Velocity and confinement of edge plasmons in HgTe-based two-dimensional topological insulators

Alexandre Gourmelon,[1,*] Elric Frigerio,[1,*] Hiroshi Kamata,[1,2,3] Lukas Lunczer,[4,5] Anne Denis,[1] Pascal Morfin,[1] Michael Rosticher,[1] Jean-Marc Berroir,[1] Gwendal Fève,[1] Bernard Plaçais,[1] Hartmut Buhmann,[4,5] Laurens W. Molenkamp,[4,5] and Erwann Bocquillon[1,6,†]

[1]*Laboratoire de Physique de l'École Normale Supérieure, ENS, PSL Research University, CNRS, Sorbonne Université, Université de Paris, 24 rue Lhomond, 75231 Paris Cedex 05, France*
[2]*NTT Basic Research Laboratories, NTT Corporation, 3-1 Morinosato-Wakamiya, Atsugi, Kanagawa 243-0198, Japan*
[3]*JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan*
[4]*Physikalisches Institut (EP3), Am Hubland, Universität Würzburg, D-97074 Würzburg, Germany*
[5]*Institute for Topological Insulators, Am Hubland, Universität Würzburg, D-97074 Würzburg, Germany*
[6]*II. Physikalisches Institut, Universität zu Köln, Zülpicher Str. 77, 50937 Köln, Germany*

(Received 23 February 2023; revised 27 May 2023; accepted 21 June 2023; published 12 July 2023)

High-frequency transport in the edge states of the quantum spin Hall (QSH) effect has rarely been explored, though it could cast light on the scattering mechanisms taking place therein. Here we report on the measurement of the plasmon velocity in topological HgTe quantum wells both in the QSH and quantum Hall (QH) regimes, using harmonic GHz excitations and phase-resolved detection. We observe low plasmon velocities in both regimes, with, in particular, large transverse widths in the QH regime despite a sharp edge confinement profile. We ascribe these observations to the prominent influence of charge puddles forming in the vicinity of edge channels. Together with other recent works, it suggests that puddles play an essential role in the edge state physics and probably constitute a main hurdle on the way to clean and robust edge transport.

DOI: 10.1103/PhysRevB.108.035405

Since its experimental discovery [1] in 2007, the quantum spin Hall (QSH) effect has been intensively studied in two-dimensional topological insulators [2]. It manifests as a pair of helical edge states, with opposite directions and opposite spins while the bulk remains insulating. Such phases of matter, driven by a topological band inversion [3], realize time-reversal symmetric topological insulators, with a $\mathbb{Z}_2$ topological invariant. Their helical edge states offer an exciting playground to study spin-polarized edge transport and topological superconductivity. As a result, possible applications in both spintronics [4] and topological quantum computation, based on the braiding of Majorana excitations [5], are envisioned. Prominent transport signatures have been observed in HgTe quantum wells (QWs) such as nonlocal and spin-polarized transport [6–9], or the fractional Josephson effect in HgTe-based Josephson junctions [10,11]. Alternatively, InAs/GaSb double QWs [12–14], or layered materials such as bismuthene [15] or WTe_2 [16], have also been successfully identified as QSH insulators.

In this context, the investigation of high-frequency transport in 2D topological insulators, such as HgTe quantum wells here, is of high interest. The charge relaxation scales of the QSH edge carriers has been measured [17] by microwave capacitance spectroscopy [18,19], revealing that the edge states have a larger than predicted density of state, possibly due to neighboring puddles. It also suggests that the QSH effect could be enhanced in dynamical studies by exploiting the difference in transport or scattering timescales between topological and bulk carriers. Here we explore another aspect, namely the velocity of plasmons propagating in the edge channels. In the quantum Hall effect of GaAs, InAs, or graphene samples, the velocities of chiral edge magnetoplasmons have been widely studied, highlighting the role of intra and interchannel Coulomb interaction, of the confinement edge potential, or of the screening of Coulomb interaction by nearby metallic gates, and of dissipation in the bulk [20–29].

Here, we report on a systematic study of the velocities of plasmons in a HgTe quantum well, in the classical and quantum Hall regime (magnetic fields B up to 8 T) of the conduction band, as well as in the topological gap where the quantum spin Hall effect takes place (at $B = 0$). The measurements are performed in a dilution refrigerator at a temperature of $T \simeq 20$ mK, for frequencies $f \simeq 3–10$ GHz. The (phase) velocity is accessed via the phase accumulated by a plasmon excitation propagating between a local source and a probe contact in the HgTe QWs. Though phase and group velocities may differ, they are known to coincide in the low-energy limit, which we experimentally confirm (see [30]). The phase shift can be rather accurately measured, even on small distances in which time-resolved techniques [26,34,35] would be inoperable due to insufficient delay. This allows for a rather short propagation length l (ranging between 3 and 7 μm), in order to approach the ballistic length which does not exceed 1 μm in our device. Finally, the electron density n is tuned via to a gate voltage V_g in a large range $n \simeq 0.5 \times 10^{11} - 5 \times 10^{11}$ cm^{-2} in the conduction band. Our main

*These authors contributed equally to this work.
†bocquillon@ph2.uni-koeln.de

observations can be summarized as follows. First, when the Fermi energy is in the conduction band and under the action of a perpendicular magnetic field B, we observe a transition in the magnetic-field dependent velocity, suggestive of the crossover between noninteracting edge states (Landauer-Büttiker picture, abbreviated as LB) and edge reconstruction under e-e interactions (Chklovskii-Shklovskii-Glazman regime [36,37], denoted CSG). From the analysis of the velocity, we conclude that the edge states have a typical width w of several microns at low fields, probably set by the electrostatic disorder, while the edge confinement itself occurs, as expected, on a typical scale $l \sim 0.1$ μm comparable to the distance d between the gate and the HgTe layer. Second, we confirm this interpretation by an analysis of the observed low velocities in the QSH gap of the device.

