0.1-μm Gate-Length AlInAs/GaInAs/GaAs MODFET MMIC Process for Applications in High-Speed Wireless Communications

Hans Rohdin
Avelina Nagy
Virginia Robbins
Chung-Yi Su
Arlene S. Wakita
Judith Seeger
Tony Hwang
Patrick Chye
Paul E. Gregory
Sandeep R. Bahl
Forrest G. Kellert
Lawrence G. Studebaker
Donald C. D'Avanzo
Sigurd Johnsen

To ensure high performance of MODFETs used in HP's high-speed communications applications, their high-frequency signal, noise, and power characteristics must be optimized.

HP has developed an ultrafast III-V (AlInAs/GaInAs/GaAs) modulation-doped field-effect transistor (MODFET) technology for use in high-speed monolithic microwave integrated circuits (MMICs). Applications for these devices include:

- Wireless millimeter-wave communications[1]
- Fiber-radio personal communications systems[2]
- Automobile collision-avoidance radar
- Optical fiber and low-noise direct broadcast satellite (DBS) communications receivers.[3]

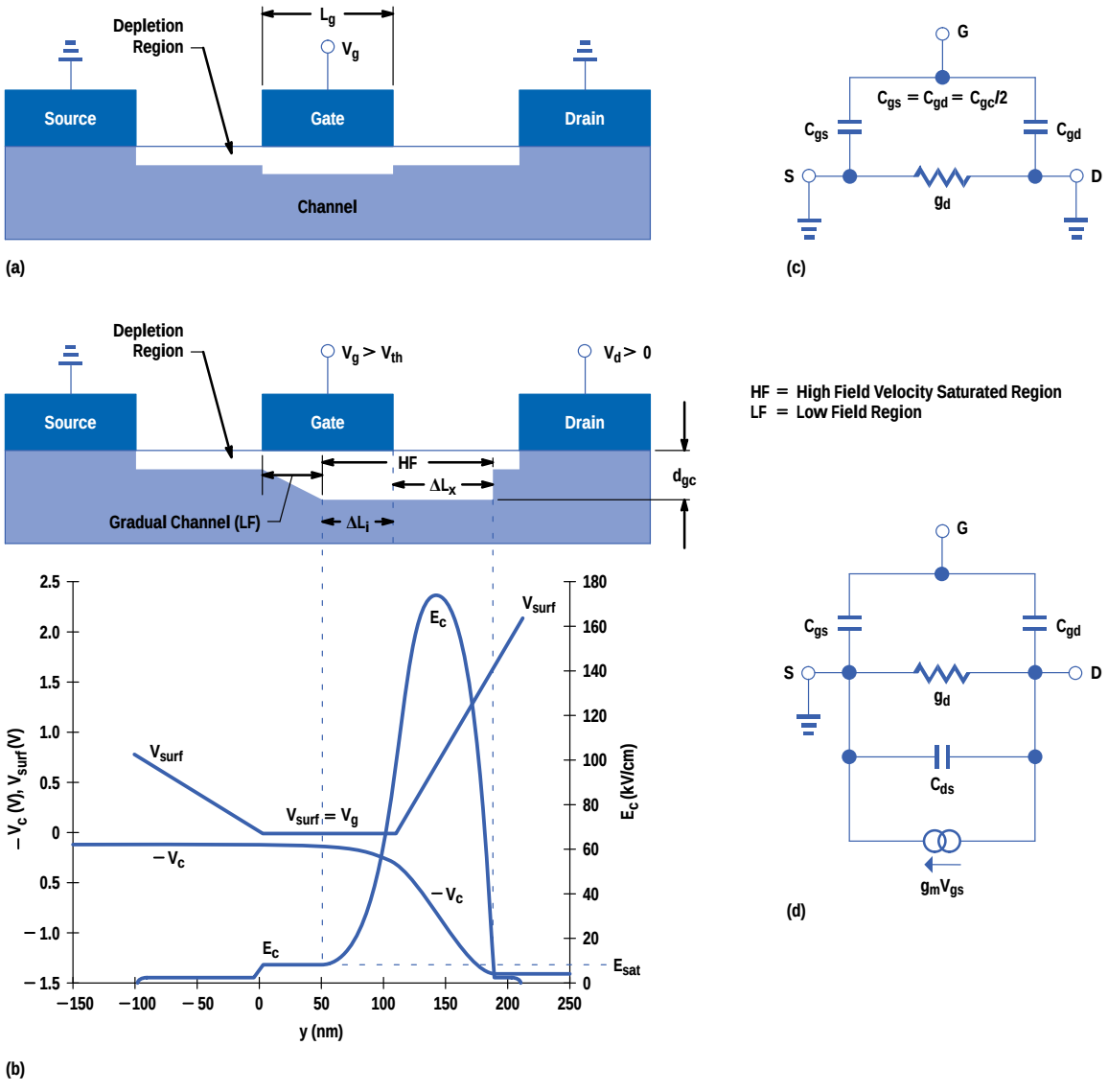
This article provides a summary of the device physics necessary to optimize the high-frequency signal and noise characteristics, semiconductor material, and output-signal power of MODFETs. It also discusses the material, processes, devices, and circuits developed and optimized at HP Laboratories for performance and manufacturability. We conclude with a discussion on process developments at HP's Communication Semiconductor Solutions Division and Microwave Technology Division.

Because of the large number of equations and associated parameters used in this article, ***Appendix A*** provides some brief definitions and typical values for many of the parameters.

Figure 1 shows a high-speed FET in its simplest form: an n-type Schottky-barrier-gate FET, or MESFET (metal-semiconductor FET). A single semiconductor material is used, with the top portion uniformly doped with donors. The source and drain electrodes are nondepleting (tunneling) ohmic contacts, letting electrons in and out with a low resistance R_c. The gate electrode (with gate length L_g in the direction of electron flow and gate width W_g in the perpendicular direction) is rectifying, and for negative to moderately positive gate bias voltage V_g, depletes the semiconductor of electrons to a V_g-dependent depth d_{gc}. The gate therefore determines the number of electrons available for conduction.

Figure 1

An n-type MESFET (a) with zero drain and (b) in saturation. Figures 1c and 1d show equivalent circuits for (a) and (b), respectively.



With the source and drain grounded (**Figure 1a**), the channel is the resistive bottom plate of a parallel-plate capacitance:

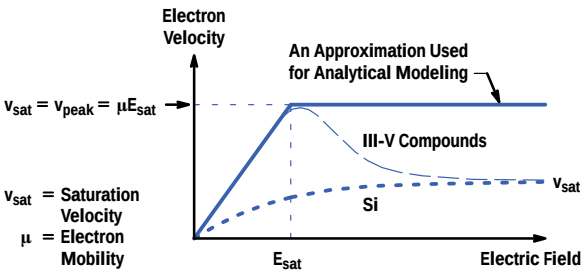
$$C_{gc} = \frac{\varepsilon W_g L_g}{d_{gc}}, \qquad (1)$$

and the top plate is the metal gate. **Figure 1c** shows the associated equivalent circuit for **Figure 1a**. Because of symmetry, we can split C_{gc} into two equal components: one to the source (C_{gs}) and one to the drain (C_{gd}). The gate capacitance C_{gc} and the channel resistance R_{ch} ($= 1/g_d$, where g_d is the output conductance $\partial I_d/\partial V_d$)are functions of V_g. With the source remaining at ground, a positive voltage V_d on the drain (**Figure 1b**) causes a flow of the available electrons from the source to the drain. V_d generates a lateral channel field E_c ($= \partial V_c/\partial y$), which is the driving force for the electrons. Because the surface potential V_{surf} is held constant by the gate while the channel potential V_c increases towards the drain, the channel is increasingly pinched down as the electrons approach the drain-side edge of the gate, more so the higher the drain voltage.

Saturation. Current continuity is maintained under the gate by two mechanisms, which are both related to the increasing lateral channel field E_c. First, the electrons move faster according to the velocity-field curve (see **Figure 2**). Second, because of the field gradient $\partial E_c/\partial y$, the finite thickness of the conducting channel, and Gauss' law, a finite electron concentration exists in the high-field region, preventing pinch-*off*. Beyond the gate, the free-surface effective gating potential V_{surf} is free to increase (see **Figure 1b**), and the pinch-*down* is gradually reduced, which reduces the field.[4] The channel field E_c, which would essentially be constant between the source and drain without a gate, is instead strongly peaked beyond the drain-side edge of the gate. E_c quickly becomes much larger than E_{sat}, and the electrons are velocity saturated as shown in **Figure 2**, leading to current saturation with a much smaller output conductance $g_d = \partial I_d/\partial V_d$.

Figure 2

Qualitative electron velocity-field curves. The solid curve is an approximation that allows analytical modeling.



The high-field region extends towards the source by a distance of ΔL_i and towards the drain by ΔL_x. The partially depleted high-field region $\Delta L_i + \Delta L_x$ separates the source from the drain in an electrostatic sense, and a drain-source capacitance C_{ds} develops as shown in the modified equivalent circuit in **Figure 1d**. Below saturation, the output conductance of the FET appears inductive because of RC and TL (transmission-line) delays in g_d:[5]

$$g_d = g_d(\omega) = \frac{g_{do} e^{-j\omega\tau_{TL}}}{1 + j\omega\tau_{RC}}, \qquad (2)$$

where g_{do} is the low-frequency output conductance. In saturation, C_{ds} overwhelms this inductive appearance. The dependence of C_{ds} on bias cannot be easily estimated analytically. C_{ds} can be quite large once it appears at a lower V_d but it typically drops as V_d (and thus $\Delta L_i + \Delta L_x$) is further increased. Asymmetry has now been introduced, and C_{gs} and C_{gd} are no longer the same. C_{gs} is approximately given by:

$$C_{gs} = \frac{\varepsilon W_g\left[(L_g - \Delta L_i) + \Delta L_i + \Delta L_x/2\right]}{d_{gc}} = \frac{\varepsilon W_g(L_g + \Delta L_x/2)}{d_{gc}}. \quad (3)$$

Equation 3 shows that there are three components that make up C_{gs}. The first, $(L_g - \Delta L_i)$, comes from the low-field resistive source-side part of the channel. This is similar to the parallel-plate capacitance in equation 1. The second component, ΔL_i, and third component, $\Delta L_x/2$, are much less obvious and involve two-dimensional electrostatic considerations.[6] The same is true for C_{gd}, which drops rapidly in saturation to a point where it is more convenient to ignore it as a separate component and lump it with the unavoidable gate fringing capacitance. ΔL_i and ΔL_x depend on drain voltage and are similar in magnitude for typical FET geometries.

The high-field depleted region is indicated in **Figure 1b**. Also shown is the lower-field source-side region under the gate, with its increased depletion towards the drain. This is the so called "gradual channel." There are two additional low-field and partially depleted regions. These are ungated regions between the source and the drain. In this case, a negative charge on the free semiconductor surface does the depleting. For a sufficiently forward-biased gate, these regions (in particular, the source-side region) limit the current that can flow.

