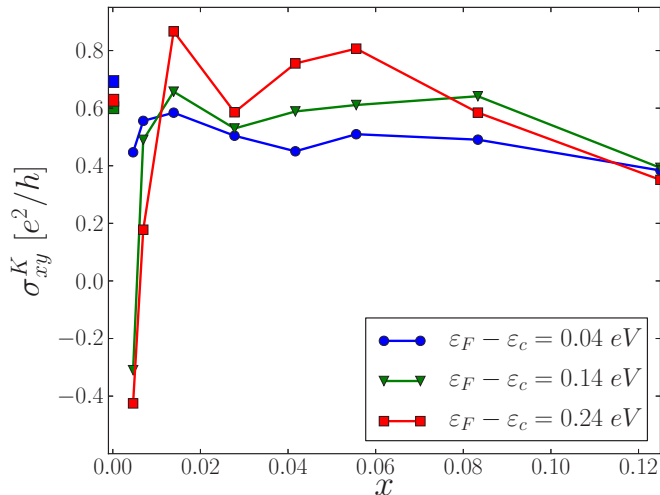
FIG. 9. (Color online) Calculated VHC as a function of impurity concentration x. We show the result for three different positions of the Fermi level ε_F with respect to the conduction-band minimum ε_c.

the standard deviation for a given supercell size increases when x becomes small, while at the same time very large supercells are needed in order to reach low impurity concentrations.

At intermediate impurity concentrations the VHC reaches a plateau, which we identify as the side-jump regime. In this regime skew scattering is insignificant, and the intrinsic VHC receives a small correction which is nearly independent of the impurity concentration. Comparing with Fig. 4, we see that there is good qualitative agreement between the model and the *ab initio* results. One major difference is that while the *ab initio* VHC decreases at high impurity concentrations, the model VHC converges towards the side-jump result. The reason is that, as emphasized in Sec. III A, Eqs. (16) and (17) were derived under the assumption of low impurity concentration, and cannot be applied in this regime. Moreover, at large Fermi-level shifts the band structure of MoS_2 starts to deviate appreciably from the bands of the Dirac model. The most important effect is the contribution from the secondary conduction-band minimum between Γ and K in Fig. 2.

Two crucial features can be observed in Figs. 4 and 9 when ε_F comes close to ε_c. First, the side-jump correction becomes small such that the intrinsic VHC becomes dominant at intermediate impurity concentrations. Second, the critical impurity concentration where skew scattering starts to dominate becomes rather small. For $\varepsilon_F - \varepsilon_c = 0.04$ eV the critical concentration is $x \sim 0.005$, and it will become even smaller as ε_F moves closer to ε_c.

In closing, we note that gating may qualitatively change the physical properties of TMDs [32,33]. Such effects are however beyond the scope of the present work, where we have assumed that the main effect of gating is to shift ε_F relative to ε_c.

E. Photoinduced anomalous Hall conductivity

The experimentally measured quantity is the photoinduced AHC $\delta\sigma_{xy}$, which we express as the sum of the two VHCs
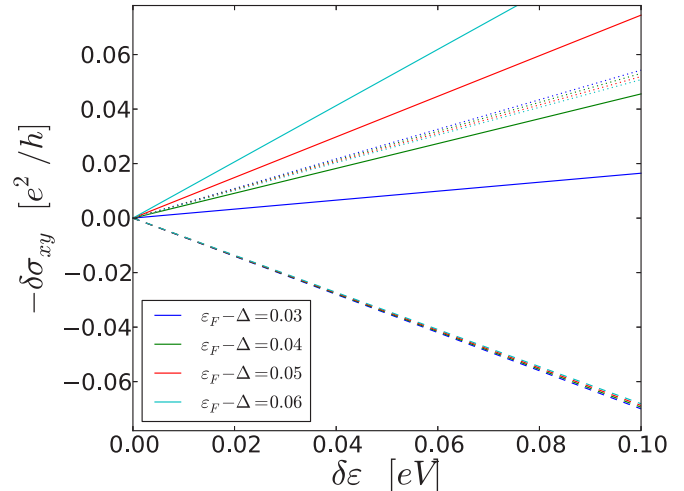

