Chapter 5
HEMT Low-Noise Amplifiers

J. Javier Bautista

5.1 Introduction—Semiconductor Conductivity

A large variety of electronic devices (such as diodes, transistors, mixers, attenuators, and photo-detectors, to name a few) are based on the unique electrical conducting properties of semiconductors. Semiconductors are solid-state conductors whose electrical conductivity ranges approximately between $10^2$ and $10^9$ mho per centimeter (mho/cm) (or its inverse, resistivity between $10^{-2}$ and $10^9$ ohm-cm) at room temperature [1]. These conductivities are higher than an insulator’s but lower than a metal’s. Although most pure single-crystal semiconductors are insulators at absolute zero with conductivities less than $10^{-14}$ mho/cm, it is impurities and imperfections that allow manipulation of semiconductor’s device properties. In fact, it is the ability to precisely engineer and finely tune a semiconductor’s conductivity at the atomic layer level that is essential to the production of state-of-the-art high electron mobility transistor (HEMT) devices [2].

Important semiconductors like silicon (Si), germanium (Ge), gallium arsenide (GaAs), indium phosphide (InP) and their ternary derivatives, like aluminum gallium arsenide (AlGaAs), indium gallium arsenide (InGaAs), and indium aluminum arsenide (InAlAs), form a crystalline diamond (or zincblende) lattice structure bonded by covalent forces. Each atom has four nearest neighbors lying at the corners of a tetrahedron, and it shares an electron with each atom. The covalent attractive potential between atoms is created by the shared electron pairs of opposite spin. The three-dimensional models for the diamond and zincblende structures are shown in Fig. 5-1.
5.1.1 Charge Carrier and Energy Band Gap

Semiconductors can have two types of charge carriers that contribute to the conductivity, electrons and holes. In pure intrinsic semiconductors at room temperature, thermal energy can free an electron (negative charge) from its bonding or lattice site, leaving behind a vacant positively ionized atom. The vacancy is called a hole, and is equivalent to a charge carrier of positive charge moving in a direction opposite to the electron’s motion.

Unlike electrons in free space, which have a continuous set of energy states, electrons in a solid have forbidden bands of energy. At low energies electrons...
in a solid behave much like free, classical electrons. However, as the electron energy increases, they interact with and are scattered by the lattice, exhibiting their wave nature. This occurs at energies where the electron’s de Broglie wavelength, \( \lambda = h/p \) (Planck’s constant/electron momentum), approaches the inter-atomic spacing, \( a \), producing a band of forbidden energy states. A graphical representation for the electron energy as a function of wave vector \( (k = 2\pi/\lambda) \) for an ideal one-dimensional solid is shown in Fig. 5-2.

5.1.2 ChargeCarrier Transport Properties

In general, the band structure or energy-momentum (E-k) relationship for a semiconductor is obtained by solving the Schrödinger equation of an approximate one-electron problem using a variety of numerical methods (e.g., the orthogonalized plane wave method, the pseudopotential method, and the \( k \cdot p \) method) and verified by experimental measurements (such as optical absorption and cyclotron resonance techniques). The intrinsic transport properties that determine microwave- and millimeter-wave performance, such as electron mobility \( (\mu_e) \), peak velocity \( (v_p) \) and effective mass \( (m^*) \), can thus be calculated and verified by measurements at room and cryogenic temperatures.

In general, near the bandgap edge the electron effective mass is inversely proportional to the second derivative of the energy with respect to wave-vector,
Thus for large values of $\frac{\partial^2 E}{\partial k^2}$, $m^*$ can be much less than the free electron mass. In addition, below room temperature the electron mobility is approximately inversely proportional to the product of the effective mass and physical temperature, that is $\mu_e \propto (m^* T)^{-1}$. While the electron velocity is in turn proportional to the product of the electron mobility and applied electric field,

$$v = \mu_e E$$  \hspace{1cm} (5.1-2)

Although silicon is the dominant material in use today for the manufacture of transistors and is the most mature technology, the III–V semiconductor compounds have far superior transport properties, especially at high frequencies and low temperatures. In fact, the $m^*$ of GaAs and InP is less than an order of magnitude smaller than Si’s, and their $\mu_e$’s are more than two orders of magnitude greater than Si’s at liquid nitrogen temperatures (77 K). In fact, the maximum intrinsic frequency of operation for a HEMT device is directly proportional to the electron velocity. Listed in Table 5-1 for comparison is the band gap energy ($E_g$), electron mobility ($\mu_e$), peak velocity ($v_p$), and lattice constant (a) for a number of important semiconductors.

