Gate-Controlled Semimetal-Topological Insulator Transition in an InAs/GaSb Heterostructure

Kyoichi Suzuki, Yuichi Harada, Koji Onomitsu, and Koji Muraki
NTT Basic Research Laboratories, NTT Corporation,
3-1 Morinosato-Wakamiya, Atsugi, Kanagawa 243-0198, Japan
(Dated: October 31, 2016)

We report a gate-controlled transition of a semimetallic InAs/GaSb heterostructure to a topological insulator. The transition is induced by decreasing the degree of band inversion with front and back gate voltages. Temperature dependence of the longitudinal resistance peak shows the energy gap opening in the bulk region with increasing gate electric field. The suppression of bulk conduction and the transition to a topological insulator are confirmed by nonlocal resistance measurements using a dual lock-in technique, which allows us to rigorously compare the voltage distribution in the sample for different current paths without the influence of time-dependent resistance fluctuations.

Topological insulators (TIs) have attracted strong interest as a new quantum state of matter categorized neither as a metal nor an insulator. The unique electronic properties of TIs originate from their band structures characterized by conduction-valance band inversion with an energy gap opening in the bulk region. The band inversion results from the interplay between the alignment of bands with different orbital characters and the strong spin-orbit interaction inherent to materials containing heavy elements, such as Bi compounds and HgTe/CdTe. Owing to their topological nature, gapless states with distinct transport characteristics emerge at the surfaces and edges of three- and two-dimensional (2D) TIs, respectively. In a 2D TI, also known as a quantum spin Hall insulator, the edge state comprises counterpropagating channels with opposite spin. Since elastic back scattering is prohibited by time-reversal symmetry, dissipationless spin-polarized transport is expected. The observation of the quantum spin Hall effect in a HgTe/CdTe quantum well has established it as a prototypical system for 2D TIs.

Recently, InAs/GaSb heterostructures, with the so-called “broken-gap” or “type-II” band alignment, have attracted interest as a route to achieving a 2D TI by a combination of materials each having a trivial band structure typical of III-V semiconductors. Different from conventional TIs whose topological properties originate from the band structure of individual materials, the band inversion in InAs/GaSb is provided by the relative band alignment of the two semiconductors, where the InAs conduction-band bottom is located below the GaSb valence-band top. Experiments have so far confirmed the existence of edge channels in InAs/GaSb heterostructures in the band-inverted regime. Additionally, it has been shown in Ref. [13] that the band alignment can be tuned by the thickness w of the InAs layer. This is manifested by the transport behavior’s changing from that of a normal band insulator through that of a TI to semimetallic as w is varied from 10 to 14 nm. However, the major advantage of the InAs/GaSb system—the band alignment can be controlled in situ by gate voltages—has not yet been fully exploited.

In this paper, we study the effects of gate electric field on the transport properties of micrometer-sized dual-gate InAs/GaSb heterostructure device using nonlocal transport measurements. Different from HgTe/CdTe, in which both the electron and hole wave functions are confined to HgTe and the band inversion is determined by the HgTe layer thickness, the electron and hole wave functions in InAs/GaSb are confined to different layers, which allows us to tune the degree of band inversion in situ by applied gate voltages. We demonstrate that an original semimetallic heterostructure can be tuned into a TI using gate voltages. Our results represent an important aspect of the InAs/GaSb system, which is useful for studies of TIs.