The article is organized in three sections. In the first section, we introduce the device geometry, and the preliminary characterization of the samples via DC magnetotransport measurements. The microwave measurement setup is briefly described in the second section, together with the post-acquisition calibration process. Finally, we detail several experimental results in the QSH and QH regime, and present a plausible interpretation based on the presence of charge puddles in the band gap of the material.

I. SAMPLE GEOMETRY AND DC TRANSPORT PROPERTIES

a. Samples. The samples are fabricated from HgTe/$Cd_{0.68}Hg_{0.32}Te$ QWs grown by molecular beam epitaxy. The thickness t of the QWs is 8.5 nm. For such a thickness, the band structure consists of light electrons in the conduction band, and heavy holes in the valence band. A topological phase transition for thickness $t > t_c \simeq 6.3$ nm enforces the presence of QSH edge states in the gap of the QWs [3]. A gap of approximately 26 meV is predicted by $\mathbf{k}\cdot\mathbf{p}$ simulations of the band structure (estimated along the $k_x = \pm k_y$ direction, in which it is minimal). Additionally, the QW is protected by a $Cd_{0.68}Hg_{0.32}Te$ capping layer of thickness 50.5 nm. The QWs are first characterized using standard Hall-bar measurements, yielding a mobility of $1-2\times10^5$ cm^2 V^{-1} s^{-1} (measured at a density $n \simeq 2-3\times10^{11}$ cm^{-2} in the conduction band). Three devices have been investigated and have given similar results. Each device comprises a rectangular mesa defined via a wet-etching technique [9] to preserve the high crystalline quality and the high mobility of the epilayer. Low-resistance ohmic contacts are evaporated on either end of the mesa, for both DC characterization and RF measurements. Two gold finger gates (denoted RFg 1 and RFg 2 in Fig. 1) are patterned with e-beam lithography, with a width $\delta \simeq 800$ nm. and are used to locally and capacitively excite the underneath QW with high-frequency signals, while an additional gate for DC tuning of the electron density (DCg) covers the rest of the mesa. All gates are evaporated on top of a 16 nm-thick HfO_2 insulating layer, grown by low-temperature atomic layer deposition (ALD) [9]. The main text focuses on one device where the global gate DCg is made of Au (denoted "Sample A"), as presented in Fig. 1(b). Another sample covered with a thin Pd global gate electrode (denoted "Sample B") is also briefly discussed in the main text, with more data presented in

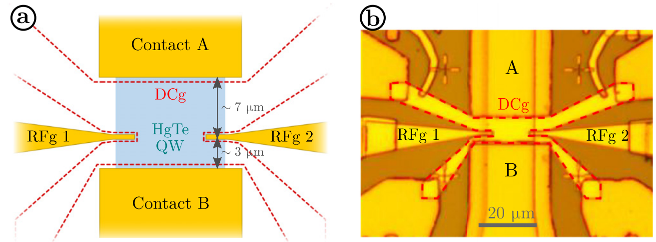

FIG. 1. Sample geometry. (a) Sketch of the device: The light blue part is the HgTe mesa while the yellow parts (finger-gate electrodes RFg 1, 2 and ohmic contacts A, B) are made of gold. The red dashed region corresponds to the space covered by the top-gate DCg. A DC voltage V_g is applied to RFg 1, 2 and DCg to uniformly tune the electron density n. (b) Image taken with an optical microscope of Sample A, showing the different gates and contacts of the sample as sketched in (a).

the Supplemental Material [30]. The results are very similar in both cases, though our observations point toward larger puddles in the electrostatic landscape of Sample B.

b. Characterization of the transport properties. The two-terminal resistance R_{2T} is measured at $B = 0$ T as a function of the gate voltage V_g (applied simultaneously to all three gates to preserve a uniform electron density), and presented in Fig. 2(a). It informs on the band structure and the position of the gap in the devices. Similarly to previous works, we identify the gap as a clear peak in R_{2T}, separating the conduction and valence band regimes. However, the peak value of R_{2T} is much larger than the expected quantized value $R_K/2 = h/2e^2 \simeq 12.9$ kΩ, and lies around 120 kΩ. The full *in situ* characterization of the mobility and electron density from the magnetoresistance of the devices is rendered impractical by the two-terminal geometry imposed by the microwave measurements. Indeed, the two-terminal resistance involves both the Hall and longitudinal resistance, which are easily separated in a four-terminal geometry. Nevertheless, one observes in Fig. 2(b) that R_{2T} clearly exhibits the quantum Hall plateaus. Assuming that these plateaus reach perfectly quantized values, we write $R_{2T} = R_K/\nu + R_c$ at the center of each plateau of filling factor $\nu \in \mathbb{N}^*$, where R_c is a contact resistance. While the contact resistance is estimated at $R_c \simeq 100\,\Omega$ in the conduction band, it appears to be much higher in the valence band (20 kΩ), presumably due to the formation of p-n junctions near the n-doped contacts. The ratio R_K/R_{2T} as a color plot in Fig. 2(b) can then be used to fit the density by adjusting the set of lines $B_\nu = \frac{eR_K}{\nu}(\frac{c}{e}V_g + n_0)$ which define the exact integer filling factors. We then obtain the density $n(V_g) = \frac{c}{e}V_g + n_0$ where $c \simeq 1$ mF m^{-2} is the gate capacitance per unit area (in agreement with the theoretical estimate given the gate layer stack), and $n_0 \simeq 4.7\times10^{15}$ m^{-2} the density at $V_g = 0$.