Current Modulation. **Figure 1b** also shows the electron potential energy $-V_c$. When the electrons "fall off the potential cliff," they no longer increase their velocity in proportion to the slope. Rather, they quickly reach a saturated velocity v_{sat}, as indicated in the velocity-field curve shown in **Figure 2**. In saturation, the drain current I_d that flows as a result of $V_d > 0$ is modulated primarily by V_g. The rate of this modulation is given by the transconductance $g_m = \partial I_d/\partial V_g$ (**Figure 1d**), which for short gates ($< 1\ \mu m$) in saturation is essentially gate-length independent:

$$g_m = \frac{\varepsilon W_g v_{sat}}{d_{gc}}. \quad (4)$$

The transconductance g_m has the same functional dependence on frequency as g_d in equation 2. However, in saturation, the delays increase because of the time $(\Delta L_x/v_{sat})$ it takes for electrons to transit the external velocity saturated region. For short gates, even with the assumed ideal velocity saturation (**Figure 2**), some output conductance $g_d = \partial I_d/\partial V_d$ will remain, allowing the drain voltage to do some modulation. This drain voltage is undesirable because it degrades the intrinsic low-frequency voltage gain:

$$G_v = \frac{g_m}{g_d}. \quad (5)$$

Unlike transconductance, the output conductance depends strongly on gate length. Empirically, the relationship is inverse, and as shown in the following equation, g_d is reduced in saturation by an empirical factor $k_{g_d} > 0$

$$g_d = g_d^{(sq)} \frac{W_g}{L_g + k_{g_d}\Delta L_x}, \quad (6)$$

where $g_d^{(sq)}$ is a constant of proportionality. Some observations indicate that $g_d^{(sq)}$ is not a constant, but is proportional to d_{gc}.[7]

The parameters g_m, g_d, C_{gs}, C_{gd}, and C_{ds} determine the FET's intrinsic high-speed performance. One figure of merit of particular importance is the cutoff frequency f_T of the current gain, which is the small-signal frequency where the gate current starts exceeding the ac-shorted drain current. This involves only three of the parameters:[6,8]

$$f_T = \frac{g_m}{2\pi(C_{gs} + C_{gd})} \approx \frac{v_{sat}}{2\pi(L_g + \Delta L_x/2)}. \quad (7)$$

$(2\pi f_T)^{-1}$ is the effective electron transit time through the gated channel and the external high-field region. As indicated in equation 5 and discussed below, the other two parameters, g_d and C_{ds}, significantly impact the power gain.

Nonstationary Electron Transport. The small-signal device model described above goes a long way in explaining the essential function of even FETs with very short gates ($<0.1\ \mu m$). This simplifies the optimization of the device. From a practical standpoint, this is convenient. However, it is also quite surprising, since electron transport over short distances in high fields is much more complicated than assumed here. The electron velocity is more correctly described as being tied to the local electron energy rather than to the local field. The net gain in energy over some distance depends on the local electric field and the scattering rate. When optimizing the device, one should, in principle, numerically solve an additional differential equation to account for the fact that it takes time for an electron (even in a constant field) to gain or lose enough energy to reach its steady-state velocity associated with this field. In a FET, as pointed out earlier, the field varies rapidly. Therefore, an electron is unlikely to have its velocity coincide at any point with the steady-state velocity associated with the local field—the velocity is always undershooting or overshooting.

Despite these fundamental considerations, the simple field dependent picture is consistent with a large body of experimental work.[9,10,11] For fields below a critical value of E_{sat} (see **Figure 2**) where the steady-state electron velocity peaks (for III-V materials) or saturates (for silicon), one can assume a linear relationship $v = \mu E$ between the velocity and the local electric field. For larger fields, one can assume for predicting the basic performance of even ultrashort FETs, that the velocity remains equal to an effective saturation velocity independent of any further increase in the field. For III-V materials, the effective saturation velocity is the peak velocity. The fact that the ultimate scattering limited velocity is no larger than in silicon is of little consequence. Thus, velocity overshoot does appear to show up, but the effective saturation velocity does not increase as the gate length is reduced, nor does it exceed the peak velocity by much. The mobility μ is a function of the electron scattering time τ_{scatt} and the electron effective mass m_{eff} in the crystal:

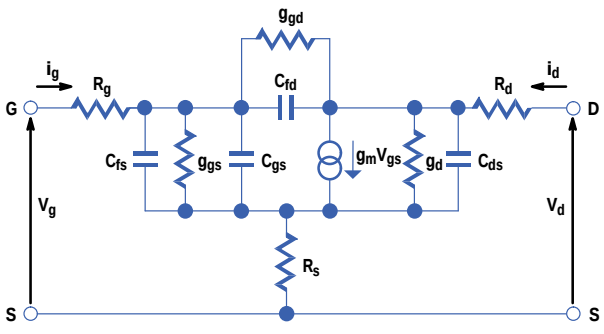
$$\mu = \frac{q\tau_{scatt}}{m_{eff}}. \qquad (8)$$

The effective mass m_{eff} is less than the free electron mass. The peak velocity is more complicated, but correlates with the mobility.[12] Peak electron velocities in III-V materials used for high-speed transistors are two to three times greater than in silicon.

Extrinsic FET Performance. **Figure 3** shows a FET equivalent circuit drawn in a more conventional manner and complete with additional elements. The two intrinsic parameters, g_{gs} and g_{gd}, model the nonzero conductance in a Schottky barrier. There are also unavoidable parasitic elements. A fraction of the source and drain resistances (R_s and R_d) comes from the contact resistance R_c. The remainder is due to channel access resistance. The source-side parasitic gate capacitance C_{fs} arises from the fringing fields between the gate and the nearby metal and conducting semiconductor. Although C_{fd} has the same origin as C_{fs}, it drops significantly in saturation because of the high depleting fields on the drain side. It

Figure 3

Equivalent circuit for the extrinsic FET with parasitic elements.



is virtually indistinguishable from C_{gd} (see **Figure 1d**), and as mentioned earlier, these two capacitances can be lumped into a single feedback/fringe capacitance:

$$C_{fd} = C_{fs}k_{C_{fd}}(\Delta L_x)\ , \tag{9}$$

where $k_{C_{fd}}$ is a function that varies monotonically from 1 in marginal saturation ($\Delta L_x = 0+$) to <1 deeper in saturation ($\Delta L_x > 0$). The extrinsic current-gain cutoff frequency is degraded by these parasitics:[13]

$$f_{Tx} = \frac{g_m}{2\pi(C_{gs} + \Delta C_g)}, \tag{10}$$

where

$$\Delta C_g = C_{fs} + C_{fd} + (R_s + R_d)\left[g_m C_{fd} + g_d(C_{gs} + C_{fs} + C_{fd})\right]. \tag{11}$$

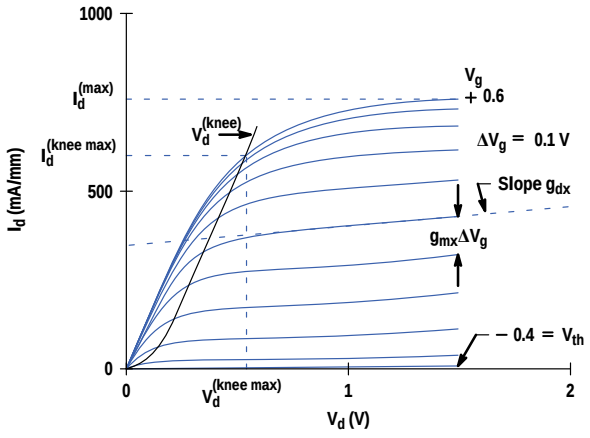
R_s and R_d also degrade the extrinsic transconductance:

$$g_{mx} = \frac{g_m}{1 + R_s g_m + (R_s + R_d)g_d}. \tag{12}$$

The extrinsic output conductance g_{dx} and input capacitance are degraded by the same denominator. The dc characteristics of a FET, showing g_{mx} and g_{dx}, are depicted in **Figure 4**. The FET channel conducts for $V_g > V_{th}$ (threshold voltage), and I_d saturates for $V_d > V_d^{(knee)}(V_g)$. This is the area of interest for small-signal, high-frequency operation. The threshold voltage V_{th} is determined by material structure (discussed later in this article). The knee voltage $V_d^{(knee)}$ should be kept as low as possible and benefits from small $R_s + R_d$ and L_g, and a large μ (mobility).

Figure 4

The dc characteristics of a FET.



y-Parameters and Power Gain. The frequency dependent performance of the FET is best described by the small-signal y-parameters, which relate the ac currents resulting from applied ac voltages (see **Figure 3**) as:

$$i_g = y_{11}V_g + y_{12}V_d \tag{13a}$$

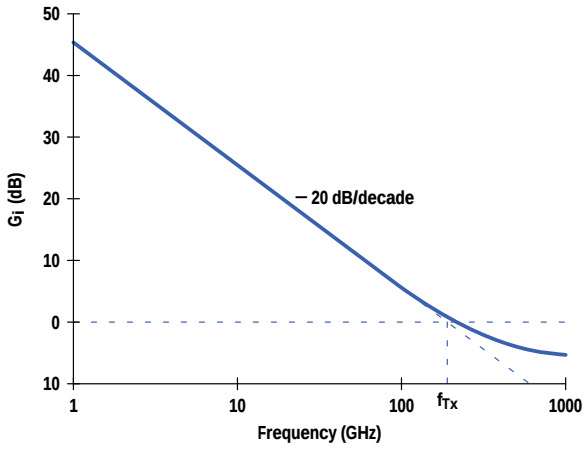
$$i_d = y_{21}V_g + y_{22}V_d\ . \tag{13b}$$

The y-parameters are functions of the equivalent circuit elements and frequency. Equation 10 for f_{Tx} is the result of extrapolating the current gain:

$$G_i = \left|\frac{y_{21}}{y_{11}}\right|^2 \tag{14}$$

Figure 5

Current gain versus frequency (modeled).



at −20 dB/decade to the frequency associated with 0 dB gain. As **Figure 5** shows, G_i does indeed follow this slope in the frequency ranges of interest, and it is an appropriately conservative way of extrapolation.[14] The same cannot be said for the two most commonly quoted power gains: the unilateral power gain[15]

$$G_u = \frac{|y_{21} - y_{12}|^2}{4(\mathrm{Re}(y_{11})\mathrm{Re}(y_{22}) - \mathrm{Re}(y_{12})\mathrm{Re}(y_{21}))} \tag{15}$$

and the maximum available gain[16]

$$G_{ma} = \left|\frac{y_{21}}{y_{12}}\right|\left(k - \sqrt{k^2 - 1}\right), \tag{16}$$

where k is the stability factor[16]

$$k = \frac{2\mathrm{Re}(y_{11})\mathrm{Re}(y_{22}) - \mathrm{Re}(y_{12}y_{21})}{|y_{12}y_{21}|}. \tag{17}$$

For $k > 1$, which occurs for sufficiently large frequencies (> 100 GHz for our technology), the FET is unconditionally stable and will not oscillate with any passive loads. For these high frequencies, equation 16 is valid for the maximum available gain. For most applications (< 100 GHz), the device is unstable but can be stabilized by input or output shunt resistors without affecting y_{21} or y_{12}.[17] Therefore, for $k \leq 1$, the maximum stable gain is:

$$G_{ms} = \left|\frac{y_{21}}{y_{12}}\right|. \tag{18}$$

Unlike the broadband stabilization to achieve G_{ms}, the higher G_u is only achievable with a passive lossless embedment, which unilateralizes the device ($y_{12} \rightarrow 0$) only in a narrow band. For instance, canceling C_{fd} in **Figure 3** would involve an

inductor L in parallel, and the combined admittance would be zero only at the LC resonance frequency $\left(2\pi\sqrt{LC_{fd}}\right)^{-1}$. Therefore, from a practical application standpoint, G_{ms} is usually more meaningful.

Process developers like to quote for G_u and G_{ma} their respective cutoff frequencies f_{max} and f_{mag}. The literature is full of expressions for these parameters, based on −20-dB/decade extrapolation. However, unlike the expression for f_{Tx}, the expressions for f_{max} and f_{mag} are questionable. The reason is clear from **Figure 6** in that neither gain has a −20-dB/decade slope in a significant enough frequency range to warrant such an extrapolation. For example, the extrapolated f_{max} from 50 GHz in **Figure 6** yields $f_{max} = 434$ GHz, while the actual (unmeasurable) value is 214 GHz. Interestingly, the actual values for f_{max} and f_{mag} are identical.[18] This, and the dangers of extrapolating G_u, are further illustrated in **Figure 7**, where C_{ds} has been varied over a wide range. The extrapolated value of f_{max} can reach very impressive values indeed, while the true value is more modest. The huge extrapolated f_{max} for $C_{ds} \approx 625$ fF/mm is due to a resonance in G_u that occurs close to 50 GHz.

