Transition from one- to two-subband occupancy in the 2DEG of back-gated modulation-doped GaAs-$Al_xGa_{1-x}As$ heterostructures

A. R. Hamilton, E. H. Linfield, M. J. Kelly,* D. A. Ritchie, G. A. C. Jones, and M. Pepper
Cavendish Laboratory, University of Cambridge, Cambridge CB3 0HE, United Kingdom
(Received 31 October 1994)

We present experimental investigations of the transition from one- to two-subband occupancy in the two-dimensional electron gas (2DEG) of back-gated modulation-doped GaAs-$Al_xGa_{1-x}As$ heterostructures. A combination of front and back gates allows us to control the subband energies in the 2DEG, so that we are able to maintain single-subband occupation for carrier densities as high as 9×10^{11} cm^{-2} and achieve double-subband occupancy for carrier densities as low as 5×10^{11} cm^{-2}. These devices are used to map out the phase diagram of subband occupancy as a function of total carrier density and back-gate bias. Our results are in good agreement with the predictions of a recent theoretical model of the electronic structure of back-gated heterostructures.

I. INTRODUCTION

Much of the work on GaAs-$Al_xGa_{1-x}As$ modulation-doped heterostructures has concentrated on the extreme quantum limit in which only the lowest bound state is occupied and the electron gas is truly two dimensional. Recently there has been a resurgence of interest in the properties of high carrier density two-dimensional electron gas (2DEG) systems, looking at the manner in which the second subband is populated,[1] the relative mobilities and scattering times of electrons in the two subbands,[2,3] and the nature of the intersubband scattering.[4]

In this paper we undertake transport investigations to probe directly the electronic structure of the 2DEG. A combination of front and back gates allows us to control the electrostatic potential confining the electron gas, and thereby the shape of the electron wave functions and the energy separation between the 2D subbands. We are thus able to study the conditions under which two-subband occupation can occur, looking in particular at the transition from one- to two-subband occupancy. We map out an experimental phase diagram for this transition, as a function of the total carrier density and back-gate electric field. Our results are in good agreement with the predictions of a recent variational theory.[5]

II. EXPERIMENTAL DETAILS

The experimental results presented here were obtained using samples from two molecular-beam-epitaxy- (MBE) grown wafers, *C*582 and *C*665. Only a brief summary of the MBE growth and sample preparation is given here, as full details have been published elsewhere.[6,7] The wafers share a common upper heterostructure, but have slightly different back-gate configurations (see Fig. 1 and its caption). Independent contact to the 2DEG and back gate was achieved with a combination of *in situ* ion-beam lithography and subsequent MBE regrowth. The MBE growth was interrupted after depositing the n^+ back-gate layer, and regions of this layer were rendered insulating by focused ion-beam damage. These regions were arranged such that Ohmic contacts placed on the subsequently grown heterostructure lay directly above them, thus making contact to the 2DEG without shorting to the n^+ back gate. After growth, a Hall-bar mesa was defined with wet-etching techniques, and annealed AuGeNi was used to make contact to the 2DEG and the back gate. The final processing stage was to deposit a PdAu Schottky gate onto the front surface of the Hall bar.

This technique of back gating was found to be far more successful[6,7] than those based on shallow Ohmic contacts and a buried n^+ layer,[8] or those using an external metal plate as the back gate.[9] Samples fabricated using these