<table>
<thead>
<tr>
<th>Semiconductor</th>
<th>$E_g$ (eV)</th>
<th>$\mu_e$ (cm$^2$/Vs)</th>
<th>$v_p$ (10$^7$cm/s)</th>
<th>a (Å)</th>
<th>a (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ge</td>
<td>0.66</td>
<td>3900</td>
<td>0.6</td>
<td>5.65</td>
<td>0.565</td>
</tr>
<tr>
<td>Si</td>
<td>1.12</td>
<td>1500</td>
<td>1.0</td>
<td>5.43</td>
<td>0.543</td>
</tr>
<tr>
<td>GaAs</td>
<td>1.42</td>
<td>8500</td>
<td>1.8</td>
<td>5.65</td>
<td>0.565</td>
</tr>
<tr>
<td>InP</td>
<td>1.35</td>
<td>4600</td>
<td>2.4</td>
<td>5.87</td>
<td>0.587</td>
</tr>
<tr>
<td>Ga$<em>{0.13}$In$</em>{0.85}$As</td>
<td>1.20</td>
<td>9500</td>
<td>2.9</td>
<td>5.85</td>
<td>0.585</td>
</tr>
<tr>
<td>Ga$<em>{0.47}$In$</em>{0.53}$As</td>
<td>0.75</td>
<td>15,000</td>
<td>3.4</td>
<td>5.85</td>
<td>0.585</td>
</tr>
<tr>
<td>InAs</td>
<td>0.36</td>
<td>33,000</td>
<td>4.4</td>
<td>6.06</td>
<td>0.606</td>
</tr>
<tr>
<td>InSb</td>
<td>0.17</td>
<td>80,000</td>
<td>5.0</td>
<td>6.48</td>
<td>0.648</td>
</tr>
</tbody>
</table>
5.1.3 Donor and Acceptor Impurities

Electronic properties of semiconductors are engineered and optimized during material growth through the introduction of impurities, a process known as “doping.” In Si and Ge phosphorous (P) or boron (B) are used as dopant impurities, while Si is used in GaAs and InP. Dopants with a valence of five (like P) replace the host atom Si or Ge adding an extra valence electron (donor state), while those with a valence of three (like B) reduce the number of valence electrons by one (acceptor state). However, in GaAs and InP the dopant Si can either introduce a donor state by replacing Ga (or In) or an acceptor state by replacing As (or P). The donor state is more energetically favorable, and therefore dominant.

Dopant impurities introduce discrete energy levels, donor and acceptor levels, within the forbidden energy band of the intrinsic semiconductor. These impurities states are similar to hydrogen atomic states with the exception that the orbital electrons or holes move through a dielectric medium. Recalling that the ionization energy for the hydrogen atom is

$$E_H = \frac{m_e e^4}{8 \varepsilon_0^2 \hbar^2} = 13.6 \text{ eV}$$

(5.1-3)

The ionization energy for the donor state $E_d$ can be calculated by replacing the electron mass, $m_e$, with its effective mass, $m^*$, and the free space permittivity, $\varepsilon_0$, with the permittivity of the semiconductor, $\varepsilon_s$, yielding

$$E_d = \left( \frac{m^*}{m_e} \right) \left( \frac{\varepsilon_0}{\varepsilon_s} \right)^2 E_H$$

(5.1-4)

The ionization energy for donors calculated with this expression are 0.006, 0.025, and 0.007 eV for Ge, Si, and GaAs, respectively. The ionization energies can be calculated in a similar manner for acceptor states with comparable values (as measured from the valence-band edge): 0.015, 0.05, and 0.05 eV for Ge, Si, and GaAs, respectively. The donor energy states lie close to the conduction band edge, while acceptor states lie close to the valence band edge. In fact, the energy required to ionize these hydrogen-like atomic states can be estimated using this simple model. Since both are loosely bound states, both can be easily ionized by thermal energy resulting in excess electrons in the conduction band and excess holes in the valence band.
5.1.4 Heterojunction—HEMT versus MESFET

Until the invention of the HEMT, the most widely used III–V transistor for both microwave and high-speed digital applications was the GaAs metal semiconductor field effect transistor (MESFET). However, since electrons must transit through the doped channel in a MESFET, it does not take full advantage of the high mobilities in GaAs. The result is more than a 50-percent reduction in electron mobility, since ionized dopants scatter electrons. Hence, separation of the dopant channel from the electron transit channel is key to the superior noise, gain and frequency performance of the HEMT. For comparison, the cross sections of a GaAs HEMT and a MESFET are shown in Fig. 5-3 and a comparison of their material properties are shown in Table 5-2.