Figure 6

Power gains versus frequency (modeled).

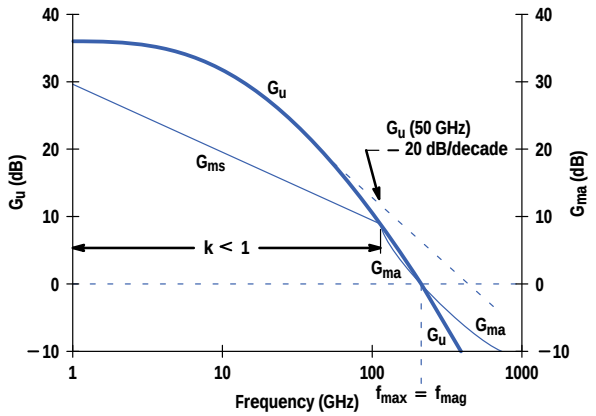
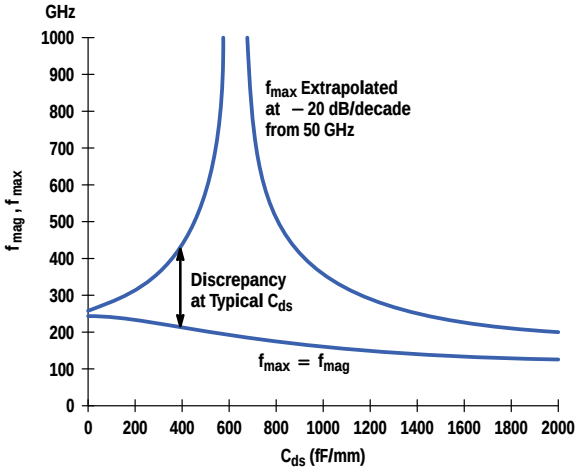


Figure 7

Actual and extrapolated power-gain cutoff frequencies (modeled).



Power gain is more sensitive than current gain to parasitic elements, feedback capacitance, output conductance, and gate leakage. In particular, power gain, unlike current gain, is degraded by the total gate resistance R_g, which is a combination of the metallization resistance along the width W_g of the gate and an interfacial component.[19] At low frequencies, G_u is limited by g_{gs} and g_{gd}.

Minimum Noise Figure. R_g and other resistive components in the equivalent circuit contribute thermal noise, which can get amplified and wind up degrading the noise figure of the FET. Fukui's well-known expression for the minimum noise figure:[20]

$$NF_{min} = 1 + K_f \frac{f}{f_T}\sqrt{g_m(R_g + R_s)} \qquad (19)$$

expresses this fact. K_f has always been considered a material-dependent fitting factor. However, by merging Pospieszalski's more recent noise model for an *intrinsic* FET[21] with equations in the classical treatise by Pucel et al,[22] which explicitly includes R_s and R_g (and is the basis for the Fukui equation), one can tie K_f to physical parameters.

In reference 21, the intrinsic FET's noise is assumed to originate from thermal noise in the output conductance g_d and the C_{gs} charging resistor R_{gs}. R_{gs} is $<20\%$ of the zero-V_d channel resistance R_{ch} $(=1/g_d)$,[5] and should appear in series with C_{gs} in **Figure 3**. T_g and T_d are noise temperatures associated with R_{gs} and g_d respectively. For the FET studied, Pospieszalski explained the noise and its frequency dependence with a value for T_g that was close to the ambient temperature T_a. For T_d, however, values exceeding 3000 Kelvin had to be chosen. Although very large, T_d has the physically appealing feature of being independent of frequency.

Using the model comparisons in reference 21 one can identify the Fukui factor as:

$$K_f = 2\sqrt{\frac{g_d T_d}{g_m T_o}}, \tag{20}$$

where T_o = 290K, the standard noise source temperature. Note that although $g_d \ll g_m$, the empirical fact that $T_d \gg T_o$ means that K_f is significant and that g_d is an important noise source. With g_m tied to the electron saturation velocity v_{sat} (equation 4), it is clear that K_f is indeed affected by material quality. The value of g_d depends on gate length (see equation 6), and thus there is also a geometry dependence in K_f. Further identification with the results in reference 21 shows that $R_{gs}(T_g/T_o)$ should be added to R_g and R_s in the Fukui equation. Since the original Fukui equation was derived under the assumption that $T_a = T_o$, and since $T_g \approx T_a$, one can arrive at the following more general form:

$$NF_{min} = 1 + a + \frac{a^2}{2}, \tag{21}$$

where

$$a = K_f \frac{f}{f_T}\sqrt{g_m(R_{gs} + R_g + R_s)}\,. \tag{22}$$

Equations 20 to 22 extend equation 19 in three respects. First, K_f has been tied to physics. Second, although typically small, R_{gs} has taken its rightful position beside R_g and R_s as a source of noise. Third, the valid frequency range has been extended by the inclusion of the second-order term $a^2/2$. This has its origin in reference 22, but can now (at room temperature) be more conveniently written as simply a higher-order version of the first two terms (1 and a). The expansion suggests an overall exponential dependence $NF_{min} = e^a$. Such an extension, however, appears to overestimate NF_{min}. The generalized room temperature Fukui equation agrees very well with the numerical predictions in reference 21. Like equation 19, equations 20 to 22 have the advantage for device design of explicitly including the important gate and source parasitic resistances.

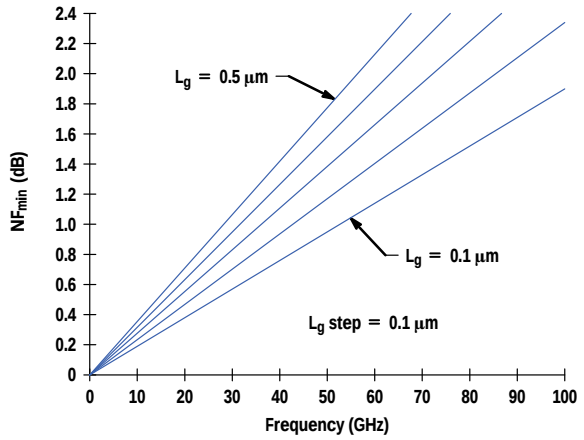
The frequency independence of the noise temperature T_d suggests that it is not merely a fitting parameter but may have physical significance. We considered in a two-step process whether the large values invoked are consistent with electron high-field transport. First, we calculated the values of T_d that produced the K_f values for eight FETs from different labs.[23] The noise temperatures ranged from 1531K to 5629K, with a median value of 3024K and an average of 3157K. Pospieszalski used two values: 3364K and 5514K in this range.

Second, we looked into some of the literature containing numerical FET models that include nonstationary transport. The output conductance g_d in saturation, which is the equivalent circuit element associated with T_d, is determined by the high-field region $\Delta L_i + \Delta L_x$, where the electrons attain their highest energy.[4] We noted this energy for drain biases that correspond to low-noise operation and translated the measurements into electron temperature. The values ranged from 2200K to 4600K, with a median value of 2700K and an average of 3028K.

The correspondence of these two sets of data coming from very different origins appears to support Pospieszalski's tractable thermal-noise model for the intrinsic FET. With $T_d = 3100K$, equations 20 to 22 predict the gate length and frequency dependence of NF_{min} shown in **Figure 8**. As in **Figures 5** to **7**, the other device parameters were chosen to represent our material system and process (described below). We have neglected the effect of the conductances g_{gs} and g_{gd}. For high-speed FETs used at lower frequencies (for instance, in DBS applications), the noise contribution from g_{gs} and g_{gd} can become significant.[24]

Figure 8

Frequency and gate-length dependence of the minimum noise figure (modeled).



Optimized Modulation-Doped AlInAs/GaInAs

From the results of our previous analysis, we now know to look for semiconductor channels with large electron mobility μ, saturation velocity v_{sat}, and full-channel sheet concentration n_{so}. In addition to being beneficial for the intrinsic and extrinsic small-signal performance, these parameters increase the maximum drain current

$$I_d^{(max)} = qW_g n_{so} v_{sat} . \qquad (23)$$

A large $I_d^{(max)}$ is good for driving interconnect capacitances in digital applications. It also increases the signal output power and improves the large-signal linearity of the device.

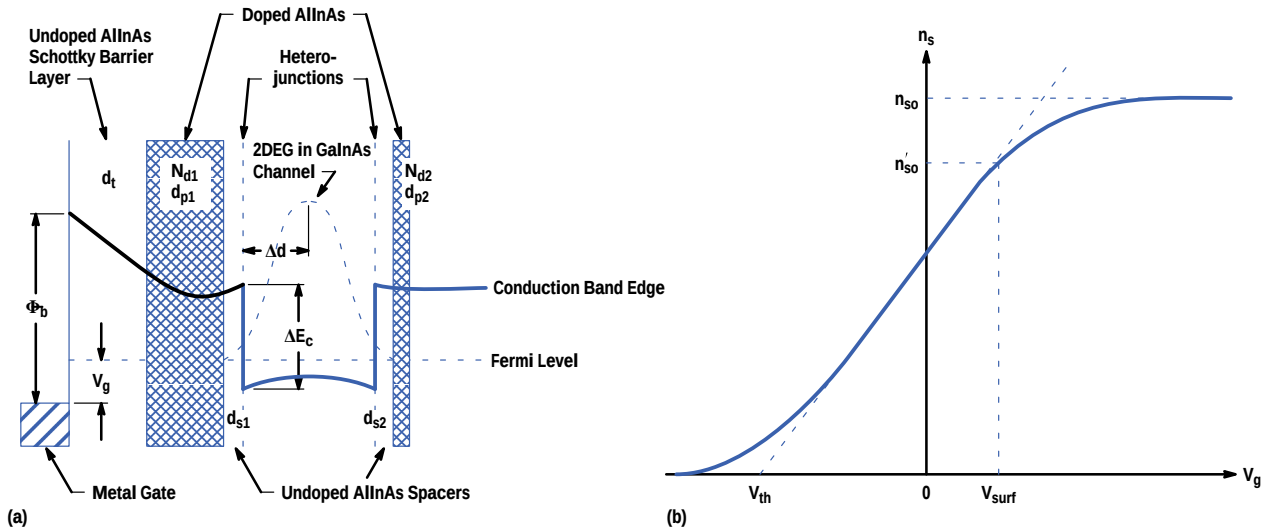
Modulation Doping. The first step taken towards higher μ and v_{sat} in FET materials was to reduce m_{eff} by fanning out in the periodic table from the column-IV element silicon to compounds made of column-III and column-V elements. GaAs was the first to be tried for the fabrication of MESFETs.[25] Later, a method for separating (by "bandgap engineering") the charged electron donors from the mobile electrons was invented. This involved growing the wide-bandgap III-V alloy AlGaAs epitaxially on GaAs and doping only the AlGaAs.[26] The electrons energetically favor the adjacent undoped GaAs where they experience less scattering, which further increases μ and v_{sat}. This method is called *modulation doping*, and FETs made on such epitaxial structures are called MODFETs (or HEMTs, for high-electron-mobility transistors).