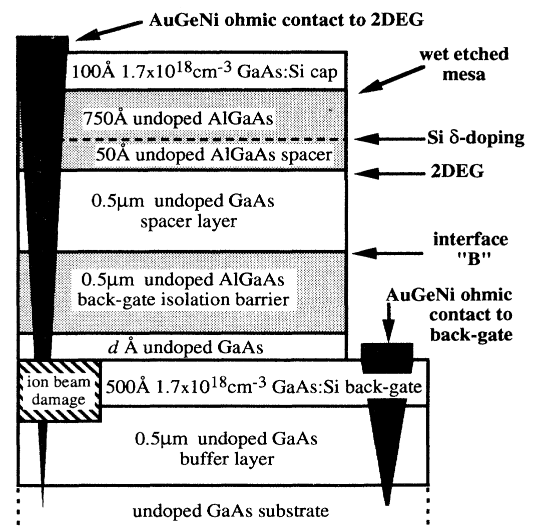

FIG. 1. Cross-sectional diagram of the back-gated heterostructures that were fabricated using the combination of ion-beam lithography and MBE regrowth. The distance *d* was 2500 and 200 Å for wafers *C*582 and *C*665, respectively. The aluminum mole fraction was 0.39 for all $Al_xGa_{1-x}As$ layers.

0163-1829/95/51(24)/17600(5)/$06.00 ©1995 The American Physical Society

alternative methods gave similar results to those presented here, but with a much more limited back-gating action due to the onset of leakage between the back gate and 2DEG.

Measurements were performed at 1.5 K using low-frequency ac lock-in techniques with a measurement current of 100 nA. Gate leakage currents were always less than 1 nA.

III. RESULTS AND DISCUSSION

The one- to two-subband-occupancy transition was studied using front and back gates to induce second-subband occupation for a range of carrier densities. There are several methods of detecting this transition, but for high mobility systems one of the most effective is to monitor the 2DEG mobility as the second subband starts to occupy, since there is then an abrupt decrease in mobility.[10,11] When only one subband is occupied, the mobility is limited by remote ionized impurity scattering and increases monotonically with increasing carrier density.[12] However, when the Fermi energy approaches the bottom of the second subband, an additional (intersubband) scattering channel is opened up, which causes the 2DEG mobility to drop sharply. Subsequently, as the second subband becomes populated, the mobility resumes its monotonic increase with increasing carrier density.

Figure 2 shows a typical set of experimental results, from sample *C665E*. The sample resistivity [Fig. 2(a)] was measured as a function of back-gate bias (V_{bg}) at $B=0$ T, the carrier density n_{Hall} [Fig. 2(b)] was obtained from the Hall resistance at 0.4 T, and the mobility [Fig. 2(c)] was obtained from the ratio of these two measurements. When only one subband is occupied ($V_{bg} < -1$ V), increasing the back-gate bias causes a monotonic increase in n_{Hall} and a steady rise in the mobility. In this regime, the Hall carrier density is equivalent to the total carrier density, and the slope of dn_{Hall}/dV_{bg} can be used to calculate the distance d_{bg} between the back gate and 2DEG using a simple parallel-plate capacitor model. For all samples this was found to be in agreement with the value expected from the growth parameters. When the second subband starts to occupy ($V_{bg} \simeq 0.5$ V), there is a slight change in gradient of the n_{Hall} vs V_{bg} curve, as the Hall carrier density is no longer simply equal to the total carrier density, but depends also upon the mobilities of the two subbands.[13] The transition from one- to two-subband occupation can be seen most clearly in the Hall mobility data in Fig. 2—the point at which it occurs lies between the characteristic maximum and minimum in the mobility.[14,15]