FIG. 10. (Color online) Photoinduced AHC as a function of the Fermi-level shift $\delta\varepsilon$ between the K and K' valleys. States in the K valley are preferentially excited ($\delta\varepsilon > 0$), corresponding to left-handed polarized light. The dotted lines denote the intrinsic contribution $\delta\sigma_{xy}^0$, the dashed lines are $\delta\sigma_{xy}^0 + \delta\sigma_{xy}^{SJ}$, and the solid lines are $\delta\sigma_{xy}^0 + \delta\sigma_{xy}^{SJ} + \delta\sigma_{xy}^{SS}$. Note the large dependence on ε_F when skew scattering is included.

according to Eq. (6). Assuming a small carrier imbalance and using $\sigma_{xy}^{K'}(\varepsilon) = -\sigma_{xy}^K(\varepsilon)$ we find

$$\delta\sigma_{xy}(\varepsilon_F, \delta\varepsilon) \simeq \frac{d\sigma_{xy}^K(\varepsilon)}{d\varepsilon}\bigg|_{\varepsilon=\varepsilon_F} \delta\varepsilon. \qquad (22)$$

A first-principles evaluation of Eq. (22) including the effect of impurities is made difficult by the fluctuations involved in the configuration-averaging procedure. However, the good agreement between the *ab initio* calculations of the VHC (Fig. 9) and the calculations for the massive Dirac model (Fig. 4) suggests that a reliable estimate of the photoinduced AHC in MoS_2 can be obtained by evaluating Eq. (22) for the Dirac model, with $\sigma_{xy}^K(\varepsilon)$ given by the sum of Eqs. (15)–(17). For the skew-scattering contribution we take $\langle V_i^3 \rangle_c / \langle V_i^2 \rangle_c^2 = -0.4$ eV as in Fig. 4. Note that this value (and in particular the sign) was chosen in order to obtain maximum agreement between Figs. 4 and 9, and thus represents a nontrivial result obtained from the *ab initio* calculations of the VHC.

In Fig. 10 we show the photoinduced AHC of the Dirac model calculated with $x = 0.01$. The first thing to note is that $\delta\sigma_{xy}^0 + \delta\sigma_{xy}^{SJ} \sim -\delta\sigma_{xy}^0$, as also noted in Ref. [10]. While in the case of the VHC [Eqs. (15)–(17)] the side jump only contributes a small fraction of the scattering-independent part, it is seen to play a major role in the photoinduced AHC. Both the side-jump and intrinsic contributions to $\delta\sigma_{xy}$ depend only weakly on the gate voltage—increasing and decreasing respectively when the Fermi level is raised. As was the case for the VHC, at low impurity concentrations the photoinduced AHC is dominated by skew scattering and scales as $1/x$. Interestingly, the skew-scattering contribution also shows a significant dependence on ε_F, which could explain the large dependence on gate voltage that is observed in experiments [10].

IV. CONCLUSIONS

A priori it is not clear that the intrinsic VHC provides a good descriptor for the VHC in a MoS$_2$ transistor setup, since the Hall conductivity is expected to diverge in the clean limit as a consequence of skew scattering [10]. Here we have used first principles calculations to investigate the effect of sulfur vacancies at different impurity concentrations. The influence of disorder was analyzed in k space in terms of the unfolded Berry curvature, and we have shown that the side-jump regime appears at moderate impurity concentrations as a concentration-independent enhancement of the Berry curvature. The skew scattering introduces divergences in the Berry curvature and the unfolded Berry curvature becomes spiky and irrelevant. Nevertheless, we were able to converge the VHC calculations in the skew scattering regime, and recover the expected divergent behavior. The *ab initio* calculations show qualitative agreement with model calculations based on a massive Dirac Hamiltonian and allow us to extract a nontrivial value of the skew scattering potential $\langle V_i^3 \rangle_c / \langle V_i^2 \rangle_c^2 = -0.4$ eV. The calculations allow us to estimate a critical impurity concentration x where skew scattering starts to dominate. For $\varepsilon_F - \varepsilon_c = 40$ meV we find $x \sim 0.005$, and below this point the intrinsic contribution becomes a poor descriptor of the VHC.