The introduction of In in the channel reduces the bandgap and improves n_{so} significantly. It also reduces m_{eff} even further. Lattice mismatch of GaInAs to GaAs prevents In mole fractions larger than approximately 30% on GaAs, unless special growth techniques are employed (discussed later). With an InP substrate, the In mole fractions can approach

100% for sufficiently thin layers.[27] In this material system, the wider-bandgap electron supply layer is AlInAs. The lattice-matched (and most common) In mole fractions for GaInAs and AlInAs on InP are 53% and 52% respectively. FETs in this material system (and variations thereof) have, because of their very high v_{sat} ($\sim 3 \times 10^7$ cm/s) and n_{so} ($\sim 3 \times 10^{12}$ cm^{-2}), shown the best noise and speed performance of any transistor so far.[11,28]

Figure 9

(a) AlInAs/GaInAs MODFET conduction band diagram and (b) 2DEG charge control.



Channel-Charge Modulation. **Figure 9a** shows a conduction band diagram for a double-heterojunction double-side pulse-doped AlInAs/GaInAs MODFET structure, with the important parameters defined. Pulse-doping[29] is particularly important in the AlInAs/GaInAs system because of the rather poor Schottky barrier of AlInAs ($\Phi_b \sim 0.7$ V). An undoped top layer thickness $d_t = 75$ to 100Å is used in our process. The additional pulse below the channel is beneficial in increasing n_{so}.[30] In the vicinity of the optimum operating point ($n_s \sim n_{so}/2$), the electron sheet concentration $n_s(V_g)$ in the channel, where the so-called "two-dimensional electron gas" (2DEG in **Figure 9a**) resides, is approximately linear (see **Figure 9b**):

$$n_s(V_g) = \frac{\varepsilon(V_g - V_{th})}{q d_{gc}}. \qquad (24)$$

The threshold voltage shown in **Figure 9b** is:

$$V_{th} = \Phi_b - \Delta E_c + E_{Fo} - \frac{qN_{d1}d_{p1}(d_t + d_{p1}/2)}{\varepsilon} - \frac{qN_{d2}d_{p2}(d_t + d_{p1} + d_{s1})}{\varepsilon}, \qquad (25)$$

where E_{Fo} is a small constant in the Fermi level and d_{gc} is the effective gate-to-channel distance:

$$d_{gc} = d_t + d_{p1} + d_{s1} + \Delta d. \qquad (26)$$

The backside doping pulse $N_{d2}d_{p2}$ is determined by considerations described below, and the top doping pulse $N_{d1}d_{p1}$ is determined by the desired threshold voltage. The parameters E_{Fo} and Δd result from the Fermi level moving up in the well as it is filled with electrons.[31] A linear relationship can be assumed:

$$E_F(n_s) = E_{Fo} + \left(\frac{q\Delta d}{\varepsilon}\right)n_s . \quad (27)$$

E_{Fo} is typically rather small, and the constant of proportionality between E_F and n_s has been written in a form that includes what turns out to be the effective distance Δd from the top heterojunction to the center of the 2DEG. In a generic MESFET like the one shown in **Figure 1**, the electron concentration in the undepleted channel remains essentially constant. Therefore a MESFET gate modulates the channel width, and consequently d_{gc}. Because of the potential well of the narrowband channel, a MODFET gate modulates the channel electron concentration, while the width of the channel, and consequently d_{gc}, is relatively V_g independent in a rather wide range. This makes the parallel plate analogy discussed earlier even more appropriate for a MODFET. In this sense, a MODFET is similar to a silicon MOSFET. The top wideband electron supply layer plays the role of the gate oxide, but does conduct current as V_g approaches Φ_b. The number of heterojunction interface states is negligible, and this is one of the reasons for the success of MODFETs.[32]

Maximum 2DEG Concentration. For single-side doping, the maximum 2DEG concentration is approximately:[31]

$$n_{so} = \sqrt{N_{dl}^2(d_{s1} + \Delta d)^2 + 2\varepsilon N_{d1}\Delta E_c^{(eff)}/q} - N_{d1}(d_{s1} + \Delta d), \quad (28)$$

$$- N_{d1}(d_{s1} + \Delta d),$$

where the effective conduction-band offset $\Delta E_c^{(eff)}$ is reduced by the donor-binding energy E_d in AlInAs:

$$\Delta E_c^{(eff)} = \Delta E_c - E_d - E_{Fo} . \quad (29)$$

This shows the advantage of a large conduction-band offset, a shallow donor level, and a high doping density. The spacer d_{s1}, inserted to further reduce Coulomb scattering, reduces n_{so}, but can be kept sufficiently small compared to the unavoidable Δd to prevent having a large effect (20Å is a typical value). To calculate the nonlinear approaches of n_s to zero and n_{so} (see **Figure 9b**), a fully numerical self-consistent calculation that solves both Schrödinger's and Poisson's equations is necessary (for instance, see reference 33). Such calculations can also be used to determine Δd and E_{Fo} by comparing the linear part of the $n_s(V_g)$ curve to equations 24 to 26. For d_c ranging from 50 to 200Å, we found that Δd is constant (at 58Å),[34] and that E_{Fo} is proportional to d_c^{-2} (like the energy levels in a square potential).These calculations also show that, to avoid reducing the electron concentration in the high-mobility channel below the value predicted by equation 28, d_c should be $\geq 2\Delta d$. This is not surprising given the interpretation of Δd above. The center of the 2DEG must be spaced from the bottom heterojunction by at least the same amount (Δd) as it is from the top. The quantum-mechanical reason is that for small d_c more of the electron wave function spills over into the wide-bandgap material. For double-sided doping, $d_c = 2\Delta d$ is a good choice. If d_c is much larger, the channel separates into two parallel channels, with nonoptimum composite gate-modulation characteristics and larger output conductance.[4]

Because of the backside doping, the backside of the channel starts filling up with electrons at a more negative gate voltage than indicated by the threshold voltage defined by linear extrapolation from the optimum gate bias (see equation 25). The shift is $-qN_{d2}d_{p2}d_c/\varepsilon$ and is another reason not to use a thicker channel than necessary. With double-side doping, n_{so} will increase, but because of the finite Δd, not by a factor of two. With $N_{d2} = N_{d1}$ and $d_{s2} = d_{s1}$, equation 28 can be used to predict the actual value after substituting $2N_{d1}$ for N_{d1}, $2\Delta d$ for Δd, and $2\Delta E_c^{(eff)}$ for $\Delta E_c^{(eff)}$. If we use the parameter values mentioned above and our typical doping density of $5.5 \times 10^{18}\ cm^{-3}$ and ignore E_d and E_{Fo}, we expect backside doping to increase n_{so} from $3.2 \times 10^{12}\ cm^{-2}$ to $4.4 \times 10^{12}\ cm^{-2}$, a 37% improvement. To avoid an undepleted backside pulse, which could cause pinch-off problems, the backside doping pulse $N_{d2}d_{p2}$ should be kept slightly smaller than the $1.2 \times 10^{12}\ cm^{-2}$ calculated for the maximum possible improvement in n_{so}. This means that d_{p2} in our example should be ≤ 20Å, and the increase in n_{so} will be somewhat less than 37%.

Large-Signal Output Power Limitations

From the numbers established above we can expect $I_d^{(max)}$ in equation 23 to exceed 1.3 A/mm in the AlInAs/GaInAs system. This is indeed possible,[35] but the FET has to be designed with either:

- A very negative threshold voltage
- A reduced distance between the gate and the source
- A gate metal that can be sintered* into the Schottky barrier layer
- A passivation that reduces the free-surface band bending.

The first approach corresponds to shifting the curve in **Figure 9b** to the left such that there is little difference between n_{so} and the lower free-surface-limited n'_{so}. This leads to a threshold voltage that is undesirable for most circuit applications.

The second approach increases the current allowed by the source-side current limiter by reducing its nonlinear resistance. The disadvantage here is that for a typical symmetric gate process, the gate-to-drain distance where ΔL_x would extend is reduced by the same amount. The electric field E_c can then become so large that the breakdown voltage is reduced to unacceptable values by impact ionization in the channel. In fact, the small bandgap ($E_g = 0.77$ eV), which makes GaInAs so attractive for high-speed operation ($E_g \downarrow \Rightarrow \Delta E_c \uparrow$, $m_{eff} \downarrow$), has the drawback of reducing the maximum operating drain voltage. Since the high field on the drain side can support a large electron concentration (through Gauss's law and the spreading of the electrons[4]), the free-surface induced current limiting on this side is less severe than on the low-field source side. If asymmetry in the gate lithography is introduced,[36] the trade-off between $I_d^{(max)}$ and the breakdown voltage BV_{ds} can be reduced.

The third approach increases the effective free-surface gating voltage V_{surf} by $q/\varepsilon(N_{d1}d_{p1} + N_{d2}d_{p2})d_{sinter}$, where d_{sinter} is the depth to which the gate metal sinters during a heat treatment. Thus, n'_{so} can approach n_{so} without shifting V_{th} negatively.

The fourth approach also increases V_{surf}, in this case by the same amount that the band bending is reduced. It has been indirectly determined that Si_3N_4 passivation of AlInAs results in a 0.55V surface potential.[37] Comparing this to the 0.7V Schottky barrier (**Figure 9a**) or the similar free-surface potential, suggests a V_{surf} of $+0.15$V. Combining this approach with the previous approach can provide a cumulative improvement in V_{surf}.

In our present process (described below), the trade-off between $I_d^{(max)}$ and BV_{ds} is shown in **Figure 10**. Note that only in the limit of small BV_{ds} does $I_d^{(max)}$ approach its theoretical ideal maximum. $BV_{ds}^{(off)}$ in **Figure 10** is the *off-state* drain-source breakdown voltage[38] when $V_g < V_{th}$ and I_d is small (10 mA/mm ≅ 1% of $I_d^{(max)}$).[19] For transmitters, the maximum saturated output power P_{sat} is an important parameter. To maximize P_{sat}, it is tempting to let $BV_{ds}^{(off)}$ and the knee voltage $V_d^{(knee\ max)}$ determine the load line (see **Figure 4**). This leads to:

$$P_{sat} = \frac{1}{8} I_d^{(knee\ max)} \left(BV_{ds}^{(off)} - V_d^{(knee\ max)} \right). \qquad (30)$$

* Sintered here refers to the gate metal controllably sinking into the Schottky barrier semiconductor layer when it is annealed.

Figure 10

Trade-off between maximum drain current and off-state breakdown voltage.

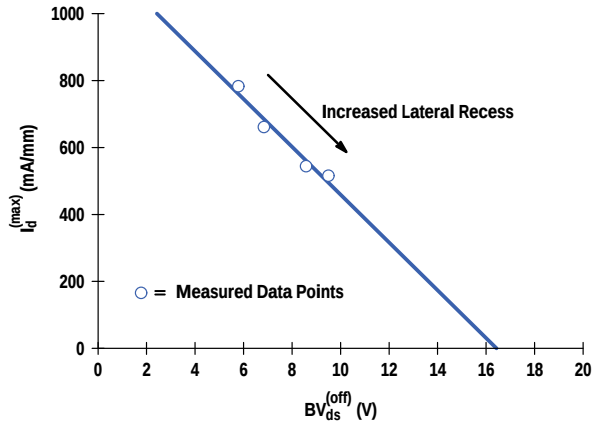
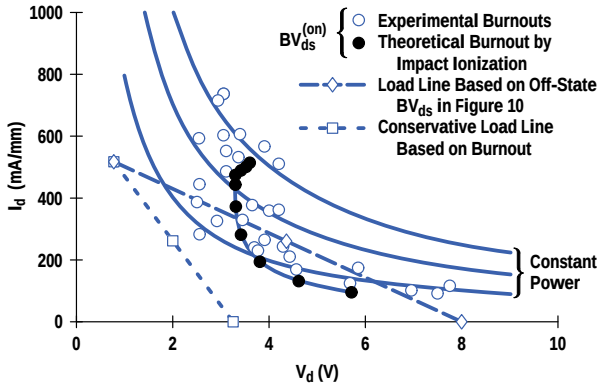


Figure 11

Experimental and theoretical I_d versus V_d burnout points, optimistic and conservative load lines, and constant dc power curves.