II. MICROWAVE MEASUREMENTS AND CALIBRATION

a. Microwave setup. The phase shift in the device under test (DUT) is measured using a standard heterodyne detection method. An RF sine wave of frequency f in the GHz regime is generated from an arbitrary wave form generator (AWG) and sent to the RFg of the sample through the microwave

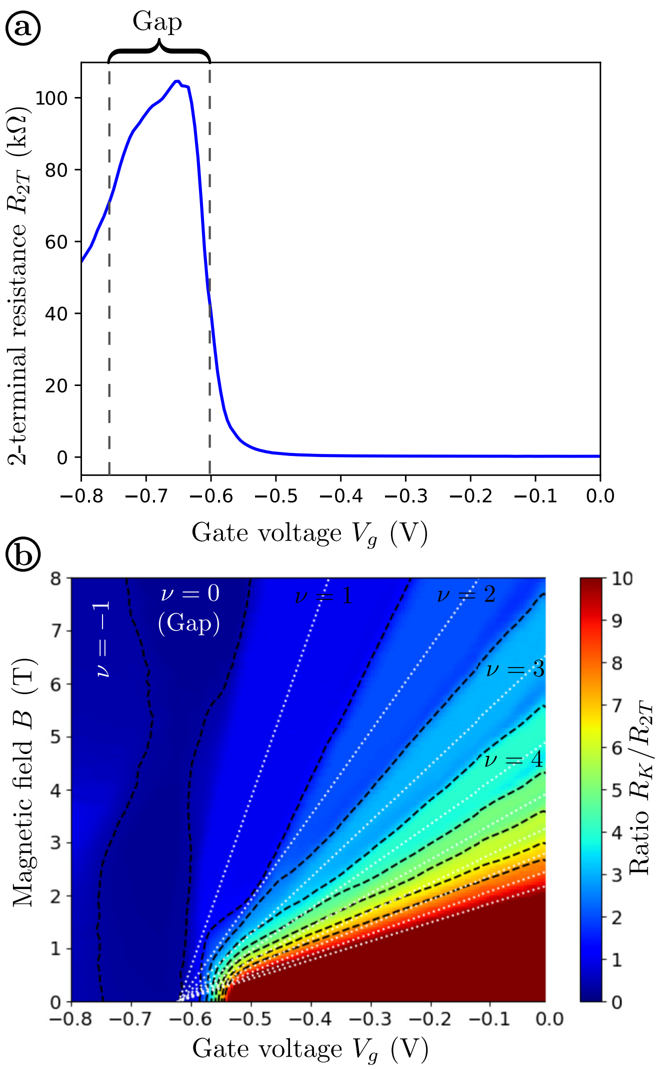

FIG. 2. DC transport properties of the sample. (a) Two-terminal resistance R_{2T} as a function of the gate voltage V_g exhibiting a peak signaling the gap (indicated by the dashed lines), and the conduction and valence bands on either sides of this peak. (b) 2D color map of the ratio R_K/R_{2T} as a function of gate voltage V_g and magnetic field B. The different filling factors ν are labeled, and the white dotted lines are the lines B_ν used to fit the carrier density n as a function of gate voltage V_g (see main text). The contours between the different QH plateaus are highlighted as dashed black lines. The color scale is intentionally saturated at a maximum value $R_K/R_{2T} = 10$ in order to distinguish more clearly the first QH plateaus.

lines of the fridge. At the excitation finger gate RFg, the signal amplitude is typically 1 mV. After being emitted by the finger gates RFg, the signal is collected by the two contacts (A and B). The signal is amplified by a cryogenic and room-temperature low-noise amplifiers before being sent to a heterodyne detection setup at room temperature. The signal is mixed with a local oscillator (LO), i.e., a sine wave generated signal generator detuned from the AWG output by 50 MHz. This mixing process converts the GHz signal coming from the sample to a 50 MHz signal which is then demodulated by a multichannel fast acquisition card to obtain the *in-phase* (I) and *in-quadrature* (Q) parts of the signal $I\cos(2\pi ft) + Q\sin(2\pi ft)$, in each contact A and B. With this setup, it is possible to simultaneously measure signals at the two contacts A and B in the range of frequencies $f \simeq 3 - 10\,\text{GHz}$ set by the cryogenic isolators placed before the cryogenic amplifiers. The full experimental setup is shown in [30].

b. Calibration of the raw data. The signal measured in the channels of the acquisition card needs further calibration and reference: (i) The measured magnitude is offset by stray couplings on the chip and sample holder, which do not contain any physical information on the topological device. This parasitic contribution is measured in a situation where the DUT is known to be perfectly insulating, and then subtracted. (ii) The phase is also affected by the propagation in the cables, and can not be directly used for computing the plasmon velocities. A phase reference needs to be defined from a situation where currents propagate at a very high velocity in the DUT (much larger than the edge plasmon velocities). We describe in Ref. [30] how we proceeded to these two steps, and how we controlled the validity of the underlying assumptions.

The calibration has been successfully conducted in numerous data sets. The chirality of the QH edge channels then manifests itself as a strong asymmetry with either the magnetic field directions, the choice of the contact (A or B), or the choice of the finger gate (RFg1 or RFg2) (see [30] for additional data and chirality maps). The phase also winds in a unique direction (clockwise). In Fig. 3, we show the resulting calibrated amplitude M of the microwave signal, its phase ϕ, and the velocity v calculated from the phase as $v = \frac{2\pi fL}{\phi}$, where L is the propagation length between finger gate and contact. The amplitude is close to zero on one half of the plane (here for $B > 0$), while it is strong on the other half (here for $B < 0$), and gradually decays with increasing field. The different filling factors are clearly visible and agree well with those determined from the DC magnetoresistance. In the regions where the amplitude M is sufficiently large, the phase can be unwrapped, allowing for the computation of the velocity in the same area.

Some samples and data sets have resisted such an analysis and exhibit asymmetric but not totally chiral behavior or do not allow to define an adequate phase reference. In agreement with our findings described later we attribute these phenomena to strong disorder in some samples, allowing for propagation of signal opposite to the expected propagation direction. We present problematic data sets in Ref. [30].

After calibrating the electron density n, the amplitude M and phase ϕ of the microwave data, we now explore the variations of the velocity v in the Hall regime as function of n, B but also the filling factor $\nu = \frac{hn}{eB}$.

