Gate-tunable high mobility remote-doped InSb/In$_{1-x}$Al$_x$Sb quantum well heterostructures

Wei Yi,1,a Andrey A. Kiselev,1 Jacob Thorp,1 Ramsey Noah,1 Binh-Minh Nguyen,1 Steven Bui,1 Rajesh D. Rajavel,1 Tahir Hussain,1 Mark F. Gyure,1 Philip Kratz,2 Qi Qian,3 Michael J. Manfra,3 Vlad S. Pribiag,4,b Leo P. Kouwenhoven,4 Charles M. Marcus,5 and Marko Sokolich1,a
1HRL Laboratories, 3011 Malibu Canyon Rd, Malibu, California 90265, USA
2Department of Applied Physics, Stanford University, Stanford, California 94305, USA
3Department of Physics and Astronomy, Purdue University, 525 Northwestern Ave., West Lafayette, Indiana 47907, USA
4Kavli Institute of Nanoscience, Delft University of Technology, 2600 GA Delft, The Netherlands
5Center for Quantum Devices, Niels Bohr Institute, University of Copenhagen, 2100 Copenhagen, Denmark

(Received 10 March 2015; accepted 26 March 2015; published online 7 April 2015)

Gate-tunable high-mobility InSb/In$_{1-x}$Al$_x$Sb quantum wells (QWs) grown on GaAs substrates are reported. The QW two-dimensional electron gas (2DEG) channel mobility in excess of 200 000 cm$^2$/V s is measured at T = 1.8 K. In asymmetrically remote-doped samples with an HfO$_2$ gate dielectric formed by atomic layer deposition, parallel conduction is eliminated and complete 2DEG channel depletion is reached with minimal hysteresis in gate bias response of the 2DEG electron density. The integer quantum Hall effect with Landau level filling factor down to 1 is observed. A high-transparency non-alloyed Ohmic contact to the 2DEG with contact resistance below 1 Ω mm is achieved at 1.8 K. © 2015 AIP Publishing LLC.

InSb has the smallest electron effective mass ($m^* = 0.014 m_e$) and the largest known room-temperature bulk electron mobility of 78 000 cm$^2$/V s of any semiconductor, making it appealing as a channel material in heterostructure transistors (HEMT, HBT) that can operate at higher frequency and lower voltage than GaAs and InP based devices.1 The strong spin-orbit interaction and a giant g factor (-51 in bulk InSb) makes InSb an interesting candidate for potential exploitation in spintronics or topological quantum computing using non-Abelian Majorana fermions.2,3 In fact, signatures of Majorana fermions were first reported in solid state systems using a gated InSb nanowire-superconductor hybrid device.5–7 However, integration of bottom-up nanowires remains a difficult task. To scale to more complicated structures for topological qubits, an alternative approach is to fabricate InSb nanowires and “T” junctions by top-down techniques starting from planar epitaxial heterostructures. Electron mobility greater than 100 000 cm$^2$/V s and a corresponding electron mean-free path in the sub-micrometer range are achievable. Although InSb/InAlSb quantum wells (QWs) with two-dimensional electron gas (2DEG) mobility in excess of 200 000 cm$^2$/V s have been developed,8,9 they are often plagued with parallel conduction channels which may be difficult to deplete. Gate tunability of the 2DEG density to full depletion is needed in split-gate quantum structures. However, making gated devices for InSb-based materials has been a major challenge. Schottky barrier gates to such narrow-gap materials are leaky due to the low barrier height.10 For metal-oxide-semiconductor (MOS) structures, previous studies found difficulties in growing high-quality gate dielectrics with appropriate interface properties to achieve depletion without hole accumulation or large hysteresis.11–14 So far only one work claimed complete 2DEG channel depletion on InSb MOS device, with the 2DEG mobility reaching ~70 000 cm$^2$/V s (at a 2DEG density of ~3.3 × 10$^{11}$ cm$^{-2}$). However, no data in the pinch-off regime was reported.14 In this letter, we show that all the aforementioned challenges can be overcome with carefully designed epitaxial growth and device processing.