However, this approach ignores the fact that the *on-state* breakdown voltage[39] $BV_{ds}^{(on)}$ has a complicated dependence on the drain current, as shown in **Figure 11**. The theoretical $BV_{d}^{(on)}$ curve is the locus of constant $(M-1)I_d$, where M is the carrier multiplication factor calculated from the numerical integral (along the channel) of the impact ionization coefficient α. This value of α increases exponentially with the channel field E_c. E_c, together with I_d, is calculated semi-analytically.[4] $E_c(y)$ in **Figure 1b** was calculated for a 0.11-µm MODFET biased for maximum g_{mx} at $V_d = 1.5V$. As **Figure 11** shows, $BV_{ds}^{(on)}$ drops rapidly as I_d increases from the off state. This is because of an increase in both the current and the channel field. $BV_{ds}^{(on)}$ reaches a minimum at an intermediate I_d close to maximum g_{mx} where the maximum channel field is large. At higher currents the field drops, and since the field is critical in determining the impact ionization, the latter is reduced despite an increasing I_d.

Circuit designers must stay away from the on-state breakdown region with some margin, particularly for the bias point. The primary reason is that on-state breakdown can degrade these FETs. In fact, the open circles in **Figure 11**, which cluster nicely around the theoretical $BV_{ds}^{(on)}$ line, are experimental burnout points, measured here without a padding series resistance, which would have reduced the spread in burnout voltage. The calculated secondary drain current, which we used to define destructive breakdown, was chosen based on the burnout of a FET on an earlier wafer.

Circuit designers are advised not to exceed 2V for $V_d^{(bias)}$, although this may still be too aggressive. We are presently investigating limits set by long-term reliability requirements. For low-noise applications, impact-ionization induced effects are of less concern. The more conservative load line in **Figure 11** leads to a more modest saturated output power expressed in terms of the maximum bias $V_d^{(bias\ max)} < BV_{ds}^{(on\ min)}$:

$$P_{sat} = \frac{1}{4} I_d^{(knee\ max)} \left(V_d^{(bias\ max)} - V_d^{(knee\ max)} \right). \qquad (31)$$

At these lower drain voltages, ΔL_x is kept small and the high-speed properties of the device are maintained.

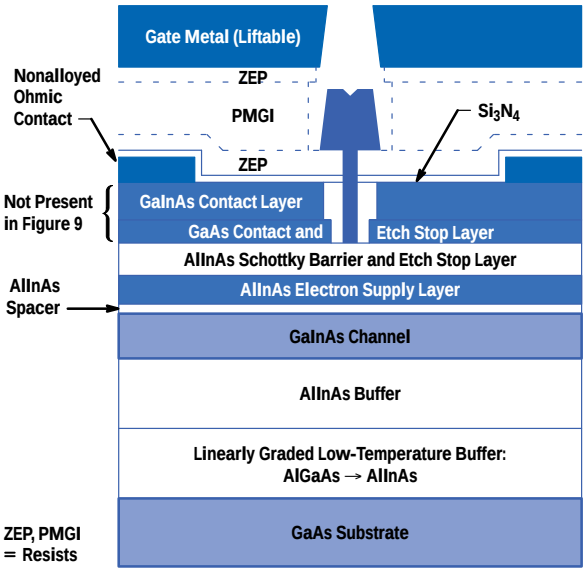
Optimized Process from HP Laboratories

Our AlInAs/GaInAs MODFET process development was guided by the lessons discussed earlier in this article and a concern for manufacturability. Typical FET characteristics are shown in **Figure 4** and listed in Appendix A. Realistic predictions of the frequency dependence of microwave gain and noise figure are shown in **Figures 5, 6** and **8**.

Epitaxial Structure, Ohmic Contacts, and Recess Etching. The epitaxial structure of our basic single-doped FET is shown in **Figure 12**. It is grown by molecular beam epitaxy (MBE). Because the top two layers are not present in **Figure 9** (they are etched away where the gate will be located), they deserve some discussion here. Their task is twofold: to provide for low parasitic resistances R_s and R_d and a well-defined, uniform, and reproducible threshold voltage. The first task is accomplished by the choice of material (GaInAs and GaAs, rather than AlInAs) and by maximum doping (6 to $12 \times 10^{18}\ cm^{-3}$). This results in a low barrier for the electrons, and therefore low tunneling resistance. We get reproducibly low R_c without alloying the contacts. The control and reproducibility of MBE and metal deposition are far better than alloying processes.

Figure 12

Epitaxial structure and gate formation.



To accomplish the second task of a uniform reproducible threshold voltage, selective recess etching is required. A two-step wet recess process has been developed for this purpose. The first solution etches GaInAs, but slows down significantly once the thin GaAs layer has been reached, enough to allow sufficient lateral etch before the GaAs layer is consumed. This lateral etch allows for a significant ΔL_x ($\sim 0.1\ \mu m$, see **Figure 1b**), so that the breakdown voltage is sufficient. The second etch consumes what is left of the GaAs layer and literally stops on the underlying In-containing layer. This is paramount for attaining threshold control.

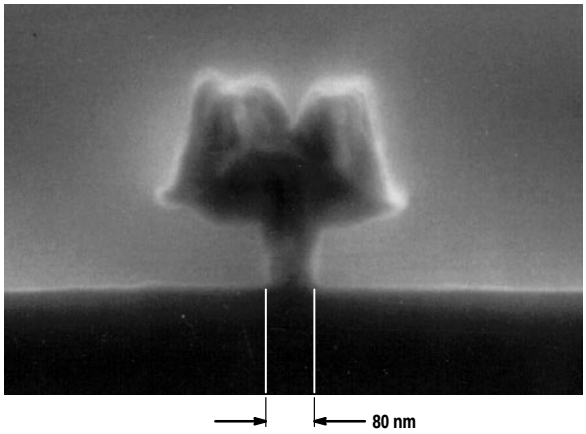
The amount of lateral etch determines where on the $I_d^{(max)} - BV_{ds}^{(off)}$ trade-off line the FETs fall (see **Figure 10**). The amount of lateral etch also affects the $V_d^{(knee)}$ of the FETs, since R_s and R_d have a nonnegligible component associated with these etched higher-resistance regions.[40] Therefore, the lateral etch has an impact on the saturated output power, as

shown in equation 31. There is no lateral-etch stop, but as the etch proceeds, it is not replenished as rapidly as it is depleted. The etch rate is eventually reduced to zero.

0.1-mm T-gates by E-beam Lithography. Several methods for defining ultrashort gates have been developed at HP, based on optical lithography. These use angle evaporation,[9,29] self-limiting oxide spacer self-alignment,[4] or phase-shifting techniques.[42] In the present process, however, the gates are defined by direct e-beam writing in a trilayer resist.[41] This results in a 0.12-μm gate cut in the bottom layer resist after development (see **Figure 12**). Shorter gates can be defined with this approach (see **Figure 13**), but the requirements for high power gain, acceptable breakdown voltage and output power, and reproducibility are met with the present 0.12-μm gates. Because fringing fields modulate the 2DEG, the electrical gate length is approximately ~30 nm longer than the 0.12 μm metallurgical value. Considering this, $g_d^{(sq)}$ in equation 6 (the intrinsic output conductance normalized to a square channel) is 11 μS. The transconductance in ***Appendix A*** corresponds to a saturation velocity (v_{sat}) of 3×10^7 cm/s. As discussed earlier, we consider v_{sat} and $g_d^{(sq)}$ to be nearly fundamental material parameters—essentially independent of gate length—and therefore essential in device optimization.

Figure 13

Sub-0.1-μm T-gate fabricated with direct e-beam writing in trilayer resist.



The opening in the top layer resist (called ZEP) is larger because of a stronger, low-contrast developer. The process produces an ultrashort gate after metal deposition and liftoff, yet has a small gate access resistance along the gate width. The formation of a T-gate (**Figures 12** and **13**) with a single e-beam exposure is made possible by the inertness of each of the two resists to the other's developer.[41]

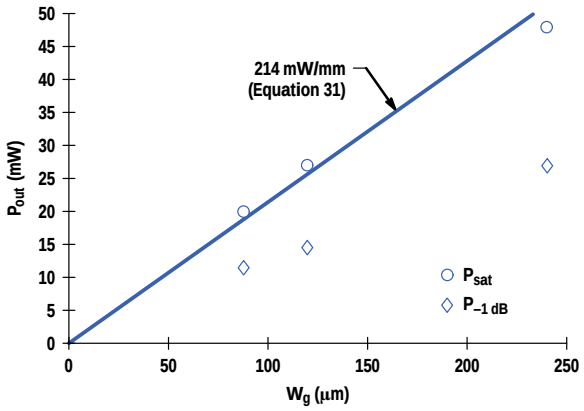
For low fringing capacitance C_{fs} (also after Si_3N_4 overcoat), the wide top of the T is sufficiently spaced (by the bottom ZEP) from the heavily doped contact layers and the source and drain contacts. This results in a very healthy extrinsic current-gain cutoff frequency f_{Tx} (**Figure 5** and ***Appendix A***), which is necessary for millimeter-wave applications. The large aspect ratio $L_g/d_{gc} \approx 5$ results in large g_m/g_d and low C_{fs}/C_{gs}, which promote a high f_{max}. To allow comparison with published values for other high-speed processes, f_{max} in Appendix A is the value extrapolated at −20 dB/decade, despite the cautionary tone in the first section of this article. The real f_{max} is probably closer to the 214 GHz as predicted in **Figure 6**.

Noise and Power. The low noise and high gain of the FETs at low drain bias make DBS (direct broadcast satellite) the primary area of application for this technology. The 12-GHz noise figure quoted in ***Appendix A*** is only 0.07 dB higher than the model predicts (see the section "Minimum Noise Figure"). Part of this discrepancy comes from not including the leakage components g_{gs} and g_{gd} in the model.

As indicated in ***Appendix A***, the FETs can also produce good millimeter-wave output power. Of the methods discussed earlier in "Large-Signal Output Power Limitations" for increasing the drain current with maintained acceptable breakdown voltage, we have presently implemented a Pt-sintered gate process. The saturated output power at 60 GHz with a 2V drain bias scales nicely with gate width (see **Figure 14**) and corresponds closely to a simple dc estimate given in equation 31. The output power at 1-dB gain compression is approximately half of the saturated power.

Figure 14

Optimum MODFET output power versus gate width at V_d = 2V and f = 60 GHz.



Reproducibility and Manufacturability. Given the deep submicrometer dimension of the gate cut, controlling the uniformity and reproducibility of the parameters with wet chemistry is nontrivial, even with the built-in vertical and lateral etch control. The controlled reproducible initiation and quenching of wet etching in small semiconfined openings that have poor aspect ratio require special techniques. The threshold voltage control evident in ***Appendix A*** is based on whole or partial two-inch wafers processed in HP Laboratories' R&D environment. Similar or better uniformity and reproducibility were demonstrated on full three-inch wafers at HP's manufacturing divisions. This similarity indicates that the etching techniques are effective and transferrable.

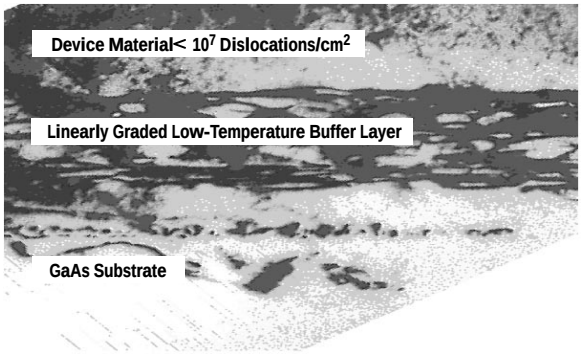
The characteristics summarized in ***Appendix A*** are particularly attractive because they represent a FET process that is designed for manufacturability. The benefits of nonalloyed ohmic contacts, single e-beam exposure, and selective recess etching have already been shown.