III. RESULTS—PLASMON VELOCITIES

In this section, we analyze the measured velocity, and discuss different interesting observations. When a perpendicular magnetic field is applied, a clear transition is observed between low and high-field regimes, which we attribute to the crossover between the LB noninteracting regime to the CSG regime [36,37] at high fields where e-e interactions are prominent. A careful study of both regimes then yields information on the role of puddles and edge confinement in the device, which is relevant for both the classical and quantum

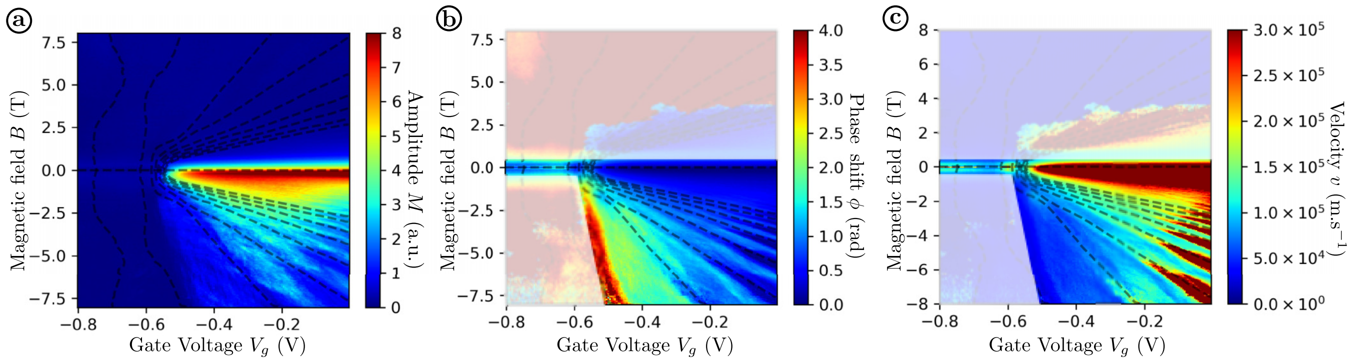

FIG. 3. Calibrated amplitude, phase, and velocity. Colormaps of the amplitude M (a), phase ϕ (b), and velocity v (c) as function of the gate voltage V_g applied on DC_g and the magnetic field B, obtained after calibration. The white shadings indicate regions where the signal amplitude M is too small, so that ϕ and v are not reliably computed.

Hall regime, but also indicative of the physics of QSH edge states. Though the data is not as clear, we also confirm these observations in the gap of the quantum well at zero magnetic field, i.e., when QSH edge states dominate transport.

a. Plasmon confinement in the quantum Hall effect. We first turn to the study of plasmon velocities in the quantum Hall regime, i.e., when a perpendicular magnetic field is applied to the sample. In gated samples, the velocity of the edge magnetoplasmons can be simply written as

$$v = \frac{ned}{\epsilon B w} = \frac{\sigma_{xy}}{C_{\mathrm{QH}}}, \tag{1}$$

where w is the transverse width of the edge plasmon, $\sigma_{xy} = ne/B$ the Hall conductance, and $C_{\mathrm{QH}} = \epsilon w/d$ the capacitance between gate and plasmon per unit length. This equation can be obtained from a microscopic derivation [38,39]. It also is a constitutive relation of a transmission line model for edge states [40], connecting the line impedance $1/\sigma_{xy}$ and the velocity v with the capacitance C_{QH}.

Through Eq. (1), the velocity v provides insights into the confinement of plasmons on a width w near the edges of the sample. In this context, the role of e-e interaction and screening in the progressive formation of edge states is well understood since pioneering works in the 90s [36,41], and have recently been numerically revisited [37].

In all measured devices, we observed two different behaviors depending on the strength of the magnetic field. At low field, the velocity is both proportional to n and $1/B$ [as illustrated in Fig. 4(a) for three different electron densities], in agreement with the Landauer-Büttiker model. In this model, a large number of edge states are uniquely defined by the edge confinement profile, while screening and reconstruction from e-e repulsion are irrelevant. This allows to define the n-independent width w_0 of the plasmon in this regime, and we find $w_0 \simeq 1.2\,\mu$m in Sample A ($w_0 \simeq 4.6\,\mu$m for Sample B).

This transverse width w_0 is much greater than the distance to the gate $d \simeq 50$ nm which controls the typical confinement length of the edge states, or than the magnetic length $l_B = \sqrt{\frac{\hbar}{eB}} \simeq 80$ nm at $B = 100$ mT. It indicates that the edge states are broadened, for example, by shallow potential fluctuations and puddles. Such very large values of w_0 have also been recently reported in Ref. [35], and similarly attributed to charge puddles. They result in an increased capacitance C_{QH} accounting for the gate-puddle coupling, an increased transverse width w_0 and equivalently to a reduced velocity v, irrespective of the edge confinement depletion length $l \simeq d$.

At higher fields [$B > B_c \simeq 2$ T in Fig. 4(a)], the velocity v strongly departs from this simple law, and shows strong oscillations. This crossover may be attributed to a reduced number of edge states, forming compressible and incompressible stripes under the influence of strong e-e interactions (CSG regime). We find that the crossover field B_c between both regimes is approximately compatible with the heuristic law [37] $B_c \propto n^{2/3}$ (see [30]). Such oscillations have already been observed in GaAs quantum wells [26] and originate from the transverse compression and decompression of plasmons when a new incompressible stripe nucleates in the bulk of the material at integer filling factors, and is progressively pushed toward the edges of the sample as ν increases [see Fig. 4(c)]. It is worth noting that oscillations of the velocity have also been observed in ungated graphene [34] and InAs quantum wells [42] with opposite behavior (minimal widths for integer filling factors), and are then ascribed to another mechanism, namely enhanced dissipation due to a conducting bulk.