A comparison with experiments [10] indicates that indeed, the intrinsic VHC cannot be applied as a descriptor for the photoinduced valley Hall conductivity. As previously noted, the side-jump contribution changes the sign of photoinduced AHC and we have shown that the skew scattering contribution is a likely explanation for the large gate dependence observed experimentally [10]. However, a reliable estimate of the skew scattering contribution requires knowledge of the impurity concentrations in the samples investigated, which is not presently available. It would be very interesting to perform measurements of the photoinduced AHC on MoS$_2$ samples with different impurity concentrations in order to unravel the roles played by side-jump and skew scattering.

As we have neglected phonon contributions, our results are only valid at low temperatures. The effect of phonons on the longitudinal mobility in MoS$_2$ has been analyzed thoroughly in Ref. [34], but the influence on the transverse conductivity has so far not been considered. Furthermore, monolayer MoS$_2$ has been shown to exhibit strong excitonic effects [35,36], due to the poor screening in two-dimensional materials. The charge imbalance utilized in the experimental realization originates from optically generated electron-hole pairs, and these effects may severely limit the carrier mobility if the Fermi level is close to the conduction-band edge. We will leave these issues for future studies.

APPENDIX A: CALCULATIONAL DETAILS

The calculations in the present work were performed with the tight-binding method using parameters obtained from *ab initio* density functional calculations and projected Wannier functions.

1. *Ab initio* calculations and construction of Wannier orbitals

The *ab initio* density functional theory calculations were performed with the `pwscf` code from the `Quantum Espresso` package [37], using the PBE functional. Norm-conserving pseudopotentials were used, and the calculations were carried out with a plane wave cutoff of 100 Ry. The lattice parameter of MoS$_2$ was set to the experimental lattice constant of 3.16 Å, and 12 Å was used to separate the periodically repeated images. All calculations were performed in a noncollinear spin framework, with fully relativistic pseudopotentials.

After converging the Kohn-Sham electronic structure, the valence and low-lying conduction Bloch bands were converted into projected Wannier functions using the `Wannier90` code package [38]. For MoS$_2$ the projected Wannier orbitals were constructed using sulfur p states and Mo d states. The sulfur s states were included in the *ab initio* calculations, but the low-lying s-like Bloch bands were excluded from the Wannierization. Unoccupied states were included by disentangling [39] bands up to 3.0 eV above the conduction-band minimum ε_c. Finally the Wannier functions were used to construct the Kohn-Sham Hamiltonian in a tight-binding basis $H_{ij}(\mathbf{R})$, where i, j denotes orbital indices within the unit cell and $\mathbf{R}$ is a lattice vector. The set of lattice vectors included were defined by the Wigner-Seitz supercell corresponding to the applied *ab initio* k-point mesh. For example, in pristine MoS$_2$ with a 8×8 k-point mesh, we have 22 orbitals (Mo d and S p) and 64 lattice vectors (some of which are equivalent, as they lie on the supercell boundary).

2. Tight-binding calculations

The majority of calculations in the present work are tight-binding calculations with parameters obtained from a Wannier representation of the Kohn-Sham Hamiltonian $H_{ij}(\mathbf{R})$. In a tight-binding framework, the calculation of band structures from $H_{ij}(\mathbf{R})$ is of course equivalent to the standard Wannier interpolation [39]. At the sampled set of k points, $H_{ij}(\mathbf{R})$ will thus yield the calculated Kohn-Sham eigenvalues and between the sampled points it smoothly interpolates. Similarly, a rigorous Wannier interpolation scheme can be constructed for the AHC [40], but this quantity cannot be calculated exactly in a bare tight-binding framework since the information contained in $H_{ij}(\mathbf{R})$ is not enough to evaluate the Berry curvature.