As mentioned earlier, the natural lattice-matched substrate for these FETs is InP. With HP's history in GaAs RFIC and MMIC manufacturability,[43] this presents a barrier for transfer to HP's divisions. Relative to GaAs, InP substrates are more expensive and more brittle. This is particularly true for three-inch wafers or larger. Therefore, there is a degree of incompatibility between the existing HP III-V FET/MMIC manufacturing infrastructure and a FET process on InP. This obstacle has been overcome, as shown in **Figure 12**, by a linearly graded low-temperature buffer technology,[44,45] which allows the use of a GaAs substrate for the fabrication of high In-mole-fraction FETs. The lattice constant is varied from GaAs to InP by gradually replacing Ga in AlGaAs with In over 1 µm, so that, at the top of the buffer, AlInAs is being grown. The vast majority of the misfit dislocations generated in this process remains confined to the buffer layer. The

threading dislocation density in the device layers is below the $10^7\ cm^{-2}$ limit observable in TEM (transmission electron microscopy) as shown in **Figure 15**. For comparison, attempts at growing AlInAs directly on GaAs leads to $\sim 10^9\ cm^{-2}$ threading dislocation density. It is possible that the reason for the confinement of the dislocations close to the GaAs substrate is the lower yield strength of material with less In.[46] We have not observed any reduction in FET performance as a result of the switch to GaAs substrate.[19] The approach allows quite a bit of freedom in the choice of In mole fraction. The conduction band offset (ΔE_c) is maximum for about 30% In,[47] which could lead to larger full-channel sheet concentration (n_{so}) and breakdown voltage (BV_{ds}). However, it would also make it harder to achieve good contact resistance (R_c) with nonalloyed ohmic contacts because of the larger band gaps involved.

Figure 15

Transmission electron microscopy (TEM) of an InGaAs/AlInAs MODFET structure grown on a GaAs substrate by linear grading of a 1-µm buffer layer. All observable dislocations are in the buffer layer.



Circuit Results

Several types of circuits have been fabricated with our FET MMIC process. By careful optimization, and using models that have now been extracted, we expect the already good performance to improve even further. **Figure 16a** shows a three-stage broadband MMIC amplifier designed with top-side coplanar transmission lines. The design is based on reference 48. **Figure 16b** shows the gain in V-band (50 to 75 GHz). The gain is 25 dB at 50 GHz and drops 3 dB at 68 GHz.

Given the uncertainty in the number of dislocations that thread through the device layers, we were interested in the yield of a more complex circuit. **Figure 17a** shows a static divide-by-4 circuit with a total of 97 FETs, 49 of which have 0.12-µm gates. The others are used for level shifting and have 0.3-µm gates.

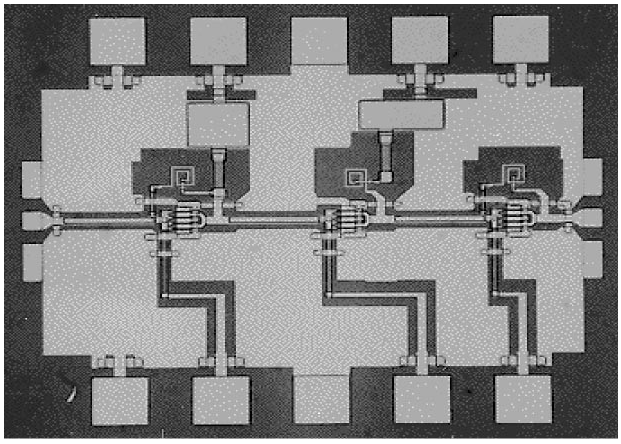
The average yield for four wafers, three on GaAs and one on InP, was 45%. Two of the three GaAs wafers had a higher yield than the InP wafer. In estimating the dislocation density from the yield numbers, we assumed that:

- There is a uniform threading dislocation density.
- The yield loss is due exclusively to dislocations.
- Only 0.12-µm switching FETs are affected by dislocations through the gate area.

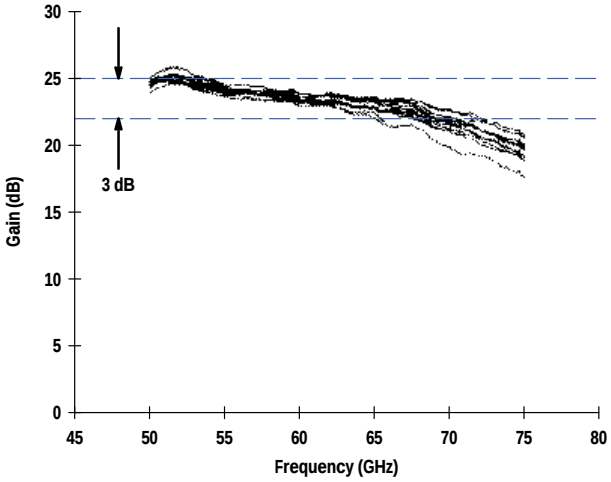
We are aware that the second assumption is not true because other mechanisms, such as gate peeling, are also yield limiters. The third assumption excludes the possibility that 0.12-µm current sources and 0.35-µm level-shifting diodes are also critically affected by dislocations through the gate area. Therefore our $5 \times 10^5\ cm^{-2}$ estimate is an upper limit of an

Figure 16

(a) Three-stage feedback MMIC amplifier, and (b) its V-band frequency response. The traces correspond to different circuit locations on the 2-inch wafer.



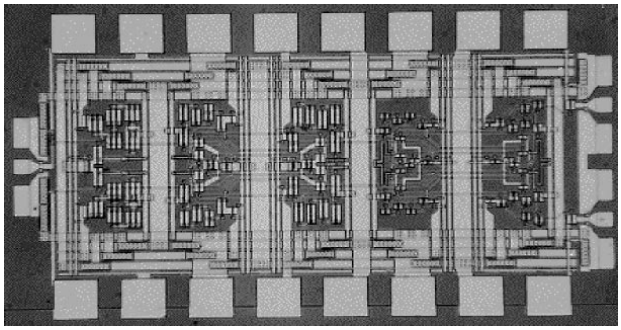
(a)



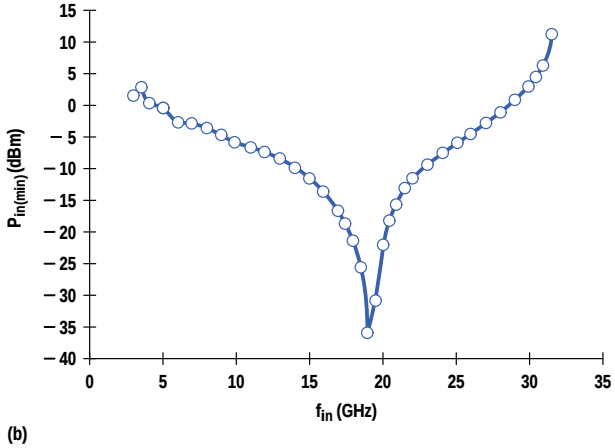
(b)

Figure 17

(a) Static divide-by-4 circuit, and (b) its input sensitivity curve.



(a)



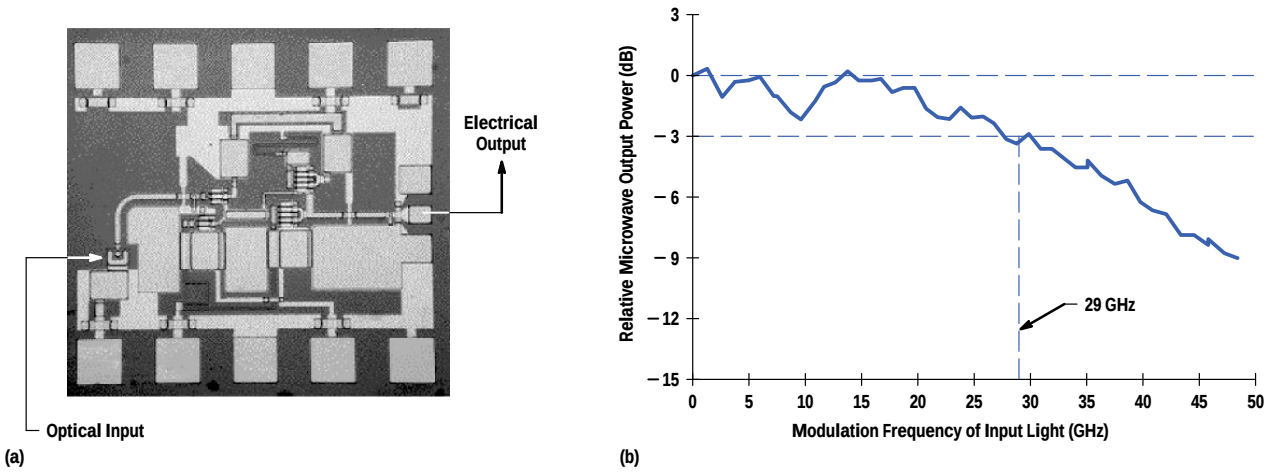
(b)

effective uniform density of harmful threading dislocations. This is indeed below the TEM detection limit, and low enough to yield useful circuits. Input sensitivity up to 31 GHz was measured on-wafer (see **Figure 17b**).

We have taken the process to a more complex level by integrating high-speed p-i-n photodetectors. We have demonstrated discrete fine-geometry detectors with bandwidths exceeding 50 GHz. **Figure 18a** shows a circuit[49] similar to the amplifier in **Figure 16a** but with a backside illuminated photodetector replacing the electronic input. The measured bandwidth of this photonic MMIC is 29 GHz (see **Figure 18b**). The response is well-modeled, and is limited by the capacitance of the p-i-n diode and a larger-than-typical R_s for the FETs used in this early attempt. We predict considerably better gain and bandwidth by using a smaller p-i-n detector along with our present FETs.

Figure 18

(a) Photonic MMIC amplifier, and (b) its frequency response. Substrate = InP.