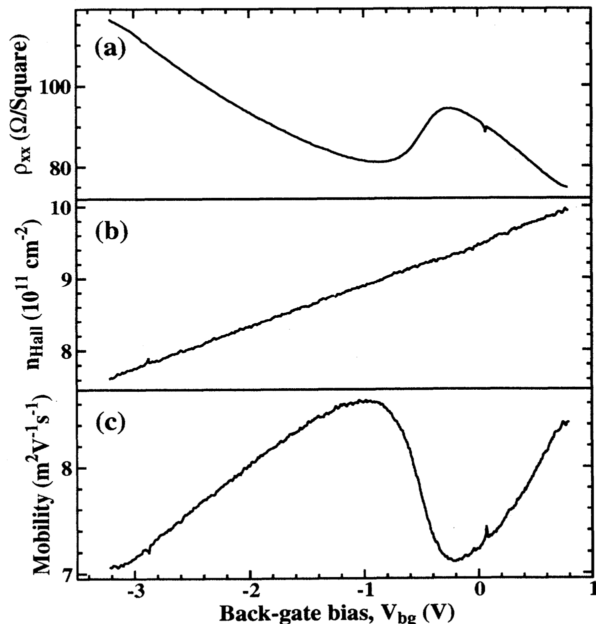

FIG. 2. The back gate induced one- to two-subband-occupation transition in sample *C665E*, with $V_{fg}=0.2$ V at $T=1.5$ K. (a) The resistivity ρ_{xx} vs V_{bg}, (b) the carrier density n_{Hall} vs V_{bg}, (c) the mobility vs V_{bg}.

An alternative method of detecting second-subband occupation is to look for two periods in the Shubnikov–de Haas (SdH) oscillations. Unambiguous observation of two periods in the SdH oscillations, however, requires a rather high carrier density n_2 in the second subband, typically more than 3×10^{10} carriers cm^{-2}. This is because the SdH oscillations cannot be observed below ~ 0.3 T, due to thermal and disorder broadening of the Landau levels, while a large magnetic field will depopulate the Landau levels of the second subband. Thus, second-subband SdH oscillations can only be observed in a limited magnetic-field range, constraining the minimum value of n_2 that can be resolved from the data. From our experience, it is not possible to pinpoint the transition with any greater precision using this technique, and so we do not discuss it any further.

The measurements shown in Fig. 2 were repeated for different front-gate biases to locate the one- to two-subband-occupation transition for a number of carrier densities. The complete data is presented in Fig. 3, showing the variation of mobility (a) with back-gate bias, and (b) with total carrier density. In Fig. 3(a) it can be seen that as the front-gate bias is decreased, the transition to two-subband occupancy occurs at more positive back-gate biases. In Fig. 3(b), each trace corresponds to the back-gate bias being swept from -3 V or less to $+1$ V. Thus, as the carrier concentration decreases, it can be seen that the transition occurs at more positive back-gate biases (i.e., more to the right of each trace).

The data of Fig. 3 also shows that we are able to achieve two-subband occupation at carrier densities as low as 5×10^{11} cm^{-2}, and single-subband occupation at densities as high as 9×10^{11} cm^{-2}. Three-subband occupation was not observed for the range of back-gate biases achievable.

A simple model giving a qualitative understanding of the effect of the two gates on the one- to two-subband transition is illustrated in Fig. 4. The front-gate bias controls the electric field at the heterointerface ($z=0$), and thus primarily affects the carrier density in the 2DEG. The back-gate bias determines the electric field for large z (i.e., $z \gg$2DEG thickness), primarily affecting