Therefore, we continue the analysis by plotting the width w obtained from Eq. (1) as a function of the filling factor ν [see Fig. 4(b)]. At high filling factors $\nu \gg 15$ (i.e., low magnetic fields), w slowly converges toward its saturation value w_0. For low filling factors, we observe that w is maximum (i.e., the velocity v reaches its minima) at integer filling factors. The oscillations are very strongly visible at high densities $n > 3 \times 10^{11}$ cm^{-2} when screening is strong and thus, when the electrostatic disorder is less influential. In contrast, the oscillations are washed out at low densities. The oscillations of v and w are also visible though much fainter in Sample B (see [30]), as can be expected in a more disordered sample.

As shown in Ref. [26], the oscillations of w allow for reconstructing the edge density profile. We define the local density as $x \mapsto n_e(x) = nf(x)$ ranging from $n_e(x=0) = 0$ at the quantum well edge to $n_e(x) = n$ deep in the bulk of the material, as depicted in Fig. 4(c). The reconstruction is based on the following principles. The plasmon width w is essentially defined by the position of the innermost edge state (compressible stripe) located at a position x_{QH} such that the

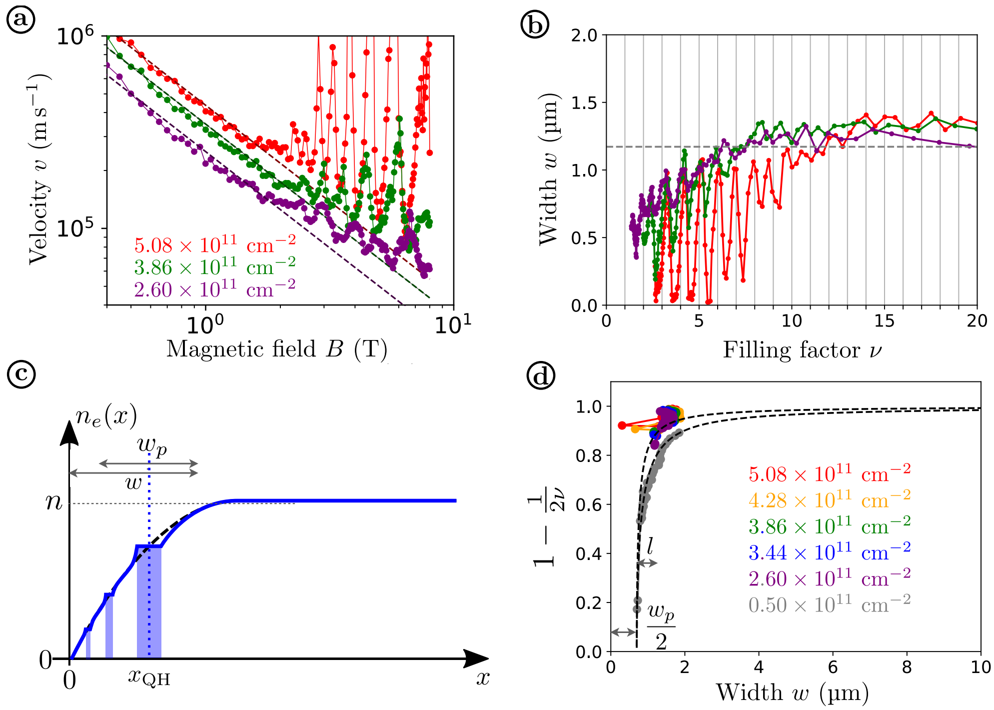

FIG. 4. Velocity and plasmon transverse width in the Hall regime. (a) Linecuts of the velocity v as function of magnetic field B for three values of the density n. The gray dashed line shows fits to the law $v \propto B^{-1}$, valid at low fields. In the high field region, v exhibits strong oscillations which become more pronounced as the density n increases. (b) Transverse width w as a function of filling factor ν. For large ν, i.e., low fields, w is approximately constant and independent of n. For low ν, the width w oscillates, showing minimum for integer filling factors $\nu \in \mathbb{Z}$. (c) Sketch of the edge density profile $n_e(x)$ as a function of the distance x from the edge: $n_e(x)$ saturates at $n_e(x) = n$ in the bulk of the material, and decreases to $n(x) = 0$ at the edge. The blue shades indicate the compressible stripes while the white stripes are the incompressible ones. The bare plasmon width is given by the position of the innermost Landau level $x_{\rm QH}$, and is further increased by $w_p/2$ due to puddles. (d) Normalized reconstructed edge profile $n_e(x)$ obtained by plotting $1 - 1/2\nu$ as a function of w for all data triplets (n, B, w). The obtained profiles are shown as colored dots for various values of the bulk density n. The dashed lines represent the heuristic edge profile $f(x)$ for two extreme admissible values of the depletion depth $l = 60$ and 150 nm.

local filling factor $\nu_e(x_{\rm QH}) = \frac{hn_e(x_{\rm QH})}{eB} = \lfloor \nu \rfloor$, i.e., is the largest integer inferior or equal to the bulk filling factor ν. As ν varies, w spans a large range of values from 0 (strongly confined plasmons) to $w \simeq w_0$ (loosely confined plasmons), reflecting the variations of $x_{\rm QH}$, thus yielding an implicit equation connecting w, B, and n_e. Accounting for a broadening w_p of the transverse width due to puddles, we find that the edge profile function $x \mapsto f(x)$ can be reconstructed using the implicit equation (see [30]):

$$f(w - w_p/2) = 1 - \frac{1}{2\nu}. \quad (2)$$

The results are presented in Fig. 4(d). For all triplets (n, B, w), we plot $1 - 1/2\nu$ as function of the measured width w. The data points describe the reconstructed edge profile, which is found to be mostly independent of the bulk density n. We then fit the reconstructed profile with the heuristic function $f(x) = \sqrt{\frac{x}{x+l}}$ used in Ref. [26] to obtain an estimate of the edge depletion length l. We find a good agreement with $l \simeq 60 - 150$ nm, and a puddle broadening $w_p \simeq 1.2$ µm (almost identical for both Sample A and B, see [30] for more data sets). In particular, the depletion occurs on a scale l on the order of 1 to 3 times the distance between the quantum well and the gate d, as anticipated from the electrostatic potential created by the gate. Besides, the different characteristic lengths l and w_0 differ by more than one order of magnitude, while they should both be of the order of d in a clean edge potential.

b. Plasmons at zero magnetic field. We now analyze the measurements at $B = 0$ when the gate voltage V_g is tuned to adjust the Fermi level in the gap of the material. Given the insulating bulk, and the much faster response times of edge states compared to bulk states [17], we argue and assume in the following that the phase response is dominated by edge transport, and that the velocity is that of the QSH edge channels. The following analysis supports this assumption.