Figure 1(a) shows schematically the band diagram of the system we study in this paper. The InAs/GaSb heterostructure is sandwiched between AlxGa1−xSb barriers and has gate electrodes on both sides. We focus on the band-inverted regime, that is, the situation where the thicknesses of the InAs and GaSb layers are such that the bottom of the electron subband in InAs (with energy E_{0}) is located below the top of the hole subband in GaSb (with energy E_{h}) as shown in Fig. 1(c). The electron and hole wave functions, which are predominantly confined to the InAs and GaSb layers, respectively, are hybridized through the heterointerface to produce an anticrossing gap Δ at a finite in-plane wave vector k [Fig. 1(b)]. As a consequence of the inverted band alignment, linearly dispersing bands connecting the upper and lower bands emerge [dotted lines in Fig. 1(b)], which correspond to the edge channels. When the Fermi level is in the energy gap, the system becomes a 2D TI, where the conductivity in the bulk region is expected to vanish and the transport is governed by the edge channels. Experimentally, while the existence of the anticrossing gap has been confirmed by capacitance spectroscopy transport measurements usually indicate a resistance value that remains fairly low when the Fermi level is swept across the expected gap region. One possible origin of this behavior is the semimetallic band structure that results from the anisotropy of the GaSb valence band [Fig. 1(e)], which we will discuss later.
Gate electric field on the dispersion relation in the presence of \( k \) the energy range of the anticrossing gap for each direction in (c) and (d), respectively. The shaded regions indicate valence-band anisotropy. (e) and (f) correspond to the situation in (c) and (d), respectively. (e) and (f) Schematics illustrating the effects of gate electric field on the sample used and (b) dispersion relations in a TI. (c) and (d) Schematics of (a) band-edge profile of the sample used and (b) dispersion relations in a TI. (c) and (d) illustrate the effects of an electric field on the dispersion relation in the presence of valence-band anisotropy. (e) and (f) correspond to the situation in (c) and (d), respectively. The shaded regions indicate the energy range of the anticrossing gap for each \( k \) direction.

Figures 1(d) and (f) illustrate the effects of an electric field on the band structure. Applying a positive voltage on the GaSb side decreases \( E_{e0} \) while a negative voltage on the InAs side increases \( E_{e0} \). This results in a lesser degree of band inversion as measured by \( E_{h0} - E_{e0} \). Consequently, the anticrossing is shifted to a smaller \( k \) and, at the same time, the larger spatial overlap of the electron and hole wave functions leads to an increase in \( \Delta \). As we will show later, both of these effects help to suppress the bulk conductivity and stabilize the TI phase.

The InAs/GaSb heterostructure we used comprises InAs and GaSb layers with nominal thicknesses of 12 and 10 nm, respectively, sandwiched between Al\(_{0.7}\)Ga\(_{0.3}\)Sb barriers [Fig. 2(a)]. The structure was grown by molecular beam epitaxy on an \( n^+ \)-GaAs substrate which serves as a back gate. To bring the Fermi level close to the energy gap region, the upper Al\(_{0.7}\)Ga\(_{0.3}\)Sb barrier is \( \delta \)-doped with \( [\text{Be}] = 5 \times 10^{11} \text{ cm}^{-2} \) at a setback of 5 nm. A small six-terminal device with AuGe Ohmic contacts [Fig. 2(b)] is formed using deep ultraviolet lithography and wet etching. The distance between adjacent contacts is 2 \( \mu \)m. A Ti/Au front gate is evaporated on a 20-nm-thick Al\(_2\)O\(_3\) gate insulating layer deposited by atomic layer deposition.

Transport measurements were performed using a lock-in technique at a temperature of \( T = 0.27 \) K, unless otherwise noted. For nonlocal resistance measurements, alternating currents with different frequencies (19 and 23 Hz) are injected from different contacts [Fig. 2(b)]. For each pair of voltage probes, two lock-in amplifiers are used to monitor the two frequency components. This dual lock-in setup allows us to simultaneously measure nonlocal resistances for different current injection paths. As we will see later, this method is particularly useful for the present purpose: it not only saves measurement time, but also allows for rigorous comparison between results for different current paths without its being influenced by irreproducible time-dependent fluctuation of the sample characteristics.