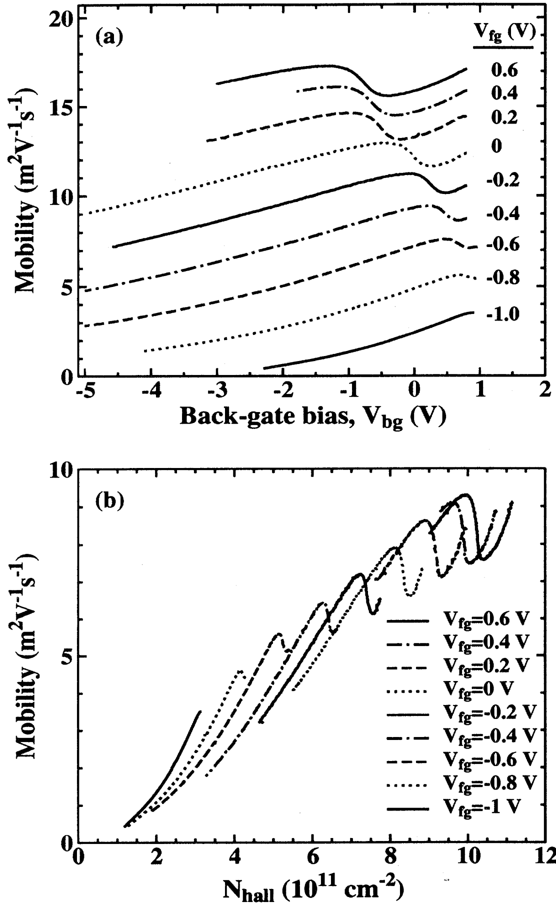

FIG. 3. The effect of front- and back-gates on the one- to two-subband transition measured in sample $C665E$ at $T=1.5$ K. (a) The 2DEG mobility vs V_{bg} at different V_{fg}; (b) The same data plotted as a function of n_{Hall}.

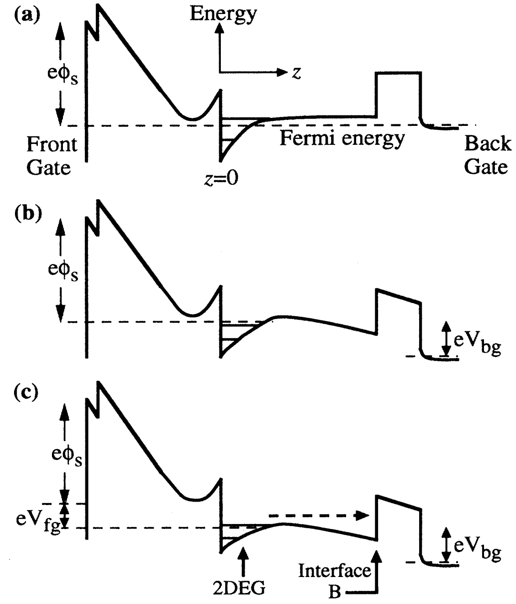

FIG. 4. Schematic diagram showing the effect of the front and back gates on the potential well confining the 2DEG. The pinning of the Fermi energy at the wafer surface is indicated by ϕ_s. The subband energies in the well are shown as horizontal lines. (a) With no applied gate biases, only one subband is occupied. (b) Applying a positive back gate bias widens the quantum well and reduces the subband separation so that two subbands are occupied. (c) Applying a negative front-gate bias reduces the 2DEG density, depopulating the second subband. Interface "B", and the dashed arrow illustrating the field-emission process, are discussed in Sec. III.

the width of the potential well in which the 2DEG is confined.[5] Thus, increasing the positive back-gate bias widens the confining potential well, reducing the subband separation until the second subband drops below the Fermi energy and starts to populate [Fig. 4(b)]. If the front-gate bias is now made more negative, the 2DEG density is reduced, and the height of the Fermi energy above the bottom of the potential well decreases [Fig. 4(c)]. With respect to the Fermi energy, the subbands then rise in energy. Eventually the bottom of the second subband rises above the Fermi energy and the subband is depopulated. In order to repopulate the second subband for this more negative front-gate bias (and hence lower 2DEG density), the back-gate bias must be further increased, reducing the subband separation even more, until the second subband once again drops below the Fermi energy. This simple explanation is in agreement with the data of Fig. 3, where it can be seen that the transition to two-subband occupation occurs at higher back-gate biases for lower carrier densities.

More detailed, quantitative, models of the electronic structure of a back-gated heterostructure have recently been presented by Kelly and Hamilton,[5] and are briefly summarized here. In the first model, Fang-Howard electron envelope wave functions are used for the first and second 2DEG subbands, taking the orthonormalized form

$$\Psi_1(\mathbf{k})=A^{-1/2}2\lambda^{3/2}ze^{-\lambda z}e^{i\mathbf{k}\cdot\mathbf{r}} ,$$

$$\Psi_2(\mathbf{k})=A^{-1/2}\sqrt{12}\lambda^{3/2}z\left[1-\frac{2\lambda z}{3}\right]e^{-\lambda z}e^{i\mathbf{k}\cdot\mathbf{r}} ,$$