The amplitude of the signal is rather weak in this regime, and consequently, the phase measurements are more scattered. Nonetheless, we reliably observe (in all samples and configurations) small velocities (see Fig. 5), on the order of $v = v_{\rm QSH} \simeq 10 \times 10^4$ ms^{-1} for Sample A (2×10^4 ms^{-1} for Sample B). We point out that these values are significantly smaller than those predicted by the band structure [43,44], which is only slightly smaller than the Fermi velocity in the conduction band $v_{\rm F}^{\rm CB} \simeq 1 \times 10^6$ ms^{-1}.

As the gate voltage V_g is driven toward the valence band ($V_g \lesssim -0.75$ V), the velocity is found to increase again ($v \gtrsim 2 \times 10^5$ ms^{-1}). This is in line with the expected Fermi velocity in the valence band $v_{\rm F}^{\rm VB} \simeq 2 \times 10^5$ ms^{-1} (though this estimation is made difficult by the camelback structure of the

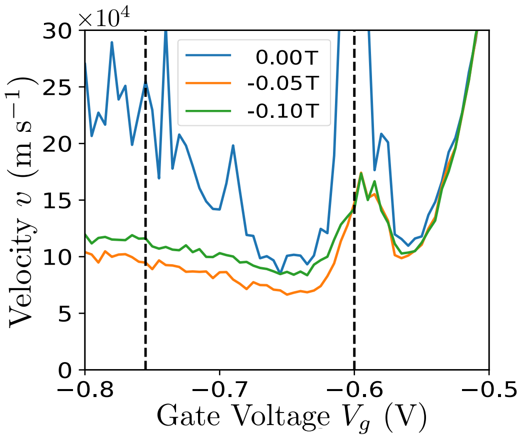

FIG. 5. Velocity measured in the QSH regime at $B \simeq 0$. Velocity v as function of the gate voltage V_g applied to DCg for three values of the magnetic field B close to $B = 0$. The gap region estimated from the resistance R_{2T} is indicated by vertical dashed lines.

valence band, and its strong variations with parameters such as the quantum well thickness t).

IV. DISCUSSION

QH and QSH edge states have different origins, namely the formation of the Landau level spectrum for QH edges states vs the topological band inversion of HgTe for the QSH ones. However, their exact properties could both be affected by electrostatic disorder, and therefore may be correlated to one another. Though the following considerations are more speculative, we put forward examples of such relations.

The velocities in these two regimes can not directly be compared ($v \propto B^{-1}$ in the QH case). However, focusing first on Sample A, we observe that the two capacitances are very close to each other, with [45] $C_{\mathrm{QH}} = \epsilon w_0/d \simeq 1.4\,\mathrm{nFm}^{-1}$ and $C_{\mathrm{QSH}} = 4e^2/hv_{\mathrm{QSH}} \simeq 1.5\,\mathrm{nFm}^{-1}$. This value is also fully compatible with the density of state previously measured in the QSH edge state [17], (with a gold-gated device, and accounting for a factor three in the distance d between the two devices). Moreover, the following ratios between the two samples $\frac{w_0(\mathrm{B})}{w_0(\mathrm{A})} \sim \frac{v_{\mathrm{QSH}}(\mathrm{A})}{v_{\mathrm{QSH}}(\mathrm{B})} \sim 3.8$ further corroborates that large values of w_0, and slow velocities in the QSH regime both originate from the electrostatic disorder and yield $C_{\mathrm{QH}} \simeq C_{\mathrm{QSH}} \simeq 5.4\,\mathrm{nFm}^{-1}$ in Sample B. We note that the characteristic puddle broadening length w_p is observed to be identical in both samples. We stress, however, that the edge profile reconstruction relies on various crude approximations, and probes a high magnetic field regime, while w_0 and v_{QSH} are obtained at low (or zero) magnetic fields.

Strikingly, our measurements suggest the existence of two different length scales differing by one order of magnitude in the QH regime: The confinement length l is of order d while the plasmon width w (determined from the low or high field regime as w_0 or w_p) is one order of magnitude larger. We interpret them in the following manner. Along the edge, the confinement potential is steep, corresponding to large energy scales (order of magnitude of the topological inversion gap ∼300 meV, or of the electron affinity ∼500 meV), resulting in abrupt changes in the electron density from the bulk value $n_e \sim n$ to $n_e \sim 0$ over a length scale $l \sim d$. Atop this sharp edge potential, shallower potential fluctuations generate large puddles of size w, more prominent far from the edge. The visibility of the quantum Hall plateaus results mostly from the disordered potential which pins the Fermi energy on the potential hills near the energy of the Landau level. Therefore, the edge states lie at rather low energies, and their position, as well as the plasmon width, can be strongly influenced by these shallow fluctuations.

These simple comparisons should not be overinterpreted. Nonetheless, all measured quantities reflect shallow fluctuations of the electrostatic potential yielding puddles to which the different types of edge states couple. They could play an important role in understanding the causes of scattering in the edge states. Modeling the edges [46] or solving of the Schrödinger-Poisson [37] equation would also be needed to determine more precisely the potential disorder in the QH or QSH region.

V. SUMMARY AND OUTLOOK

Puddles play a minor part in archetypical studies of the quantum Hall effect in GaAs heterostructures thanks to larger gaps, optimized electrostatic disorder, and the natural protection of QH edge states against scattering [47]. However their role has been recently stressed in the quantum spin Hall effect [17,35,48,49] or the quantum anomalous effect [50], where characteristic energy scales are much smaller.