Previous studies showed that for remote-doped InSb/InAlSb QWs grown on lattice-mismatched GaAs substrates with typical rms surface roughness of the order of a few monolayers, the contribution of interface roughness scattering has a strong dependence on the well width and is negligible for a QW thickness of 30 nm or greater. The low-temperature InSb 2DEG mobility is limited by remote ionized impurity (RI) scattering originating from ionized dopants (e.g., Te or Si) in the δ-doping layer and charged defects in the buffer layer.9 Screening by electrons in both the QW and the δ-doping plane improves the mobility. However, electrons in the δ-doping plane and higher subbands in the QW act as parallel conduction channels and are undesirable for certain device operations that require a single electron channel, e.g., the quantum Hall effect (QHE). Magnetotransport of a Hall device is well suited to diagnose the parallel conduction. In the absence of parallel conduction, a clean QHE with the longitudinal resistance showing zero-resistance states and quantized plateaus in Hall resistance should be observed. Otherwise, the magnetoresistance oscillations are superimposed on a background of parallel conduction and never become zero (see Refs. 15 and 16). Due to the rather low

---

a)Authors to whom correspondence should be addressed. Electronic addresses: wy@hrl.com and MSokolich@hrl.com
b)Current address: School of Physics and Astronomy, University of Minnesota, Minneapolis, Minnesota 55455, USA.
conduction band offsets in InSb/InAlSb QWs, the maximum 2DEG density that can be achieved without introducing parallel conduction is typically less than \(~4 \times 10^{11}\) cm\(^{-2}\).\(^{19}\) Inserting multiple doping layers with a goal of increasing the 2DEG density can produce undesirable parallel conduction channels. We have prepared and measured both symmetrically doped InSb QWs (with n-type doping layers at both above and below the QW) and asymmetrically doped QWs (i.e., with only n-type doping layer(s) above the QW). We found that the 2DEG density in a symmetrically doped QW is nearly doubled as compared with an asymmetrically doped QW. However, the parallel conduction in the bottom doping layer cannot be eliminated without first depleting the 2DEG channel. In contrast, asymmetrically doped samples do not have such issues. Therefore, in this work, we focus on results from asymmetrically doped samples. We report that by optimizing the design, epitaxial growth, and fabrication process, an asymmetrically doped InSb/InAlSb QW can achieve both 2DEG mobility higher than 200 000 cm\(^2/V\) s and a nearly hysteresis-free gate bias response to full depletion.

Fig. 1 shows the schematic diagram of the heterostructure growth and gated Hall device design used in this work. Samples were grown in a Varian Gen II molecular beam epitaxy (MBE) system equipped with valved crackers for As, Sb and effusion cells for Ga, In, and Al. The InSb/InAlSb single QW heterostructure (Fig. 1(a)) was grown on a semi-insulating GaAs (100) substrate, on top of a GaAs buffer, an InSb nucleation layer, and a 4 \(\mu\)m thick strain-relaxed In\(_{1-x}\)Al\(_x\)Sb buffer layer (\(x = 8\%\)). The InSb QW thickness (30nm), the Si \(\delta\)-doping level (5 \(\times\) 10\(^{11}\) cm\(^{-2}\)) and the distance \(d\) to the QW (20nm) were selected after several rounds of test growths and Schrödinger-Poisson (SP) simulations to achieve high 2DEG mobility. The surface depth of the Si \(\delta\)-doping was in the range of 20–50nm (20nm used for the data presented). Transmission electron microscopy (TEM) of the sample cross sections (Fig. S1 in Ref. 15) showed that the defective region extends to about 3 \(\mu\)m into the In\(_{1-x}\)Al\(_x\)Sb buffer layer and the buffer region beyond that thickness has relatively low density of dislocations. To inspect the effect of buffer thickness, we grew several samples with identical layer structures and growth conditions, except that the buffer thickness was changed to either 1 \(\mu\)m or 6 \(\mu\)m. For 1 \(\mu\)m-buffer samples, the low-temperature 2DEG mobilities dropped from \(~200\) 000 cm\(^2/V\) s to \(~60\) 000 cm\(^2/V\) s, whereas for 6 \(\mu\)m-buffer samples the mobility improvement over 4 \(\mu\)m-buffer samples was marginal. The rms surface roughness of the as-grown sample surface is 1.3 nm for 1 \(\mu\)m-buffer samples and 1.9 nm for 4 \(\mu\)m-buffer samples over an area of 10 \(\mu\)m \(\times\) 10 \(\mu\)m.\(^{15}\)