where $\mathbf{k}$ is the 2D in-plane wave vector, A is the area of the 2DEG, and λ^{-1} is a measure of the thickness of the 2DEG. The total potential and kinetic energies of the interacting electrons in the 2DEG is expressed within the Hartree approximation in terms of the variational parameter λ and the back-gate voltage, for a fixed carrier density and number of occupied subbands. This total energy is then minimized with respect to λ, and a comparison made of the energies obtained for different subband occupations. Thus, for a given carrier density and biasing arrangement, a unique value of λ is obtained, which gives the electron wave functions, the shape of the electrostatic confining potential, and the subband energies.[16]

The use of a single variational parameter, however, constrains the range of electron subband structures that can be considered in the energy minimization. Therefore, a more general variational approach was introduced in

which separate variational parameters, λ_1 and λ_2, were used for the two subbands. The extra degree of freedom introduced by the use of separate variational parameters allows the envelope wave functions to be altered independently, and so a much wider range of wave functions can be considered in the calculations. The orthonormalized envelope functions now take the form

$$\Psi_1(\mathbf{k}) = A^{-1/2} 2\lambda_1^{3/2} z e^{-\lambda_1 z} e^{i\mathbf{k}\cdot\mathbf{r}} ,$$

$$\Psi_2(\mathbf{k}) = A^{-1/2}\sqrt{12}\lambda_2^{3/2} z \times \left[1 - \frac{(\lambda_1+\lambda_2)z}{3}\right] \frac{e^{-\lambda_2 z} e^{i\mathbf{k}\cdot\mathbf{r}}}{\sqrt{1-\lambda_1/\lambda_2+\lambda_1^2/\lambda_2^2}} .$$

Calculations then proceed as before to obtain unique values of λ_1 and λ_2 that minimize the total energy for a given carrier density and biasing arrangement.

The dashed and solid lines in Fig. 5 show the predicted one- to two-subband-occupancy transition as a function of total carrier density and applied back-gate electric field ($E_{bg} = V_{bg}/d_{bg}$), for the one- and two-parameter variational theories, respectively. Both theories show qualitatively the same features: At high carrier densities, two subbands are occupied, and the transition moves to lower densities with increasing E_{bg}; at some point the transition shows a marked downturn, reaches a maximum E_{bg}, and finally turns over completely, heading back towards the origin (dotted sections of the curves). The behavior of the predicted transitions at high carrier densities can be understood by simple kinetic-energy considerations, as shown in Fig. 4. The "downturn" of the curves as they approach the maximum E_{bg} is due to the increasing importance of the electrostatic Coulomb energy in determining the subband occupancy: for wider wells (i.e., larger E_{bg} or lower n), the spatial overlap between the first- and second-subband wave functions is reduced and occupation of the second subband carries less electrostatic energy penalty. At very low carrier densities, the transitions show a "turnover" (dotted lines). In this region, the second subband is no longer confined in the well and the 2DEG is unstable. This instability arises because at low carrier densities the application of a positive back-gate bias lowers the barrier confining the 2DEG and eventually it becomes electrostatically favorable for the second subband to exist outside the quantum well. The lower the carrier density, the less the electric field required to cause this instability, and so the "transitions" move to lower carrier density with decreasing E_{bg}.