The model HEMT structure can be formed of two distinct semiconductor layers [4]. The bandgap difference results in the formation of conduction and valence band discontinuities at the layer interface or heterojunction creating a quantum well in the conduction band. The wider band gap semiconductor is
HEMT Low-Noise Amplifiers

5.2 The Many Acronym-ed Device (MAD)—A Brief HEMT History

The commercial HEMT evolved from the GaAs and AlGaAs superlattice (multiple heterorstructures) research conducted in the late 1960s by Leo Esaki and Ray Tsu at the IBM Thomas J. Watson Research Center in Yorktown Heights, New York [5]. It was not until 1978 when Raymond Dingle, Horst Stormer, and Arthur Gossard at Bell Laboratories, Murray Hill, New Jersey, first demonstrated the high mobilities possible in these superlattices that worldwide interest developed at a number of university, industrial, and government laboratories [6,7]. Among them were Cornell University, University of Illinois, University of Michigan, University of Tokyo, University of Duisburg, Germany, Rockwell International, General Electric Company, TRW Inc., Fujitsu Ltd., Japan, Thomson CSF, France, U.S. Naval Laboratory, and the U.S. Air Force Avionics Laboratory, to name but a few. The first cryogenic, microwave HEMTs were reported in 1983 by Thompson CSF, France [8], and Fujitsu, Ltd., Japan [9]. The HEMT is known by other names, such as modulation-doped field effect transistor (MODFET), two-dimensional electron gas field effect transistor (TEGFET), selectively-doped heterostructure transistor (SDHT), and heterojunction field effect transistor (HFET) reflects the number of laboratories involved world-wide in its development and the device property they chose to emphasize.

5.2.1 HEMTs in the Deep Space Network and Radio Astronomy—Voyager at Neptune

One of the great technical challenges faced by the Jet Propulsion Laboratory’s (JPL’s) Deep Space Network (DSN) is to receive signals from spacecraft that are up to billions of kilometers from Earth. Since the transmitted power is limited and fixed, the communication burden is placed on the ground-based antenna receive systems, which must detect an extremely weak signal in
the presence of a nearly overwhelming amount of noise. The Deep Space
Network (DSN) approach is to simultaneously maximize the received signal
collection area and receiver sensitivity. This key figure of merit for a
telecommunications link is the ratio of antenna gain to operational noise
temperature \( G/T_{op} \) of the system.

For the Voyager encounter with Neptune, the 27 antennas of the National
Radio Astronomy Observatory’s (NRAO) Very Large Array (VLA) in
Socorro, New Mexico, were successfully arrayed with the antennas of the
DSN’s Goldstone complex in California at 8.4 GHz. The VLA was equipped
with cryogenic, HEMT low-noise amplifiers (LNAs) while the DSN was
operated with its extraordinarily sensitive cryogenic, ruby maser LNAs.

Since masers were too expensive, the VLA was to be originally equipped
with cryogenic, GaAs field effect transistor (FET) LNAs. However, in 1983
during the definition phase of the VLA/Goldstone Telemetry Array project, the
first cryogenic, HEMT was announced. The results, reported by Thompson
CSF [8] and Fujitsu [9], suggested that the cryogenic HEMT was capable of
significantly lower noise than FETs at cryogenic temperatures. Their results
further suggested that much of the system noise performance lost with the use
of FETs instead of masers could be regained if the new HEMT device could be
developed in time for VLA implementation. In 1984, a cooperative program
was initiated among NRAO, JPL, General Electric Co. (GE), and Cornell
University to develop a HEMT device and HEMT amplifiers optimized for
cryogenic use in the 1- to 10-GHz range. GE was contracted by JPL to design
and fabricate the devices, while NRAO assumed responsibility for device
evaluation and amplifier development at 8.4 GHz for the VLA. JPL was
similarly responsible for device evaluation and amplifier development for DSN
applications at 2.3 GHz. During the program, GE HEMT noise temperatures
and device yields steadily improved. At the start of the program in 1985, the
best device cryogenic noise temperature demonstrated at 8.4 GHz was 8.5 K
with an associated gain of 12 dB. Near the end of the program in 1986, the
device noise temperature had dropped to 5.5 K with more than 14-dB
associated gain [10].