Figure 3(a) shows longitudinal resistance \( R_{14,23} \) at different back-gate voltages \( (V_{BG}) \) plotted as a function of front-gate voltage \( (V_{FG}) \). The suffixes of \( R_{ij,kl} \) (= \( V_{kl}/I_{ij} \)) indicate the contacts used for driving the current \( (i, j) \) and measuring the resultant voltage \( (k, l) \). Because of the Be doping in our sample, the Fermi level is in the valence band at \( V_{FG} = 0 \) V. With increasing \( V_{FG} \), the Fermi level moves toward the conduction band. The resistance curve shows a peak when the Fermi level passes through the gap region. When the back gate is unbiased \( (V_{BG} = 0 \) V), the peak resistance is low (< 1 k\( \Omega \)). As \( V_{BG} \) is increased, the resistance peak shifts to lower \( V_{FG} \) and grows noticeably, reaching \( \sim 30 \) k\( \Omega \) at \( V_{BG} = 10 \) V. At the same time, spike-like fluctuations superimposed on the peak become conspicuous. These fluctuations are
mostly reproducible for repeated gate sweeps, except for irreproducible slow time-dependent ones.\(^{13}\)

The height of the \(R_{14,23}\) peak deduced by fitting it with a Gaussian is plotted in Fig. 3(b) as a function of \(V_{BG}\). The horizontal dashed line in the figure indicates the value, \(h/2e^2\) (\(\sim 13 \text{ k}\Omega\)), expected from the quantized edge transport for the geometry used (\(h\): Planck’s constant, \(e\): elementary charge). At low \(V_{BG}\), the peak resistance is much lower than this value, indicating that considerable bulk conduction exists even when the Fermi level is in the gap region. As \(V_{BG}\) is increased, the peak resistance increases significantly, exceeding \(h/2e^2\) at \(V_{BG} \sim 7 V\), and then tends to saturate at higher \(V_{BG}\). Although edge transport in a 2D TI should be protected from backscattering, in actual systems the conductance quantization is disrupted by processes that equilibrate the counterpropagating edge channels, as seen from the length dependence of the edge conductance.\(^{8,9,16}\) The most probable cause is the existence of electron and hole puddles coming from the spatial potential fluctuations,\(^{27,28}\) which can terminate the edge transport even in small samples and make the resistance higher than the expected quantized value.\(^{10,13,29,30}\)

![Figure 3](image-url)

**FIG. 3.** (Color online) (a) Longitudinal resistance \((R_{14,23})\) at different back gate voltages \((V_{BG})\) as a function of front gate voltage \((V_{FG})\). Inset: magnified view of the the \(R_{14,23}\) traces in the low-resistance regime. (b) Peak value of \(R_{14,23}\) obtained from Gaussian fitting of the \(R_{14,23}\) traces in (a), plotted as a function of \(V_{BG}\). (c) Temperature dependence of \(R_{14,23}\) peak at different \(V_{BG}\). The data are plotted as a function of inverse temperature \((1/T)\). The short-dashed line is the Arrhenius fit to the data at \(V_{BG} = 10 \text{ V}\), with the slope corresponding to a gap of 0.2 meV.

![Figure 4](image-url)

**FIG. 4.** (Color online) (a)-(d) Nonlocal resistances (a) \(R_{16,34}\), (b) \(R_{16,45}\), (c) \(R_{21,34}\), and (d) \(R_{21,45}\), measured at \(V_{BG} = 0 \text{ and } 10 \text{ V}\) as a function of \(V_{FG}\). (e) and (f) Ratios of nonlocal resistances (voltages) measured with adjacent voltage-probe pairs. Results for different current paths are compared for (c) \(V_{BG} = 10 \text{ and } 0 \text{ V}\). (g) and (h) Equivalent circuits for TI and normal conductor, respectively.

Figure 3(c) shows the temperature dependence of the \(R_{14,23}\) peak, plotted as a function of inverse temperature. At \(V_{BG} \leq 2 \text{ V}\), the peak resistance is almost temperature independent, indicating that the system does not have a full energy gap. At \(V_{BG} = 4 \text{ V}\) the resistance is thermally activated in the high-\(T\) regime, suggesting the opening of an energy gap. However, the temperature dependence saturates at low-\(T\) regime (\(\lesssim 1 \text{ K}\)), and the system remains fairly conductive, with the resistance well below \(h/2e^2\). With increasing \(V_{BG}\), further, the activated behavior in the high-\(T\) regime becomes more pronounced,\(^{31}\) and the resistance value in the low-\(T\) regime increases.