Processing of narrow-gap InSb-based materials poses another major challenge. At elevated process temperatures, several mechanisms can deteriorate the device characteristics, including, but not limited to, interface roughening, preferential evaporation of Sb atoms on exposed surfaces, and diffusion of In and Sb into gate dielectrics.\(^{18}\) As a precaution, we have kept the process thermal budget conservatively at or below 180 °C. For high-frequency transistors and superconductor-semiconductor hybrid quantum devices, a low resistance Ohmic contact to the InSb 2DEG is desired for various reasons. Conventional Ohmic contacts on the top surface by eutectic alloys such as AuGe/Ni or AuGe/Pt\(^{19}\) require a post-deposition anneal, which may cause material damage and surface leakage. Instead, we took the approach of etching into the InSb QW using an Ar ion etch with secondary ion end-point detection followed by non-alloyed Ti/Au evaporation (500 Å Ti/3000 Å Au). The contact resistance is characterized by a transmission line method and is found to be less than 1 \(\Omega\)-mm at both room temperature and low temperatures.\(^{15}\)

For fabricating gated MOS devices, a major challenge is growing high-quality gate dielectrics on (Al)InSb with low gate leakage. It was reported that the typical interface trap density \(D_{it}\) for Al\(_2\)O\(_3\) gate oxide by atomic layer deposition (ALD) was in the range of \(1.1 \times 10^{12}\) cm\(^{-2}\)eV\(^{-1}\) to \(1.7 \times 10^{13}\) cm\(^{-2}\)eV\(^{-1}\), and the gate depletion and hysteresis was improved by replacing the InSb surface layer with a larger bandgap Al\(_{0.1}\)In\(_{0.9}\)Sb.\(^{14}\) ALD HfO\(_2\) was shown to be a superior diffusion barrier than ALD Al\(_2\)O\(_3\) for InSb.\(^{18}\) Another advantage of ALD HfO\(_2\) is its much larger dielectric constant (\(\kappa \approx 18–21\)) than ALD Al\(_2\)O\(_3\) (\(\kappa \approx 7–9\)). We deposited a 20–40nm high-\(\kappa\) ALD HfO\(_2\) thin film as the gate dielectric using remote plasma enhanced ALD (PEALD) at a low substrate temperature of 175 °C. PEALD is expected to reduce impurities and increase the film density as compared with thermal ALD process. A dielectric breakdown field of \(~3\) MV/cm was measured at room temperature.

For magnetotransport characterizations, materials were processed into 6-lead Hall bars using standard photolithography techniques. Hall bar mesas (200 \(\mu\)m long, 20 \(\mu\)m wide) with Hall voltage lead spacing of 100 \(\mu\)m were wet etched using citric-acid based solution (see Refs. 15 and 20). A top view micrograph of an actual device is shown in Fig. 1(b) with a schematic cross section shown in Fig. 1(c). Processed devices were tested at 1.7–1.8 K in a Lakeshore 9709A Hall measurement system equipped with a 9 T magnet using AC lock-in technique (frequency at 9 Hz and time constant at 30 ms). The drain/source of the Hall bar were connected in series with a 25 kΩ load resistor under a 10 mV AC bias, which restricts the channel current to below 400 nA to avoid localized Joule heating. Measured gate leakage was at or below the noise floor of the instrument (~20 pA) for all DC biases applied (up to ±5 V).