To see this explicitly we will briefly state the relevant expressions below. The starting point is the Berry curvature in its spectral representation where it can be written

$$\Omega_{\alpha\beta}(\mathbf{k}) = \sum_{m,n} (f_{n\mathbf{k}} - f_{m\mathbf{k}})$$
$$\times \frac{\langle u_{m\mathbf{k}} | \nabla_\alpha H(\mathbf{k}) | u_{n\mathbf{k}} \rangle \langle u_{n\mathbf{k}} | \nabla_\beta H(\mathbf{k}) | u_{m\mathbf{k}} \rangle}{(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}})^2}, \quad \text{(A1)}$$

with $\nabla_\alpha \equiv \partial/\partial k_\alpha$. We let φ_i denote a set of localized orbitals and expand the Bloch states as

$$|\psi_{n\mathbf{k}}\rangle = \sum_i C_{ni\mathbf{k}}|\chi_{i\mathbf{k}}\rangle = \sum_{i\mathbf{R}} C_{ni\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{R}+\mathbf{t}_i)}|\varphi_{i\mathbf{R}}\rangle, \tag{A2}$$

where $\mathbf{t}_i = \langle\phi_{i\mathbf{0}}|\hat{\mathbf{r}}|\phi_{i\mathbf{0}}\rangle$. Note that the inclusion of $\mathbf{t}_i$ is purely a matter of convention. However, the present convention will prove highly convenient below. Matrix elements of the Bloch Hamiltonian and their gradients can now be written as

$$H_{ij\mathbf{k}} = \langle\chi_{i\mathbf{k}}|\hat{H}|\chi_{j\mathbf{k}}\rangle = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{t}_i+\mathbf{t}_j)} H_{ij}(\mathbf{R}), \tag{A3}$$

$$\nabla_{\mathbf{k}} H_{ij\mathbf{k}} = i\sum_{\mathbf{R}} (\mathbf{R}-\mathbf{t}_i+\mathbf{t}_j) e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{t}_i+\mathbf{t}_j)} H_{ij}(\mathbf{R}) \tag{A4}$$

and in terms of these, the matrix elements appearing in Eq. (A1) become

$$
\begin{aligned}
J_{\alpha mn}(\mathbf{k}) &\equiv \langle u_{m\mathbf{k}}|\nabla_\alpha H(\mathbf{k})|u_{n\mathbf{k}}\rangle \\
&= \sum_{ij\mathbf{R}} C_{mi\mathbf{k}}^* C_{nj\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{t}_i+\mathbf{t}_j)} \langle\varphi_{i\mathbf{0}}|e^{i\mathbf{k}\cdot\hat{\mathbf{r}}}\nabla_\alpha H(\mathbf{k}) e^{-i\mathbf{k}\cdot\hat{\mathbf{r}}}|\varphi_{j\mathbf{R}}\rangle \\
&= -i\sum_{ij\mathbf{R}} C_{mi\mathbf{k}}^* C_{nj\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{t}_i+\mathbf{t}_j)} \langle\varphi_{i\mathbf{0}}|[\hat{\mathbf{r}},\hat{H}]|\varphi_{j\mathbf{R}}\rangle \\
&= -i\sum_{ijl\mathbf{R}\mathbf{R}'} C_{mi\mathbf{k}}^* C_{nj\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{t}_i+\mathbf{t}_j)} [\langle\varphi_{i\mathbf{0}}|\hat{\mathbf{r}}|\varphi_{l\mathbf{R}'}\rangle\langle\varphi_{l\mathbf{R}'}|\hat{H}|\varphi_{j\mathbf{R}}\rangle - \langle\varphi_{i\mathbf{0}}|\hat{H}|\varphi_{l\mathbf{R}'}\rangle\langle\varphi_{l\mathbf{R}'}|\hat{\mathbf{r}}|\varphi_{j\mathbf{R}}\rangle] \\
&= -i\sum_{ijl\mathbf{R}\mathbf{R}'} C_{mi\mathbf{k}}^* C_{nj\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{R}+\mathbf{R}'-\mathbf{t}_i+\mathbf{t}_j)} [\mathbf{r}_{il\mathbf{R}'} H_{lj\mathbf{R}} - H_{il\mathbf{R}'}(\mathbf{r}_{lj\mathbf{R}}+\mathbf{R}'\delta_{\mathbf{R0}}\delta_{lj})] \\
&= -i\sum_{ijl} C_{mi\mathbf{k}}^* C_{nj\mathbf{k}} [\mathbf{r}_{il\mathbf{k}} H_{lj\mathbf{k}} - H_{il\mathbf{k}}\mathbf{r}_{lj\mathbf{k}}] + i\sum_{ij\mathbf{R}} C_{mi\mathbf{k}}^* C_{nj\mathbf{k}} \mathbf{R} e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{t}_i+\mathbf{t}_j)} H_{ij\mathbf{R}} \\
&= \sum_{ij} C_{mi\mathbf{k}}^* C_{nj\mathbf{k}} \nabla_\alpha H_{ij\mathbf{k}} + i\sum_{ijl} C_{mi\mathbf{k}}^* C_{nj\mathbf{k}} [(\mathbf{t}_i\delta_{il}-\mathbf{r}_{il\mathbf{k}}) H_{lj\mathbf{k}} - H_{il\mathbf{k}}(\mathbf{t}_l\delta_{lj}-\mathbf{r}_{lj\mathbf{k}})].
\end{aligned}
\tag{A5}
$$