In this context, our analysis of plasmon velocities in the classical and quantum Hall regime ($B \neq 0$) and in the QSH gap ($B = 0$) examines the interplay of puddles with high-frequency edge channel transport in HgTe quantum wells. It consistently points toward the picture of edge states coupled to puddles that form due to electrostatic disorder. Though the steep edge confinement takes place over a distance $l \sim d$, the quantum Hall edge states spread at low fields over a width w_0 of order 1–4 µm. In addition, we find that the velocity v_{QSH} in the QSH edge state regime is strongly reduced compared to the anticipated Fermi velocity of the edge channels, in agreement with recent measurements of the edge density of state in similar quantum wells.

This body of works suggest that before e-e interaction [51] or other mechanisms, puddles play a prominent role in the physics of topological edge states, and constitute a serious hurdle in order to investigate the topological physics of pristine edge states. We hope that the progress in the growth and lithography of existing materials, the development of new platforms with enhanced gaps [52], and lesser electrostatic disorder will help overcome disorder.

The data sets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

ACKNOWLEDGMENTS

The authors warmly thank S. Shamim and N. Kumada for insightful discussions, and W. Beugeling for technical support with $\mathbf{k} \cdot \mathbf{p}$ simulations. This work has been supported by the ERC under Contract No. ERC-2017-StG "CASTLES" and ERC-2017-Adv "4TOPS;" the DFG (SFB 1170 and Leibniz Program); Germany's Excellence Strategy (Cluster of Excellence Matter and Light for Quantum Computing ML4Q, EXC 2004/1 - 390534769, and the Würzburg-Dresden Cluster of Excellence on Complexity and Topology in Quantum Matter, EXC 2147, 39085490) and the Bavarian Ministry of Education (Elitenetzwerk Bayern Graduate school on Topological Insulators, and the Institute for Topological Insulators); and finally by the JST, PRESTO Grant No. JPMJPR20L2, Japan.

A.G. and E.F. performed the measurements and the data analysis under the supervision of H.K. and E.B. A.G. fabricated the samples with help from H.K., and based on MBE layers grown by L.L., H.K., and E.B., supervised the project. All authors participated in the analysis of the results and to the writing of the manuscript.

[1] M. König, S. Wiedmann, C. Brune, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science **318**, 766 (2007).

[2] X. L. Qi and S. C. Zhang, Rev. Mod. Phys. **83**, 1057 (2011).

[3] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Science **314**, 1757 (2006).

[4] D. Pesin and A. H. MacDonald, Nat. Mater. **11**, 409 (2012).

[5] J. Alicea, Rep. Prog. Phys. **75**, 076501 (2012).

[6] A. Roth, C. Brune, H. Buhmann, L. W. Molenkamp, J. Maciejko, X.-L. Qi, and S.-C. Zhang, Science **325**, 294 (2009).

[7] C. Brüne, A. Roth, H. Buhmann, E. M. Hankiewicz, L. W. Molenkamp, J. Maciejko, X.-L. Qi, and S.-C. Zhang, Nat. Phys. **8**, 485 (2012).

[8] M. R. Calvo, F. de Juan, R. Ilan, E. J. Fox, A. J. Bestwick, M. Mühlbauer, J. Wang, C. Ames, P. Leubner, C. Brüne, S. C. Zhang, H. Buhmann, L. W. Molenkamp, and D. Goldhaber-Gordon, Phys. Rev. Lett. **119**, 226401 (2017).

[9] K. Bendias, S. Shamim, O. Herrmann, A. Budewitz, P. Shekhar, P. Leubner, J. Kleinlein, E. Bocquillon, H. Buhmann, and L. W. Molenkamp, Nano Lett. **18**, 4831 (2018).

[10] E. Bocquillon, R. S. Deacon, J. Wiedenmann, P. Leubner, T. M. Klapwijk, C. Brüne, K. Ishibashi, H. Buhmann, and L. W. Molenkamp, Nat. Nanotechnol. **12**, 137 (2017).

[11] R. S. Deacon, J. Wiedenmann, E. Bocquillon, F. Domínguez, T. M. Klapwijk, P. Leubner, C. Brüne, E. M. Hankiewicz, S. Tarucha, K. Ishibashi, H. Buhmann, and L. W. Molenkamp, Phys. Rev. X **7**, 021011 (2017).

[12] I. Knez, R.-R. Du, and G. Sullivan, Phys. Rev. Lett. **107**, 136603 (2011).

[13] T. Akiho, F. Couedo, H. Irie, K. Suzuki, K. Onomitsu, and K. Muraki, Appl. Phys. Lett. **109**, 192105 (2016).

[14] H. Irie, T. Akiho, F. Couedo, K. Suzuki, K. Onomitsu, and K. Muraki, Phys. Rev. Mater. **4**, 104201 (2020).

[15] F. Reis, G. Li, L. Dudy, M. Bauernfeind, S. Glass, W. Hanke, R. Thomale, J. Schäfer, and R. Claessen, Science **357**, 287 (2017).

[16] S. Wu, V. Fatemi, Q. D. Gibson, K. Watanabe, T. Taniguchi, R. J. Cava, and P. Jarillo-Herrero, Science **359**, 76 (2018).

[17] M. C. Dartiailh, S. Hartinger, A. Gourmelon, K. Bendias, H. Bartolomei, H. Kamata, J.-M. Berroir, G. Feve, B. Plaçais, L. Lunczer, R. Schlereth, H. Buhmann, L. W. Molenkamp, and E. Bocquillon, Phys. Rev. Lett. **124**, 076802 (2020).

[18] E. Pallecchi, A. C. Betz, J. Chaste, G. Fève, B. Huard, T. Kontos, J.-M. Berroir, and B. Plaçais, Phys. Rev. B **83**, 125408 (2011).