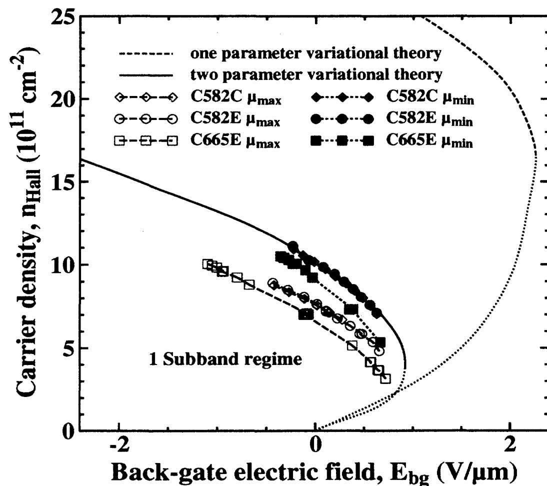

FIG. 5. Phase diagram for one- to two-subband occupation. The dashed and solid lines are the transitions predicted by the one- and two-parameters variational models, respectively, and the dotted lines mark regions in which the 2DEG is predicted to be unstable. The dashed lines with symbols indicate the bounds on the measured transitions for different samples. The actual transition lies between the points obtained from the maximum (solid symbols) and minimum (hollow symbols) in the 2DEG mobility.

The symbols in Fig. 5 mark the phase diagram that we are able to map out experimentally from data such as that shown in Fig. 3. The one- to two-subband-occupancy transition lies between the minimum (hollow symbols) and maximum (solid symbols) in the mobility. It can be seen that the experimental data is far better described by the two-parameter variational theory, indicating that one-parameter variational models are only applicable to single-subband occupancy. However, the agreement between the data and the two-parameter theory is remarkable, especially as there are no adjustable parameters in the model.

We find that samples from the same wafer give identical results (samples $C582C$ and $C528E$), although there is a slight difference between wafers. This discrepancy is thought to be due to differences in the (unintentional) background doping level in the GaAs and $Al_xGa_{1-x}As$ between the back gate and 2DEG. At liquid-helium temperatures, the ionized donors are frozen into place,[17] so the total electric field at the back-gate side of the 2DEG (E_{bg}) is simply the sum of the electric field due to the back gate and a constant electric field due to these fixed charges. Thus, the effect of this background doping is to offset the back-gate bias experienced by the 2DEG, introducing slightly different E_{bg} offsets between wafers. This causes a slight horizontal offset between the measured phase transition, but leaves the slope of the curves unaffected. A modest net concentration of ionized impurities, of order 10^{14} cm^{-3}, is sufficient to account for the offset seen in the experimental data of Fig. 5.

Finally, we turn o the behavior of the transition at low carrier densities and large positive V_{bg}. Some evidence of the theoretically predicted "downturn" of the transition can be seen in the data of Fig. 5, particularly for device $C665E$, where the slope of the measured transition curves increases at low carrier densities. We were also able to observe the predicted instability of the 2DEG at low carrier densities and positive E_{bg}: the application of large positive back-gate biases ($E_{bg} > 1$ V/μm) induced a second 2DEG at the inverted interface marked "B" in Figs. 1 and 4.[6,8] The formation of the second 2DEG was detected by a plateau in the n_{Hall} vs V_{bg} characteristics at $V_{bg} > 1$ V; once the second 2DEG had formed, it screened the original 2DEG from any further increases in V_{bg}.[6,18] The back-gate voltage at which the second 2DEG started to accumulate corresponds to a field of ~ 1

V/μm, which is very close to the electric field at which the two-parameter variational theory predicts that the 2DEG is unstable.

IV. CONCLUSIONS

We have investigated the transition from one- to two-subband occupancy in the 2DEG of a back-gated GaAs-$Al_xGa_{1-x}As$ heterostructure. A combination of front and back gates allows us to control the electronic structure of the 2DEG, so that we are able to maintain single-subband occupation for carrier densities as high as 9×10^{11} cm^{-2}, and achieve two-subband occupancy for carrier densities as low as 5×10^{11} cm^{-2}. This is relevant to the design of high mobility, high carrier-density structures, as intersubband scattering can be minimized.[19]

We have used this flexible biasing arrangement to map out an experimental phase diagram for the one- to two-subband-occupation transition, which we compare with theoretical predictions. We find that this transition moves to more positive back-gate biases as the carrier density is increased, as anticipated by a simple qualitative model. Comparisons with quantitative theories show that a single-parameter variational model is unable to describe our data, implying that such models cannot be applied to a 2DEG with more than one occupied subband. A more general variational theory, however, is found to be in good agreement with the data. In particular it successfully predicts the conditions under which the 2DEG is found to become unstable to electron escape via field emission process.