In the present work we have made the diagonal approximation where $\mathbf{r}_{ij\mathbf{R}} = \delta_{\mathbf{R0}}\delta_{ij}\mathbf{t}_i$ and we thus use $J_{\alpha mn}(\mathbf{k}) = \sum_{ij} C_{mi\mathbf{k}}^* C_{nj\mathbf{k}} \nabla_\alpha H_{ij\mathbf{k}}$. With this approximation, the problem can be mapped exactly to a tight-binding calculation with parameters obtained from the Kohn-Sham Hamiltonian in a basis of Wannier functions. This turns out to be an excellent approximation. The Berry curvature calculated with this method cannot be distinguished with the naked eye from *ab initio* calculations (obtained by Wannier interpolation) and anomalous Hall conductivities calculated within the diagonal approximation differ from *ab initio* results by less than 2%.

3. Mapping the supercell onto the normal cell

Unfolding band structures and curvatures involve the calculation of matrix elements between pristine and supercell systems. A localized basis set allow us to perform the unfolding without direct reference to the pristine system [28].

We denote pristine band, orbital indices, and crystal momentum by $n,i,\mathbf{k}$, respectively, and supercell band, orbital, and crystal momentum indices by $N,I,\mathbf{K}$, respectively. For the present purpose we will use $\mathbf{r}$ for pristine lattice vectors and $\mathbf{R}$ for supercell lattice vectors. We can thus consider the matrix element

$$\langle\psi_{n\mathbf{k}}|\psi_{N\mathbf{K}}\rangle = \sum_{iI\mathbf{r}\mathbf{R}} C_{ni\mathbf{k}}^* C_{NI\mathbf{K}} e^{-i\mathbf{k}\cdot(\mathbf{r}+\mathbf{t}_i)+i\mathbf{K}\cdot(\mathbf{R}+\mathbf{t}_I)} \langle\varphi_{i\mathbf{r}}|\varphi_{I\mathbf{R}}\rangle. \tag{A6}$$

Following Ku *et al.* [28] we introduce a map that uniquely identifies orbitals in the supercell with corresponding orbitals in the pristine system. The map thus takes $I \to \mathbf{r}'(I), i'(I)$ and we have

$$\langle\varphi_{i\mathbf{k}}|\varphi_{I\mathbf{R}}\rangle = \langle\varphi_{i\mathbf{r}}|\varphi_{i'(I)\mathbf{R}+\mathbf{r}'(I)}\rangle = \delta_{ii'(I)}\delta_{\mathbf{r}\mathbf{R}+\mathbf{r}'(I)}. \tag{A7}$$

The matrix element can then be written as

$$
\begin{aligned}
\langle\psi_{n\mathbf{k}}|\psi_{N\mathbf{K}}\rangle &= \sum_{I\mathbf{R}} C_{ni'(I)\mathbf{k}}^* C_{NI\mathbf{K}} e^{-i\mathbf{k}\cdot[\mathbf{r}'(I)+\mathbf{R}+\mathbf{t}_{i'(I)}]+i\mathbf{K}\cdot[\mathbf{R}+\mathbf{t}_I]} \\
&= \sum_I C_{ni'(I)\mathbf{k}}^* C_{NI\mathbf{K}} e^{-\mathbf{k}\cdot[\mathbf{r}'(I)+\mathbf{t}_{i'(I)}]+i\mathbf{K}\cdot\mathbf{t}_I} \delta_{\mathbf{K}[\mathbf{k}]}, \tag{A8}
\end{aligned}
$$

where $[\mathbf{k}]$ is the set of crystal momenta that downfolds to $\mathbf{K}$.