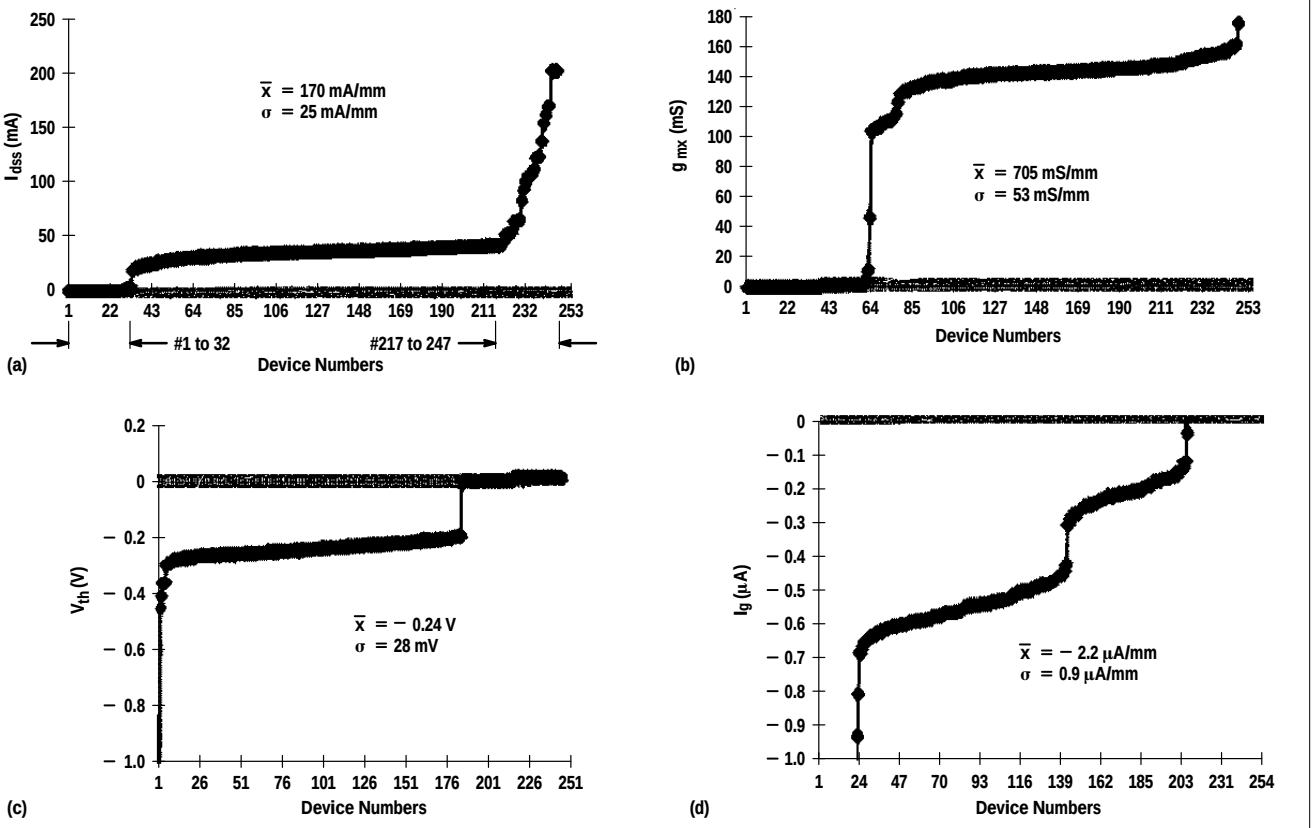
Process Development at HP Manufacturing Divisions

The demonstrated performance of MMICs, small digital demonstration circuits, photonic MMICs, and the demands for increased speed and volume of wireless communications make this technology attractive to HP for many applications. HP's Communication Semiconductor Solutions Division (CSSD) has a primary interest in the low-noise and high-gain quality of the discrete FETs. The noise and gain data in Appendix A were taken by CSSD at the most common DBS frequency of 12 GHz. CSSD has successfully transferred the MBE structure, the nonalloyed ohmic contacts, and the gate-recess etch solutions to its 3-in GaAs fabrication facility. **Figure 19** is an example of the level of control that has been achieved. It shows the sorting charts of four important dc parameters over a 3-in wafer for 247 200-µm-wide FETs. I_{dss} is the drain current in saturation with zero V_g. I_{dss} is an important parameter for classical depletion-mode processes with negative V_{th} and low g_{mx} at zero V_g. I_{dss} is then essentially equal to $I_d^{(max)}$, the parameter that best describes the available current in a FET. In our process, however, I_{dss} occurs close to maximum g_{mx}. Having g_{mx} peak at $V_g \backsim 0$ is desirable for circuit design and is made possible because of the process optimization described in this article. However, since g_{mx} is quite large, a standard deviation $\sigma_{g_{mx}}$ or $\sigma_{V_{th}}$ can contribute significantly to $\sigma_{I_{dss}}$. In fact, they will dominate, given that MBE-induced variations are typically negligible compared to those induced by the gate process. The fraction of $\sigma_{I_d^{(max)}}$ that also contributes is small. Given the sensitivity of I_{dss} to the high g_{mx}, we were pleased to see the good control exhibited in **Figure 19a**. A sizable fraction of the 31 FETs (217 to 247 in **Figure 19a**) had an abnormally large I_{dss} because they were too close to the wafer periphery where the material quality is lower or the gates are not exposed. The 32 FETs (1 to 32 in **Figure 19a**) with zero I_{dss} failed for undetermined reasons and had to be considered a 15% yield loss. This corresponds to a $5 \times 10^5\ cm^{-2}$ maximum uniform threading dislocation density, which is consistent with the earlier estimate. The 63 deviant FETs show up at different places in the sorting curves for g_{mx} (**Figure 19b**), V_{th} (**Figure 19c**), and the reverse gate leakage (**Figure 19d**). Using the values from **Figures 19b** and **19c** one finds that $\sigma_{I_{dss}}$ (= 25 mA/mm) is indeed dominated by $\sigma_{V_{th}}$ and $\sigma_{g_{mx}}$, since:

$$\sqrt{\left(\langle g_{mx}\rangle \sigma_{V_{th}}\right)^2 + \left(\langle V_{th}\rangle \sigma_{g_{mx}}\right)^2} = 24\ \text{mA/mm}\ .$$

Figure 19

CSSD dc sorting charts for (a) I_{dss}, (b) g_{mx}, (c) V_{th}, and (d) reverse gate leakage (V_d = 1V, I_d = 50 mA/mm) for 247 200-µm-wide FETs over a 3-inch wafer.



The uniformity exhibited in **Figure 19** by a 3-in wafer processed in CSSD's manufacturing environment is better than typically achieved on full or partial 2-in wafers in HP Laboratories' R&D environment. The low reverse gate leakage (< 3.5 µA/mm in **Figure 19d**) allows CSSD to achieve a low noise figure at the relatively low frequency of 12 GHz (see ***Appendix A***).[24]

HP's Microwave Technology Division (MWTD) addresses different businesses than CSSD. Microwave output power is of primary importance for most of the present markets. Despite the inherently lower breakdown voltage of InGaAs FETs, the high-speed performance of these devices still make them attractive at high frequencies,[1] where most higher-power processes run out of steam. The InGaAs-channel makes up, to some extent, for the low $BV_{ds}^{(on\ min)}$ by having a large $I_d^{(knee\ max)}$. With optimum lateral recess (see **Figure 10**) and output load match, useful output power can still be achieved as shown in equation 31 and **Figure 14**. This provided the motivation for a cooperative effort between HP Laboratories and MWTD, which has demonstrated that the process can be reproduced in MWTD's 3-in GaAs manufacturing environment (see ***Appendix B***).

Standard GaAs manufacturing unit processes account for about 85% of the steps. Other steps, such as gate lithography and recess, are kept nominally identical to those of HP Laboratories. To achieve 0.1-µm gates, the gate lithography is done at HP Labs with the e-beam process discussed earlier. FET performance and uniformity have been reproduced.

As with CSSD, the MBE growth expertise has been transferred. We believe that the manufacturability of the process has been successfully demonstrated, and hope that it will become an asset to HP's high-frequency circuit designers.

Acknowledgments

The process development has benefited greatly from cooperation with many of our colleagues. We are grateful to Nick Moll for sharing his insight in device physics and for reviewing this article, to Antoni Niedzwiecki for advice on noise modeling, to Alice Fischer-Colbrie, Dave Reed, Midori Kanemura, Shelli Nelsen, Mitchell Kido, Eleazar Ramirez, Howard Eng and George Patterson for contributions to the material growth, and to Marge Pustorino, Nancy Caldwell, Alan Kashiwagi, Virender Makker, and Yogesh Desai for the thin-film depositions. For their contributions to processing and process development we thank Roshan Merchant, Sue Harris, Debbie Ritchey, Denise Davis, Ed Wong, Alan Quash, Jerry Wang, Hengchang Chou, and Ho-Fai Wong. For their contributions to the 0.1-µm gate e-beam lithography, thanks go to Nadine Whittaker, Adrian Lee, and Ines Stolberg. For their process-equipment support thanks to Bill Collins and Warren Hargrave. For their help in testing and characterization, thanks to Mike Kauffman, Jeff Raggio, Rick Powell, David Briscoe, and JoAnn Peterson. For circuit designs and discussions thanks to Chris Madden, Rory Van Tuyl, and Jean Tillinghast. For their helpful advice and discussions, thanks to Karen Seaward, Francoise Mertz, Hans Queisser, Ben Keppeler, and David Kuhn. Finally, thanks to Rolf Jaeger, Jeff Miller, Gary Baldwin, Ding Day, Craig Snapp, Stretch Camnitz, Charles Stolte, Noel Fernandez, Jerry Gladstone, and Derry Hornbuckle for their managerial support.

References

1. R.L. Van Tuyl, "Unlicensed Millimeter Wave Communications: A New Opportunity for MMIC Technology at 60 GHz," *Technical Digest of the 1996 GaAs IC Symposium*, p. 3.

2. D. Polifko and H. Ogawa, "The Merging of Photonic and Microwave Technologies," *Microwave Journal*, March 1992, p. 75.

3. S. Yajima and A. Niedzwiecki, "Direct Broadcast Satellite Applications," *Hewlett-Packard Journal*, Vol. 49, no. 1, February 1998.

4. H. Rohdin and A. Nagy, "A 150 GHz Sub-0.1-µm E/D MODFET MSI Process," *Technical Digest of the 1992 IEDM*, p. 327.

5. H. Rohdin, "Analytical Forward and Reverse Modeling of MODFETs," in *Heterojunction Transistors and Small Size Effects in Devices*, M. Willander (Editor), Studentlitteratur/Chartwell-Bratt, 1992, p. 107.

6. N. Moll, "Delay and Current Due to Charged Particles in Generalized Space-Charge Regions," *to be published*.

7. J. Braunstein, P. J. Tasker, A. Hulsmann, K. Kohler, W. Bronner, and M. Schlechtweg, "G_{ds} and f_T Analysis of Pseudomorphic MODFETs with Gate Lengths down to 0.1 µm," *Proceedings of the GaAs and Related Compounds Conference*, 1993.

8. P.H. Ladbrooke, "Reverse Modelling of GaAs MESFETs and HEMTs," *GEC Journal of Research*, Vol. 6, 1988, p. 1.

9. N. Moll, M.R. Hueschen, and A. Fischer-Colbrie, "Pulse-Doped AlGaAs/InGaAs Pseudomorphic MODFETs," *IEEE Transactions on Electron Devices*, Vol. 35, 1988, p. 879.

10. H. Rohdin, "Reverse Modeling of E/D Logic Submicrometer MODFETs and Prediction of Maximum Extrinsic MODFET Current Gain Cutoff Frequency," *IEEE Transactions on Electron Devices*, Vol. 37, 1990, p. 920.

11. L.D. Nguyen, A.S. Brown, M.A. Thompson, and L.M. Jelloian, "50-nm Self-Aligned-Gate Pseudomorphic AlInAs/GaInAs High Electron Mobility Transistors," *IEEE Transactions on Electron Devices*, Vol. 39, 1992, p. 2007.

12. J. Xu and M. Shur, "Velocity-Field Dependence in GaAs," *IEEE Transactions on Electron Devices*, Vol. 34, 1987, p. 1831.

13. P.J. Tasker and B. Hughes, "Importance of Source and Drain Resistance to the Maximum f_T of Millimeter-Wave MODFETs," *IEEE Electron Device Letters*, Vol. 10, 1989, p. 291.

14. L.D. Nguyen, P.J. Tasker, and W.J. Schaff, "Comments on 'A New Low-Noise AlGaAs/GaAs 2DEG FET with a Surface Undoped Layer'," *IEEE Transactions on Electron Devices*, Vol. 34, 1987, p. 1187.

15. S.J. Mason, "Power Gain in Feedback Amplifier," *Transactions of the IRE*, Vol. CT-1, 1954, p. 20.

16. J.M. Rollett, "Stability and Power-Gain Invariants of Linear Twoports," *Transactions of the IRE*, Vol. CT-9, 1962, p. 29.

17. P. Wolf, "Microwave Properties of Schottky-barrier Field-effect Transistors," *IBM Journal of Research and Development*, Vol. 9, 1970. p. 125.

18. R.J. Trew and M.B. Steer, "Millimetre-Wave Performance of State-of-the-Art MESFET, MODFET and PBT Transistors," *Electronics Letters*, Vol. 23, 1987, p. 149.