[19] A. Inhofer, S. Tchoumakov, B. A. Assaf, G. Fève, J. M. Berroir, V. Jouffrey, D. Carpentier, M. O. Goerbig, B. Plaçais, K. Bendias, D. M. Mahler, E. Bocquillon, R. Schlereth, C. Brüne, H. Buhmann, and L. W. Molenkamp, Phys. Rev. B **96**, 195104 (2017).

[20] R. C. Ashoori, H. L. Stormer, L. N. Pfeiffer, K. W. Baldwin, and K. West, Phys. Rev. B **45**, 3894 (1992).

[21] N. B. Zhitenev, R. J. Haug, K. v. Klitzing, and K. Eberl, Phys. Rev. Lett. **71**, 2292 (1993).

[22] G. Sukhodub, F. Hohls, and R. J. Haug, Phys. Rev. Lett. **93**, 196801 (2004).

[23] J. Gabelli, G. Fève, T. Kontos, J.-M. Berroir, B. Plaçais, D. C. Glattli, B. Etienne, Y. Jin, and M. Büttiker, Phys. Rev. Lett. **98**, 166806 (2007).

[24] H. Kamata, T. Ota, K. Muraki, and T. Fujisawa, Phys. Rev. B **81**, 085329 (2010).

[25] M. Hashisaka, H. Kamata, N. Kumada, K. Washio, R. Murata, K. Muraki, and T. Fujisawa, Phys. Rev. B **88**, 235409 (2013).

[26] N. Kumada, H. Kamata, and T. Fujisawa, Phys. Rev. B **84**, 045314 (2011).

[27] E. Bocquillon, V. Freulon, J.-M. Berroir, P. Degiovanni, B. Plaçais, A. Cavanna, Y. Jin, and G. Fève, Nat. Commun. **4**, 1839 (2013).

[28] N. Kumada, P. Roulleau, B. Roche, M. Hashisaka, H. Hibino, I. Petković, and D. C. Glattli, Phys. Rev. Lett. **113**, 266601 (2014).

[29] V. Freulon, A. Marguerite, J.-M. Berroir, B. Plaçais, A. Cavanna, Y. Jin, and G. Fève, Nat. Commun. **6**, 6854 (2015).

[30] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.108.035405 for information about the RF measurement setup and its calibration, additional data sets on Sample B, additional studies of the crossover magnetic field, and frequency-dependent studies. It also contains Refs. [31–33].

[31] B. I. Halperin, Phys. Rev. B **25**, 2185 (1982).

[32] M. Büttiker, Phys. Rev. B **38**, 9375 (1988).

[33] N. H. Tu, M. Hashisaka, T. Ota, Y. Sekine, K. Muraki, T. Fujisawa, and N. Kumada, Solid State Commun. **283**, 32 (2018).

[34] H. Kamata, N. Kumada, M. Hashisaka, K. Muraki, and T. Fujisawa, Nat. Nanotechnol. **9**, 177 (2014).

[35] H. Kamata, H. Irie, N. Kumada, and K. Muraki, Phys. Rev. Res. **4**, 033214 (2022).

[36] D. B. Chklovskii, B. I. Shklovskii, and L. I. Glazman, Phys. Rev. B **46**, 4026 (1992).

[37] P. Armagnat and X. Waintal, J. Phys. Mater. **3**, 02LT01 (2020).

[38] I. L. Aleiner and L. I. Glazman, Phys. Rev. Lett. **72**, 2935 (1994).

[39] M. D. Johnson and G. Vignale, Phys. Rev. B **67**, 205332 (2003).

[40] P. Burke, IEEE Trans. Nanotechnol. **1**, 129 (2002).

[41] C. de C. Chamon and X. G. Wen, Phys. Rev. B **49**, 8227 (1994).

[42] N. Kumada, N.-H. Tu, K.-i. Sasaki, T. Ota, M. Hashisaka, S. Sasaki, K. Onomitsu, and K. Muraki, Phys. Rev. B **101**, 205205 (2020).

[43] S. S. Krishtopenko and F. Teppe, Phys. Rev. B **97**, 165408 (2018).

[44] J. Strunz, J. Wiedenmann, C. Fleckenstein, L. Lunczer, W. Beugeling, V. L. Müller, P. Shekhar, N. Traverso Ziani, S. Shamim, J. Kleinlein, H. Buhmann, B. Trauzettel, and L. W. Molenkamp, Nat. Phys. **16**, 83 (2020).

[45] For $\epsilon \simeq 8.8\epsilon_0$, largely dominated by the contribution of the capping CdTe layer.

[46] T. L. van den Berg, M. R. Calvo, and D. Bercioux, Phys. Rev. Res. **2**, 013171 (2020).

[47] C. J. Lin, M. Hashisaka, T. Akiho, K. Muraki, and T. Fujisawa, Phys. Rev. B **104**, 125304 (2021).

[48] J. I. Väyrynen, M. Goldstein, Y. Gefen, and L. I. Glazman, Phys. Rev. B **90**, 115309 (2014).

[49] S. Shamim, P. Shekhar, W. Beugeling, J. Bottcher, A. Budewitz, J. B. Mayer, L. Lunczer, E. M. Hankiewicz, H. Buhmann, and L. W. Molenkamp, Nat. Commun. **13**, 2682 (2022).

[50] G. Lippertz, A. Bliesener, A. Uday, L. M. C. Pereira, A. A. Taskin, and Y. Ando, Phys. Rev. B **106**, 045419 (2022).

[51] J. Wang, Y. Meir, and Y. Gefen, Phys. Rev. Lett. **118**, 046801 (2017).

[52] H. Deng, Z. Chen, A. Wolos, M. Konczykowski, K. Sobczak, J. Sitnicka, I. V. Fedorchenko, J. Borysiuk, T. Heider, L. Plucinski, K. Park, A. B. Georgescu, J. Cano, and L. Krusin-Elbaum, Nat. Phys. **17**, 36 (2021).