The idea of a map allows one to only work with the supercell system and avoid the explicit calculation of overlap matrices. However, the procedure does require that the supercell system considered is somewhat similar to the pristine reference system and becomes ill defined if there is not a unique way of relating orbitals in the two systems. Furthermore, it is important to construct the tight-binding Hamiltonian from projected Wannier functions as opposed to maximally localized Wannier functions, since the latter may differ significantly in otherwise similar systems.

In Figs. 5 and 6 we showed examples of the band structures and Berry curvature of MoS_2 in a 3×3 unit cell with a single sulfur vacancy, unfolded onto the normal BZ.
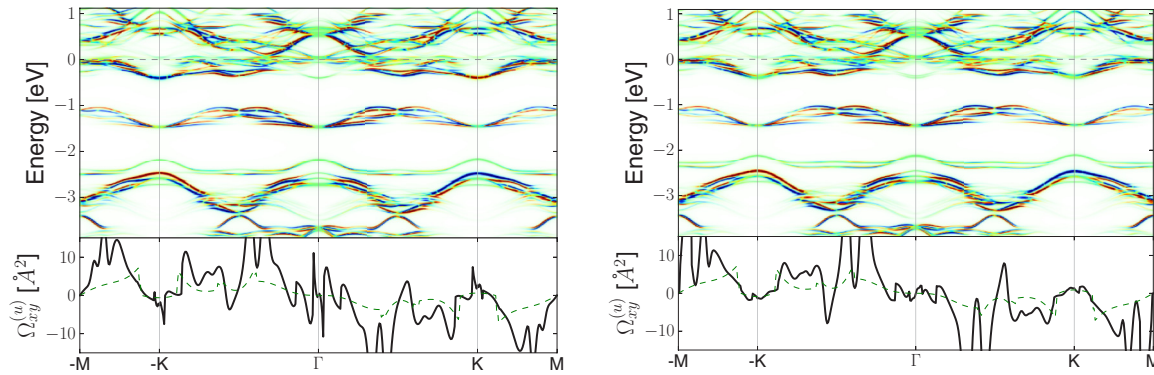
FIG. 11. (Color online) Unfolded spectral function (top panels) and unfolded Berry curvature (bottom panels) of MoS$_2$ in a 3×3 supercell with two sulfur vacancies at next-nearest-neighbor sites. Energies are measured from the Fermi level. Left: *ab initio* calculation. Right: calculation using the effective Hamiltonian method.

APPENDIX B: EFFECTIVE HAMILTONIANS

Here we briefly summarize the construction of effective Hamiltonians as proposed in Ref. [27]. In a tight-binding framework, the effective Hamiltonian with N impurities is constructed as

$$H_{IJ}^{Eff}(\mathbf{R}) = H_{IJ}^{SC}(\mathbf{R}) + \sum_{s=1}^{N} P(I,J,\mathbf{r}_s,\mathbf{R}) H'_{i'(I)i'(J)}$$
$$\times (\mathbf{r}'(I) - \mathbf{r}_s, \mathbf{r}'(J) + \mathbf{R} - \mathbf{r}_s) \quad \text{(B1)}$$

where $I,J,\mathbf{R}$ denotes supercell orbitals and lattice vectors, respectively. $\mathbf{r}$ is a normal cell lattice vector and $i',\mathbf{r}'$ are maps from the supercell orbitals and lattice to the normal cell. $\mathbf{r}_s$ denotes the position of impurity s. The influence Hamiltonian is given by