The field dependent magnetoresistance \(R_{xx}\) and Hall resistance \(R_{xy}\), allow for extraction of carrier mobility and sheet charge density. The linearity of \(R_{xy}\) vs. B (not shown) suggests that single B field measurement is adequate. Fig. 2
shows the electron density and mobility of the InSb 2DEG as a function of gate bias $V_g$ measured at $B = 0.1$ T (low field). In Fig. 2(a), repeated negative $V_g$ double sweeps (0 V to $-0.3$ V then back to 0 V) were applied to a Hall bar to deplete/replenish the 2DEG channel. A small hysteresis of $\sim 20$ mV was seen in the initial gate bias swing after the Hall device was cooled to 1.8 K with all the electrical contacts grounded (data not shown). In all subsequent sweeps, the hysteresis was reduced to only $\sim 10$ mV and remained the same irrespective of the voltage swing rate. The mobility vs. sheet charge density (Fig. 2(a) inset) follows a linear dependence, which suggests that carrier scattering is dominated by the interplay of remote dopants and background impurities. The 2DEG channel is completely pinched off at a small negative gate bias of $V_p = -0.17$ V ($V_p$ varies slightly from device to device and is in the $-0.1$ V to $-0.25$ V range). The channel pinch-off always occurs abruptly when the 2DEG electron density drops below certain thresholds (2 to $6 \times 10^{10}$ cm$^{-2}$, varies with device), suggesting a percolation type of behavior with a minimum electron density for the metal-insulator transition to occur. No sign of hole accumulation was observed in samples using a top Si $\delta$-doping layer, as shown by SP simulation (see Fig. 3). The slope of $n-V_g$ vs. charge density (channel-partitioned) for the 2DEG channel charge density as a function of the fixed charge density $n_f$ in the doping layer. Shaded region depicts the desired single channel regime. (d) Mobile charge density (channel-partitioned) vs top gate bias. Experimental data for two different devices are shown by solid and dotted lines, respectively. (a) CBD at $V_g = 0$ when only a single channel (the QW first subband) is populated. Mobile charge density is shown as shaded area. Position of the intentional top Si doping ($n_f = 5 \times 10^{11}$ cm$^{-2}$) along with that of the weakly $V_p$-dependent dielectric interface charge density (for $D_e = 7.5 \times 10^{11}$ cm$^{-2}$ eV$^{-1}$, total $n_{\text{surface}} \sim 8 \times 10^{10}$ cm$^{-2}$) and a weak intrinsic bulk doping ($n_{\text{bulk}} = 1 \times 10^{13}$ cm$^{-3}$) are depicted schematically. Envelope function density for electron states at the bottom of the first and second subbands are shown by solid (blue) and dotted (cyan) lines, respectively. (b) CBD at $V_g = 0.15$ V, wherein two channels—one in the QW and one in the doping layer—are populated appreciably. (c) Channel charge density as a function of the fixed charge density $n_f$ in the doping layer. Shaded region depicts the desired single channel regime. (d) Mobile charge density (channel-partitioned) vs top gate bias. Experimental data for two different devices are shown by solid and dotted lines.

The slope of $n-V_g$ can be used to estimate the interface trap density using an equivalent circuit capacitor model of $\frac{dn}{dV_g} = \frac{C_{\text{surf}}}{e[1 + C_{\text{sc}}(C_{\text{SC}} + C_{2D})]/(C_{\text{SC}}C_{2D})]}$ in the absence of parallel conduction and hole accumulation. Here, $C_{\text{sc}} = e^2D_{\text{sc}}$ is the capacitance associated with interface traps. $C_{\text{ox}}$ and $C_{\text{SC}}$ are the capacitance per unit area.
of the HfO₂ (κ = 18.5) and the semiconductor layers above the QW (κ = 16.4). $C_{2D} = e^2m^*/\pi\hbar^2$ is the quantum capacitance of the 2DEG. $C_{tot}$ is the total capacitance of $C_{ss}$, $C_{sc}$, $C_{it}$, and $C_{2D}$. In negative $V_g$ sweeps (Fig. 2(a)), the $D_{it}$ value of $7.5 \times 10^{11}$ cm$^{-2}$ eV$^{-1}$ is estimated from the n-$V_g$ slope of $8.4 \times 10^{11}$ cm$^{-2}$/V. This $D_{it}$ level in depletion-mode is close to a reported mean $D_{it}$ value of $8 \times 10^{11}$ cm$^{-2}$ eV$^{-1}$ in HfO₂/InSb MOS capacitors, and is lower than a reported $D_{it}$ value of $1.1 \times 10^{12}$ cm$^{-2}$ eV$^{-1}$ in Al₂O₃/InSb devices. In positive $V_g$ sweeps (Fig. 2(b)), $D_{it}$ levels of $1.5 \times 10^{12}$ cm$^{-2}$ eV$^{-1}$ and $3.3 \times 10^{12}$ cm$^{-2}$ eV$^{-1}$ are estimated from the n-$V_g$ slopes of $7.3 \times 10^{11}$ cm$^{-2}$/V and $5.5 \times 10^{11}$ cm$^{-2}$/V for up and down sweeps, respectively. The large increase in $D_{it}$ under positive gate bias sweeps indicates filling of empty interface traps towards the conduction band as electrons tunnel from the QW into the surface under positive $V_g$. This adds negative interface charges and shifts the 2DEG turn-on threshold higher. It is also noteworthy that our PEALD process consistently produced negative pinch-off thresholds. For devices with pinch-off voltage near $-0.2$ V, the 2DEG mobility already reaches $200,000\text{cm}^2/\text{V s}$ at zero gate bias. Therefore, a positive gate bias sweep is not needed in depletion mode operation and electron injection can be avoided. To overcome the issue of electron injection under positive gate bias, a higher Al-content barrier can be used above the QW to reduce the electron tunneling across the InSb/InAlSb interface.