In our previous study, we compared samples with different InAs layer thicknesses \((w = 10, 12, \text{ and } 14 \text{ nm})\) while keeping \(V_{BG} = 0 \text{ V}\).\(^{32}\) The temperature depen-
ence for $V_{BG} = 2$ and 6 V shown in Fig. 3(c) is similar, respectively, to that of the $w = 12$ and 14 nm samples (at $V_{BG} = 0$ V) in the previous report, which we identified as being a TI ($w = 12$ nm) and a semimetal ($w = 14$ nm). As seen in Fig. 3(c), the resistance increases further at higher $V_{BG}$ (≥ 8 V). However, the behavior at $V_{BG} = 8$ and 10 V is distinct from that of the normal band insulator previously reported for the $w = 10$ nm sample (at $V_{BG} = 0$ V), in which the resistance increases to ~$\Omega$. Here, it must be noted that the sample studied here was fabricated from a wafer different from the one used for the $w = 12$ nm sample in the previous study. Comparing the results for the two samples with $w = 12$ nm, we infer that the actual InAs layer thickness is larger in the sample used in this study.

The results in Fig. 3(c) show that the contribution of bulk conduction can be reduced by applied gate voltages. The different contributions of bulk and edge transport at low and high $V_{BG}$ can be further highlighted by nonlocal measurements. Figures 4(a)-(d) display nonlocal resistances measured in four different configurations. The black and red curves in each panel represent the results obtained for $V_{BG} = 0$ and 10 V, respectively. The four traces for each $V_{BG}$ were all measured simultaneously using the dual lock-in technique. The horizontal dashed line in each panel indicate the value, $h/6e^2$, expected for the quantized edge transport in the geometry used for the nonlocal measurements. The nonlocal resistances for $V_{BG} = 0$ V (which are enlarged in the figures by a factor of 20) are far below this value, $h/6e^2$. Similar to the case of longitudinal resistance, the nonlocal resistance peak grows significantly when $V_{BG} = 10$ V is applied, with the peak values exceeding $h/6e^2$.

As shown in Ref. 13, that the transport is dominated by edge channels can be confirmed by examining if the ratio of the nonlocal resistances (or voltages) measured with adjacent voltage-probe pairs depends on the current path. The results for $V_{BG} = 10$ V are shown in Fig. 4(e), where we compare the voltage ratios $V_{23}/V_{34}$ for two current paths $I_{16}$ and $I_{21}$ as a function of $V_{FG}$. In the $V_{FG}$ range corresponding to the peak in the nonlocal resistance ($0.3 < V_{FG} < 0.7$ V), the voltage ratios for the two current paths completely agree with each other. Such agreement has been confirmed for all contact geometries [Fig. 5(d) and Fig. 6 (Appendix)]. As shown in Ref. 13, these results indicate that the current flows only along the edges of the sample, as illustrated in Fig. 4(g). This can be understood by noting that if the current also flows through the bulk region, the voltage distribution in the sample should depend on the current path [Fig. 4(h)]. In fact, the latter behavior is seen in the low- and high-$V_{FG}$ regions of Fig. 4(e), where the voltage ratios no longer agree with each other. This shows that in these $V_{FG}$ ranges the central bulk region becomes conductive, with the majority carrier type’s being $p$ and $n$ on the low- and high-$V_{FG}$ sides, respectively. We note that also in these $V_{FG}$ ranges the voltage ratio depends on $V_{FG}$, indicating that edge transport coexists with bulk conduction.

Similar results were obtained from nonlocal resistance measurements at $V_{BG} = 8$ V.

At $V_{BG} = 0$ V [Fig. 4(f)], the voltage ratios are only a weak function of $V_{FG}$, and different current paths yield different voltage ratios. Similar results were confirmed for other contact geometries. We note that when the bulk transport is dominant the voltage distribution in the sample is determined solely by the contact geometry and does not depend on sample conductivity. Thus, the results in Fig. 4(f) indicate that at $V_{BG} = 0$ V the transport is dominated by bulk conduction even when the Fermi level is in the expected gap region, implying a semimetallic band structure.