19. H. Rohdin, A. Nagy, V. Robbins, C.-Y. Su, C. Madden, A. Wakita, J. Raggio, and J. Seeger, "Low-Noise, High-Speed $Ga_{.47}In_{.53}As$ / $Al_{.48}In_{.52}As$ 0.1-μm MODFETs and High-Gain/Bandwidth Three-Stage Amplifier Fabricated on GaAs Substrate," *Proceedings of the 1995 IPRM*, p. 73.

20. H. Fukui, "Optimal Noise Figure of Microwave GaAs MESFETs," *IEEE Transactions on Electron Devices*, Vol. 26, 1979, p. 1032.

21. M.W. Pospieszalski, "Modeling of Noise Parameters of MESFETs and MODFETs and their Frequency and Temperature Dependence," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 37, 1989, p. 1340.

22. R.A. Pucel, H.A. Haus, and H. Statz, "Signal and Noise Properties of Gallium Arsenide Microwave Field-Effect Transistors," *Advances in Electronics and Electron Physics*, Vol. 38, L. Marton (Editor), Academic Press, 1975, p. 195.

23. A. Cappy, "Noise Modeling and Measurement Techniques," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 36, 1988, p. 1.

24. R. Reuter, S. van Waasen, and F.J. Tegude, "A New Noise Model of HFET with Special Emphasis on Gate Leakage," *IEEE Electron Device Letters*, Vol. 16, 1995, p. 74.

25. W.W. Hooper and W.I. Lehrer, "An Epitaxial GaAs Field-Effect Transistor," *Proceedings of the IEEE*, Vol. 55, 1967, p. 1237.

26. R. Dingle, H.L. Stormer, A.C. Gossard, and W. Wiegmann, "Electron mobilities in modulation-doped semiconductor heterojunction superlattices," *Applied Physics Letters*, Vol. 33, 1978, p. 665.

27. A.S. Brown, A.E. Schmitz, L.D. Nguyen, J.A. Henige, and L.E. Larson, "The Growth of High Performance $In_xGa_{1-x}As(.52 < x < .9)$ – $Al_{0.48}In_{0.52}As$ High Electron Mobility Transistors by MBE," *Proceedings of the 1994 IPRM*, p. 263.

28. L.D. Nguyen, L.E. Larson, and U. K. Mishra, "Ultra-High-Speed Modulation-Doped Field-Effect Transistors: A Tutorial Review," *Proceedings of the IEEE*, Vol. 80, 1992, p. 494.

29. M. Hueschen, N. Moll, E. Gowen, and J. Miller, "Pulse Doped MODFETs," *Technical Digest of the 1984 IEDM*, p. 348.

30. M. Hueschen, N. Moll, and A. Fischer-Colbrie, "High-Current GaAs/AlGaAs MODFETs with f_T over 80 GHz," *Technical Digest of the 1987 IEDM*, p. 596.

31. K. Lee, M. Shur, T.J. Drummond, and H. Morkoc, "Electron density of the two-dimensional gas in modulation doped layers," *Journal of Applied Physics*, Vol. 54, 1983, p. 2093 .

32. N. Moll, "HFETs: A Study in Developmental Device Physics," *Compound Semiconductor Transistors: Physics and Technology*, S. Tiwari, Editor, IEEE Press, 1993, p. 3.

33. B. Vinter, "Subbands and charge control in a two-dimensional electron gas field-effect transistor," *Applied Physics Letters*, Vol. 44, 1984, p. 307.

34. N. Moll and H. Yeager, unpublished, program developed for internal HP use.

35. F. Gueissaz, T. Enoki, and Y. Ishii, "High current density double modulation-doped $Al_{0.48}In_{0.52}As$ – $Ga_{0.35}In_{0.65}As$ millimetre-wave HEMT," *Electronics Letters*, Vol. 29, 1993, p. 2222.

36. D.G. Ballegeer, I. Adesida, C. Caneau, and R. Bhat, "Physics and Behavior of Asymmetrically Recessed InP-based MODFETs Fabricated with an Electron Beam Resist Process," *Proceedings of the 1994 IPRM*, p. 331.

37. M. Arps, H.G. Bach, W. Passenberg, A. Umbach, and W. Schlaak, "Influence of SiN_x Passivation on the Surface Potential of GaInAs and AlInAs in HEMT Layer Structures," *Proceedings of the 1996 IPRM*, p. 308.

38. S.R. Bahl and J.A. del Alamo, "Physics of Breakdown in InAlAs/n^+ – InGaAs Heterostructure Field-Effect Transistor," *IEEE Transactions on Electron Devices*, Vol. 41, 1994, p. 2268.

39. J. Dickmann, S. Schildberg, A. Geyer, B.E. Maile, A. Schurr, S. Heuthe, and P. Narozny, "Breakdown Mechanisms in the On-State Mode of Operation of InAlAs / $In_xGa_{1-x}As$ Pseudomorphic HEMTs," *Proceedings of the 1994 IPRM*, p. 335.

40. G.T. Cibuzar, "Effects of Gate Recess Etching on Source Resistance," *IEEE Transactions on Electron Devices*, Vol. 42, 1995, p. 1195.

41. A.S. Wakita, C.-Y. Su, H. Rohdin, H.-Y Liu, A. Lee, J. Seeger, and V.M. Robbins, "Novel high-yield trilayer resist process for 0.1 μm T-gate fabrication," *Journal of Vacuum Science and Technology B*, Vol. 13, 1995, p. 2725.

42. H.-Y. Liu, C.-Y. Su, N. Farrar, and B. Gleason, "Fabrication of 0.1 μm T-shaped Gates by Phase-shifting Optical Lithography," *Digest of the 1994 U.S. Conference on GaAs Manufacturing Technology*, p. 41.

43. D.C. D'Avanzo, D.B. Estreich, T.W. Taylor, H. Kondoh, J. Orr, A. Quach, E.H. Wong, R. Yeats, V.K. Makker, and R.A. Fischer, "A Manufacturable, 26 GHz GaAs MMIC Technology," *Technical Digest of the 1988 GaAs IC Symposium*, p. 317.

44. A. Fischer-Colbrie, G.G. Zhou, and G. Hasnain, "High-Quality $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ MODFETs and PINs Grown on GaAs Substrates," *1993 Electronic Materials Conference*, paper A4.

45. R.S. Goldman, J. Chen, K.L. Kavanagh, H.H. Weider, V.M. Robbins, and J.N. Miller, "Structural and Magnetotransport Properties of InGaAs/InAlAs Heterostructures Grown on Linearly Graded Al(InGa)As Buffers on GaAs," *Institute of Physics Conference Series*, No. 141, Chapter 3.

46. J.-I. Chyi, J.-L. Shieh, C.-S. Wu, R.-M. Lin, J.-W. Pan, Y.-J. Chan, and C.-H. Lin, "Characteristics of $In_{0.3}Ga_{0.7}As/In_{0.29}Al_{0.71}As$ Heterostructures grown on GaAs Using InAlAs Buffers," *Japanese Journal of Applied Physics*, Vol. 33, 1994, p. L1574.

47. P. Win, Y. Druelle, Y. Cordier, D. Adam, J. Favre, and A. Cappy, "High-Performance $In_{0.3}Ga_{0.7}As/In_{0.29}Al_{0.71}As$ /GaAs Metamorphic High-Electron-Mobility Transistor," *Japanese Journal of Applied Physics*, Vol. 33, 1994, p. 3343.

48. C.J. Madden, R.L. Van Tuyl, M.V. Le, and L.D. Nguyen, "A 17 dB Gain, 0.1-70 GHz InP HEMT Amplifier IC," *1994 IEEE ISSCC Technical Digest*, p. 178.

49. R.L. Van Tuyl, unpublished circuit design.

Hans Rohdin

Hans Rohdin is a member of the technical staff in the device technology department in HP Laboratories. He joined HP in 1982 after receiving his DSc degree in electrical engineering in 1982 from Washington University in St. Louis, Missouri.

Avelina Nagy

Avelina Nagy is an R&D specialist in HP Laboratories working on semiconductor device processing. She joined HP in 1974.

Virginia M. Robbins

Virginia Robbins is a process development engineer in HP Laboratories. She is responsible for developing compound semiconductor materials used in electronic electronic and optical devices. She received a PhDEE degree from the University of Illinois in 1988. She enjoys outdoor activities, especially bicycing.

Chung-Yi Su

Chung-Yi Su joined HP Laboratories in 1981 after receiving his PhDEE degree from Stanford University. He is now GaAs fabrication manager at the HP Communications Semiconductor Solutions Division. He has published over 67 papers in the areas of semiconductor devices, solid surfaces, and interfaces. Born in Taiwan, he is married and has three children.

Arlene S. Wakita

A member of the technical staff at HP Laboratories, Arlene Wakita joined HP in 1984. She is responsible for e-beam lithography process development and device reliability. She received a PhD degree in materials science from Stanford University in 1984.

Judith W. Seeger

Judith Seeger is an R&D specialist in e-beam lithography in HP Laboritories. She joined HP Laboratories in 1981. She received a BA degree in journalism and mass communications from New Mexico State University in 1974.

Tony Hwang

Tony Hwang is a senior member of the technical staff at HP's Communication Semiconductor Solution Division. He has been with HP since 1991 and is responsible for improving and fabricating devices for DBS, microwave, and millimeter-wave applications. He received a PhdEE degree from Colorado State University in 1985. He is married, has one son, and enjoys camping, touring, and playing bridge.

Patrick W. Chye

An R&D project manager at HP's Communications Semiconductor Solutions Division, Patrick Chye is responsible for GaAs fabrication. He joined HP in 1991. He has a PhD degree in applied physics from Stanford University.

Paul E. Gregory

Paul Gregory is a manufacturing development engineer at HP's Communication Semiconductor Solution Division. He came to HP in 1991 and is responsible for epitaxial growth by MBE for microwave devices. He received a PhDEE degree from Stanford University in 1976. He is married and has two children.

Sandeep R. Bahl

Sandeep Bahl is a member of the technical staff at HP Laboratories, which he joined in 1993. He is working on the development of the next generation of heterojunction bipolar technologies. He transferred the InGaAs process to HP's Microwave Technology Division, and he provides support and consulting for this activity. He received his PhdEE degree from the Massachusetts Institute of Technology in 1993. He is married, has one child, and enjoys skiing, hiking, frisbee, tennis, and photography.

Forrest G. Kellert

Forrest Kellert is a manufacturing engineering project manager at HP's Microwave Technology Division, where he is working on semiconductor wafer fabrication. He joined HP in 1980 after receiving a PhD degree in physics from Rice University. He was born in Passaic, New Jersey, and he is married and has two children.

Lawrence G. Studebaker

Lawrence Studebaker is an integration engineer for millimeter-wave GaAs IC development and manufacturing at HP's Microwave Technology Division. Larry joined HP in 1979. He has an MS degree in materials science from Stanford University. He was born in Elmhurst, Illinois, is married and has two children. Participating in local school activities and amateur radio (K06KP) are among his outside-of-work activities.

Donald C. D'Avanzo

Don D'Avanzo is a new process development and integration engineering manager for development of new GaAs IC processes at HP's Microwave Technology Division. Don joined HP in 1979, just around the time he was finishing his PhDEE degree from Stanford University. He was born in Cranston, Rhode Island, is married and has two children. Outside activities include coaching little league baseball, snow skiing, gardening, and wine making.

Sigurd W. Johnsen

Sig Johnsen is a semiconductor marketing manager at HP's Microwave Technology Division, which he joined in 1979. He received a BS degree in physics from Rochester Institute of Technology in 1977.

▶ Go to Appendix A
▶ Go to Appendix B
▶ Go to Next Article
▶ Go to Journal Home Page