Better understanding of the self-consistent steady-state charge distribution and formation of the mobile channels in the top-gate biased structure has been achieved by employing SP simulations. Electron spatial quantization and Thomas-Fermi filling of the resulting 2D subbands are described in the effective mass approximation, accounting for the conduction band non-parabolicity (which is prominent in narrow-gap materials like InSb). Material parameters for InAlSb alloy-based heterostructures can be found in Ref. 23. Subplots (a) and (b) of Fig. 3 provide examples of the conduction band profile at two distinct gate biases ($V_g = 0$ and $0.15 \text{V}$ here, with an activated top Si doping level of $n_t = 5 \times 10^{11} \text{cm}^{-2}$), allowing substantial electron density (shown as shaded area) accumulation in only a single conducting channel or two parallel channels, respectively. Electrons in the first channel are always strongly localized in the QW (envelope function density at the bottom of the first subband is shown by the blue line). Simulations suggest that electrons in the emerging second channel experience strong hybridization between states in the doping layer and the QW second subband (cyan line) and thus are expected to have lower mobility. The amount of fixed charge in the doping layer (due to different dose and/or fraction of activated dopants) is not fully known and would require extensive additional characterization. In Fig. 3(c), charge density in the first channel is shown as a function of the top gate bias for a set of $n_t$ values which is treated here as a free input parameter. Its main effect is a parallel shift of gate biases that would be required to bring the structure to a similar charge distribution (even identical when no mobile charges are present above the QW). Kinks mark sequential activation of additional channels. Dashes mark the approximate low threshold for metal insulator transition.

Results of the simulations at $n_t = 5 \times 10^{11} \text{cm}^{-2}$ are compared with experimental n-$V_g$ Hall data for two devices in Fig. 3(d), where electron density is explicitly partitioned between channels. Below $V_g \approx -0.05 \text{V}$, only a single conduction channel is present and correspondence between simulation and experiment is good all the way down to the channel pinch-off observed at $V_{2DEG}$ of a few $10^{10} \text{cm}^{-2}$.

After establishing gate tunability of the 2DEG density to full depletion with minimal hysteresis, we performed Hall measurements at high B field to examine the existence of a parasitic conduction channel, which is a common issue for InSb based narrow-gap heterostructures. High B field sweeps manifest no sign of parallel conduction between the 2DEG channel pinch off and the mobility peak (at $V_g = -0.08 \text{V}$) resulting in a clearly observable QHE. The Hall resistance shows well-defined quantized plateaus with a Landau level filling factor down to 1 and correspondingly dissipation-less zero-resistance edge states in magnetoresistance, as shown in Figs. 4(a)–4(d). For most of the QHE regime, the low-field mobility of 2DEG channel is greater than $100,000\text{cm}^2/\text{V s}$ (green bars in Figs. 4(c) and 4(d)).

In conclusion, we have demonstrated efficient gate control and full depletion of high mobility InSb/InAlSb quantum
wells grown on GaAs substrates. A high-$\kappa$ PEALD HfO$_2$ dielectric forms an excellent interface with InAlSb and delivers nearly hysteresis-free gate response at very low operating voltages. Combined with a 2DEG mobility greater than 100 000 cm$^2$/V s in most of the QHE regime and a low Ohmic contact resistance of less than 1 $\Omega$·mm, this development facilitates fabricating top-down InSb-based nanostructures suitable for spin and Majorana fermion based quantum computation schemes.

We would like to acknowledge Gunjana Sharma, Jack A. Crowell, Elias A. Flores, Ivan Alvarado-Rodriguez, Marcel M. Musni, Kasey C. Fisher, John S. Yeah, Adele E. Schmitz, Helen K. Fung, Fiona C. Ku, and David H. Chow at HRL Laboratories and Morten Kjaergaard, Henri Suominen, and Fabrizio Nichele at Center for Quantum Devices, University of Copenhagen, for technical support and fruitful discussions. The authors are grateful for support from Microsoft Station Q.

15See supplementary material at http://dx.doi.org/10.1063/1.4917027 for details.