$$H'_{ij}(\mathbf{r}_1,\mathbf{r}_2) = H_{ij}^{\text{Imp}}(\mathbf{r}_1,\mathbf{r}_2) - H_{ij}^{0}(\mathbf{r}_2 - \mathbf{r}_1), \quad \text{(B2)}$$

where $H_{i,j}^{\text{Imp}}(\mathbf{r}_1,\mathbf{r}_2)$ is constructed with a map from the impurity to the normal cell. Since this is expressed in a basis of normal cell lattice vectors H^{Imp} and therefore H' is not periodic in simultaneous translations of $\mathbf{r}_1$ and $\mathbf{r}_2$. We have used the partition function of Liu and Vanderbilt [41]

$$P(I,J,\mathbf{r}_s,\mathbf{R}) = e^{-(d/d_c)^8}$$
$$d = (|\mathbf{r}'(I) - \mathbf{r}_s| + |\mathbf{r}'(J) + \mathbf{R} - \mathbf{r}_s|)/2, \quad \text{(B3)}$$

with $d_c = 9.0$ Å.

For applications of the method we have performed *ab initio* calculations of the pristine and impurity systems and constructed H^{imp} and H^0 using projected Wannier functions. H^{SC} is then constructed as a straightforward repetition of H^0.

As a nontrivial test of the method we have performed an *ab initio* calculation of a 3×3 unit cell of MoS$_2$ and with sulfur vacancies at two next-nearest-neighbor sites. We have then constructed the same system from the effective Hamiltonian: first we construct the tight-binding model of a 3×3 unit cell of MoS$_2$ by repeating the tight-binding Hamiltonian obtained from calculations of pristine MoS$_2$ (normal unit cell). We have then constructed the influence Hamiltonian (B2) from the pristine calculation and a calculation of MoS$_2$ in a 3×3 unit cell with a *single* sulfur vacancy. The influence Hamiltonian is then added to the 3×3 tight-binding Hamiltonian at the two nearest neighbor sites to obtain the effective Hamiltonian of a system with two impurities. The unfolded bands and curvature of *ab initio* and effective Hamiltonian calculations are shown in Fig. 11. The unfolded spectral function is nearly indistinguishable in the two cases. However, the unfolded Berry curvature is highly sensitive to the exact positions of bands near avoided crossings and therefore exhibits larger deviations in the two methods. Nevertheless, the Berry curvature obtained from the effective Hamiltonian reproduces the main qualitative features (for example, a vanishing contribution in the vicinity of K and $-K$) and we believe that the resulting configuration-averaged Berry curvature provides a reliable measure of the effects of disorder on the Berry curvature and VHC.

[1] X. Xu, W. Yao, D. Xiao, and T. F. Heinz, Nat. Phys. **10**, 343 (2014).

[2] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, Phys. Rev. Lett. **105**, 136805 (2010).

[3] D. Xiao, W. Yao, and Q. Niu, Phys. Rev. Lett. **99**, 236809 (2007).

[4] D. Xiao, G.-B. Liu, W. Feng, X. Xu, and W. Yao, Phys. Rev. Lett. **108**, 196802 (2012).

[5] H. Zeng, J. Dai, W. Yao, D. Xiao, and X. Cui, Nat. Nanotechnol. **7**, 490 (2012).

[6] K. F. Mak, K. He, J. Shan, and T. F. Heinz, Nat. Nanotechnol. **7**, 494 (2012).

[7] T. Cao, G. Wang, W. Han, H. Ye, C. Zhu, J. Shim, Q. Niu, P. Tan, E. Wang, B. Liu, and J. Feng, Nat. Commun. **3**, 887 (2012).

[8] D. Xiao, M.-C. Chang, and Q. Niu, Rev. Mod. Phys. **82**, 1959 (2010).

[9] N. Nagaosa, J. Sinova, S. Onoda, A. H. MacDonald, and N. P. Ong, Rev. Mod. Phys. **82**, 1539 (2010).

[10] K. F. Mak, K. L. McGill, J. Park, and P. L. McEuen, Science **344**, 1489 (2014).

[11] N. A. Sinitsyn, Q. Niu, and A. H. MacDonald, Phys. Rev. B **73**, 075318 (2006).