ACKNOWLEDGMENTS

The authors wish to thank M. P. Grimshaw for MBE wafer growth, J. E. F. Frost for assistance with sample fabrication, and I. M. Castleton for the provision of self-consistent modeling programs. This work has been funded by EPSRC.

---

*Present address: Now at Department of Physics, University of Surrey, Guildford GU2 5XH, United Kingdom.

[1]K. Ensslin, D. Heitmann, R. R. Gerhardts, and K. Ploog, Phys. Rev. B **39**, 12 993 (1989).

[2]T. P. Smith III, F. F. Fang, U. Meirav, and M. Heiblum, Phys. Rev. B **38**, 12 744 (1988).

[3]R. M. Kusters, F. A. Wittekamp, J. Singleton, J. A. A. J. Perenboom, G. A. C. Jones, D. A. Ritchie, J. E. F. Frost, and J. P. André, Phys. Rev. B **46**, 10 207 (1992).

[4]P. T. Coleridge, Semicond. Sci. Technol. **5**, 961 (1991).

[5]M. J. Kelly and A. Hamilton, Semicond. Sci. Technol. **6**, 201 (1991).

[6]E. H. Linfield, G. A. C. Jones, D. A. Ritchie, and J. H. Thompson, Semicond. Sci. Technol. **8**, 415 (1993).

[7]E. H. Linfield, A. R. Hamilton, N. Iredale, D. A. Ritchie, and G. A. C. Jones, J. Vac. Sci. Technol. B **11**, 982 (1993).

[8]D. A. Ritchie, J. E. F. Frost, D. C. Peacock, E. H. Linfield, A. Hamilton, and G. A. C. Jones, J. Cryst. Growth **111**, 300 (1991).

[9]H. Störmer, A. Gossard, and W. Wiegmann, Appl. Phys. Lett. **39**, 493 (1981).

[10]R. Fletcher, E. Zaremba, M. D'Iorio, C. T. Foxon, and J. J. Harris, Phys. Rev. B **38**, 7866 (1988).

[11]D. Cantrell and P. Butcher, J. Phys. C **18**, 5111 (1985).

[12]T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).

[13]M. J. Kane, N. Apsley, D. A. Anderson, L. L. Taylor, and T. Kerr, J. Phys. C **18**, 5629 (1985).

[14]W. de Lange, F. A. P. Blom, P. J. van Hall, P. M. Koenrad, and J. H. Wolter, Physica B **184**, 216 (1993).

[15]The distance between the characteristic maximum and minimum in the mobility is not due to thermal broadening, but is believed to be due to the broadening of the second-subband band edge by the random impurity potential, as proposed in Ref. 10.

[16]Note that the treatment of Ref. 5 was for the limiting case of no background doping in the channel. Although the subband energies are sensitive to the background doping level, numerical calculations show that the subband separation is increased by an amount which is almost independent of the carrier density [F. Stern and S. Das Sarma, Phys. Rev. B **30**, 840 (1982)]. Therefore, the main effect of introducing a finite background doping level will be to alter the carrier density at which second-subband occupation occurs by a constant amount. Effects due to the background dopants between the back gate and 2DEG are addressed later in the paper.

[17]A. Kumar, S. E. Laux, F. Stern, A. Zaslavsky, J. M. Hong, and T. P. Smith III, Phys. Rev. B **48**, 4899 (1993).

[18]Our unique ability to pattern the back gate allows us to deduce that the electrons in this second 2DEG have been emitted from the first 2DEG [dashed arrow in Fig. 4(c)], rather from the Ohmic contacts. This is because there is no back-gate in the vicinity of the Ohmic contacts and hence there is no direct electrical connection between the second 2DEG and the Ohmic contacts.

[19]S. Pearton and N. Shah, in *High Speed Semiconductor Devices*, edited by S. Sze (Wiley, New York, 1990), p. 283.