We study the evolution from a semimetal to a TI in more detail using nonlocal measurements. Figures 5(a)-(d) display how the voltage ratios $V_{23}/V_{34}$ for two current paths $I_{16}$ and $I_{65}$, measured at (a) $V_{BG} = 0$, (b) 4, (c) 6, (d) 10 V.

Knez et al. also examined the effects of an electric field on residual bulk conductivity $g_{bulk}$ in GaSb/InAs de-
sices and observed that $g_{\text{bulk}}$ decreased with applied gate voltages.\cite{12,32} However, in their undoped samples $g_{\text{bulk}}$ remained significant $(\gtrsim 2e^2/h)$ even when the largest electric field was applied\cite{12} The results were interpreted in terms of the theory in Ref.\cite{33} which states that even weak scattering can destroy the anticrossing gap and result in finite bulk conductivity. For $\Gamma \ll \Delta \ll E_{h0} - E_{e0}$ (where $\Gamma$ is the level broadening), the theory provides a crude estimate of the residual conductivity: $g_{\text{bulk}} \sim (e^2/h)(E_{h0} - E_{e0})/\Delta$. The effect of the electric field is then understood to decrease $(E_{h0} - E_{e0})/\Delta$ However, the theory does not account for the vanishing of bulk conductivity we observed at high $V_{BG}$. This suggests that the assumption $\Gamma \ll \Delta \ll E_{h0} - E_{e0}$ no longer holds in the high-$V_{BG}$ regime of our experiments, where $E_{h0} - E_{e0}$ could be comparable to $\Delta$ depending on the InAs layer thickness.

Another possible mechanism that gives rise to residual bulk conductivity is the anisotropy of the GaSb valence band.\cite{22} In the presence of band anisotropy, the energy position of the anticrossing gap depends on the $k$ direction. If the anisotropy is large as compared to $\Delta$, the anticrossing gaps for different $k$ directions may not overlap in energy, resulting in a semimetallic band structure [Figs. 1(c) and (e)]. The effects of the electric field are to shift the anticrossing to a smaller $k$, where the effect of the GaSb valence band anisotropy is smaller [Fig. 1(d) and (f)], and enlarge $\Delta$.\cite{17} Both effects would work to suppress the bulk conductivity and stabilize the TI phase. Most importantly, our nonlocal measurements clearly show that the bulk conduction vanishes at higher $V_{BG}$ (= 8 and 10 V), demonstrating that an originally semimetallic heterostructure can be tuned into a TI. Although further studies are necessary to clarify the role of disorder\cite{16,34} our results clearly represent an important aspect of the InAs/GaSb system, which will be useful for the study of TIs.

In summary, we have shown that a semimetallic InAs/GaSb heterostructure can be tuned into a TI by applying gate voltages. Nonlocal resistance measurements using a dual lock-in technique enabled us to confirm the suppression of bulk conductance and the transition to a topological insulator without influence from time-dependent resistance fluctuations. Our results warrant further investigations to explore the transition between nontopological and topological insulators.

ACKNOWLEDGMENTS

We sincerely thank Y. Ishikawa and H. Murofushi for sample preparation. This work was supported by JSPS KAKENHI Grant Number 26287068.

Appendix

Figure 6 summarizes the results of nonlocal resistance measurements at $V_{BG} = 10$ V for contact geometries other than those in Figs. 4(e) and 5(d). For all the geometries, the nonlocal resistance ratios for different current paths precisely match each other in the shaded region, demonstrating that the bulk region is insulating and the transport is governed by edge channels in the corresponding $V_{FG}$ range.

FIG. 6. (Color online) Nonlocal resistance measurements at $V_{BG} = 10$ V for various contact geometries shown in the insets. In each panel, nonlocal resistance ratios for different current paths are plotted as a function of $V_{FG}$ ($T = 0.27$ K).