[12] N. A. Sinitsyn, A. H. MacDonald, T. Jungwirth, V. K. Dugaev, and J. Sinova, Phys. Rev. B **75**, 045315 (2007).

[13] W. Zhou, X. Zou, S. Najmaei, Z. Liu, Y. Shi, J. Kong, J. Lou, P. M. Ajayan, B. I. Yakobson, and J.-C. Idrobo, Nano Lett. **13**, 2615 (2013).

[14] D. Liu, Y. Guo, L. Fang, and J. Robertson, Appl. Phys. Lett. **103**, 183113 (2013).

[15] Y. Ma, Y. Dai, M. Guo, C. Niu, J. Lu, and B. Huang, Phys. Chem. Chem. Phys. **13**, 15546 (2011).

[16] Y. Zhou, P. Yang, H. Zu, F. Gao, and X. Zu, Phys. Chem. Chem. Phys. **15**, 10385 (2013).

[17] J.-w. Wei, Z.-w. Ma, H. Zeng, Z.-y. Wang, Q. Wei, and P. Peng, AIP Adv. **2**, 042141 (2012).

[18] M. Ghorbani-Asl, A. N. Enyashin, A. Kuc, G. Seifert, and T. Heine, Phys. Rev. B **88**, 245440 (2013).

[19] H.-P. Komsa and A. V. Krasheninnikov, Phys. Rev. B **91**, 125304 (2015).

[20] B. Velický, Phys. Rev. **184**, 614 (1969).

[21] J. S. Faulkner and G. M. Stocks, Phys. Rev. B **21**, 3222 (1980).

[22] W. H. Butler, Phys. Rev. B **31**, 3260 (1985).

[23] S. Lowitzer, D. Ködderitzsch, and H. Ebert, Phys. Rev. Lett. **105**, 266604 (2010).

[24] B. Zimmermann, K. Chadova, D. Ködderitzsch, S. Blügel, H. Ebert, D. V. Fedorov, N. H. Long, P. Mavropoulos, I. Mertig, Y. Mokrousov, and M. Gradhand, Phys. Rev. B **90**, 220403 (2014).

[25] J. Kudrnovský, V. Drchal, and I. Turek, Phys. Rev. B **88**, 014422 (2013).

[26] J. Weischenberg, F. Freimuth, J. Sinova, S. Blügel, and Y. Mokrousov, Phys. Rev. Lett. **107**, 106601 (2011).

[27] T. Berlijn, D. Volja, and W. Ku, Phys. Rev. Lett. **106**, 077005 (2011).

[28] W. Ku, T. Berlijn, and C.-C. Lee, Phys. Rev. Lett. **104**, 216401 (2010).

[29] R. Bianco, R. Resta, and I. Souza, Phys. Rev. B **90**, 125153 (2014).

[30] F. D. M. Haldane, Phys. Rev. Lett. **61**, 2015 (1988).

[31] M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).

[32] M. Rösner, S. Haas, and T. O. Wehling, Phys. Rev. B **90**, 245105 (2014).

[33] T. Brumme, M. Calandra, and F. Mauri, Phys. Rev. B **91**, 155436 (2015).

[34] K. Kaasbjerg, K. S. Thygesen, and K. W. Jacobsen, Phys. Rev. B **85**, 115317 (2012).

[35] F. Hüser, T. Olsen, and K. S. Thygesen, Phys. Rev. B **88**, 245309 (2013).

[36] D. Y. Qiu, F. H. da Jornada, and S. G. Louie, Phys. Rev. Lett. **111**, 216805 (2013).

[37] P. Giannozzi *et al.*, J. Phys.: Condens. Matter **21**, 395502 (2009).

[38] N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, Rev. Mod. Phys. **84**, 1419 (2012).

[39] I. Souza, N. Marzari, and D. Vanderbilt, Phys. Rev. B **65**, 035109 (2001).

[40] X. Wang, J. R. Yates, I. Souza, and D. Vanderbilt, Phys. Rev. B **74**, 195118 (2006).

[41] J. Liu and D. Vanderbilt, Phys. Rev. B **88**, 224202 (2013).