5.2.2 InP HEMT LNAs in the Deep Space Network

To date, the noise, gain, and maximum frequency of InP HEMTs at room
temperature is steadily improving as the technology is being internationally
developed and commercialized. Although device (commercial and research)
noise temperatures continue to fall at ambient, there is no guarantee that an
attendant improvement at cryogenic temperatures will be realized. To develop
ultra-low-noise microwave amplifiers for cryogenic applications, one must
have a reliable source of state-of-the-art cryogenic devices and the capacity to
accurately characterize them at the device or wafer level at cryogenic
temperatures. The work on cryogenic, InP HEMTs was based on another partnership among TRW, Inc., the Georgia Institute of Technology (GIT), and the Jet Propulsion Laboratory (JPL). TRW was responsible for device fabrication and optimization, GIT cryogenic device noise parameter characterization, JPL cryogenic device scattering parameter characterization, LNA module development, and receiver implementation.

The development of cryogenic, InP HEMTs has enabled the demonstration of state-of-the-art LNA modules that are yielding noise temperatures less than ten times the quantum noise limit from 1 to 100 GHz (0.5 to 50 K). The state-of-the-art noise temperature of cryogenic, HEMT-based amplifiers has steadily improved since the invention of the HEMT. Notable examples at physical temperatures near 20 K are 5.5 K at 8.4 GHz in 1986 [10], 15 K at 43 GHz in 1993 [11], and 30 K at 102 GHz in 1999 [12].

The DSN is in the process of implementing this technology to meet its current and future (2010) navigation, telemetry, radar, and radio science needs at 8.4 and 32 GHz. To date, typical InP HEMT LNA modules developed for the DSN have demonstrated noise temperatures of 3.5 K at 8.4 GHz and 8.5 K at 32 GHz. Front-end receiver packages employing these modules have demonstrated operating system noise temperatures of 17 K at 8.4 GHz on a 70-m Cassegrain antenna and 39.4 K at 32 GHz on a 34-m beam wave-guide (BWG) antenna, both at zenith [13].

5.3 HEMT Growth Technology

The material optimization of HEMT structures, sometimes referred to as bandgap engineering, is complex and typically focused on increasing the electron’s mobility, velocity, and density and on improving its confinement to the channel layer. The material quality is key to achieving optimal device performance and can be significantly degraded by impurities, defects, heterojunction roughness, and lattice mismatch strain. These criteria dictate relatively slow growth rates of 1 to 2 atomic layers per second (epitaxial growth) of lattice matched materials on heated semi-insulating GaAs or InP wafers. Molecular beam epitaxy (MBE) [14–16] and metal-organic chemical vapor deposition (MOCVD), sometimes referred to as metal-organic vapor phase epitaxy (MOVPE) [17,18], are the two primary techniques used for HEMT epitaxial growth. (For digital and power applications a high transconductance over a large gate bias range is desirable, while for microwave devices the peak transconductance is more important.)

5.3.1 Molecular Beam Epitaxy

MBE growth of HEMT structures such as AlGaAs/GaAs, AlGaAs/InGaAs/GaAs, and InAlAs/InGaAs/InP HEMTs is performed in a vacuum chamber held at ultra-low pressures (less than $10^{-8}$ torr
(1.33 \times 10^{-6} \text{Pa})). High purity elemental sources (e.g., Al, Ga, and In) are evaporated while As and antimony (Sb) are sublimated when heated to high temperatures in furnaces known as effusion or Knudsen cells by atomic and molecular beams, respectively. For example, the epitaxial growth of GaAs is performed by heating and evaporating atomic Ga from the melt, while solid arsenic is heated to sublimate As$_4$ which is “cracked” by further heating to produce molecular, As$_2$. The furnace temperatures are adjusted so that the flux ratio of gallium to arsenic is 2 to 1. The atomic (Ga), and molecular (As$_2$), beams are aimed at a heated, substrate (GaAs or InP) and adsorbed on the heated substrate. They react to form GaAs; facilitating the growth of GaAs one atomic layer at a time. Typical growth rates range from 0.1 to 5 $\mu$m per hour. During growth n-type (Si) and p-type (beryllium, Be) dopants (heated in effusion cells) can be incorporated into the layers. The epilayer composition is controlled with shutters placed in front of the sources and by variation of the substrate temperature. In addition, epilayers must be closely lattice matched to avoid crystalline defects, such as dislocations. The lattice constant and the energy gap of some III–V semiconductor compounds are plotted in Fig. 5-4, and those compounds lattice matched to GaAs and InP are noted.

### 5.3.2 Metal-Organic Chemical Vapor Deposition (MOCVD)

Gaseous sources, such as trimethyl or triethyl organometallics, of the elemental compounds are used for the epitaxial layer growth of HEMT.
HEMT Low-Noise Amplifiers

structures in MOCVD. The gaseous compounds are “pyrolyzed” at the heated surface of a substrate like GaAs. For example, to grow GaAs trimethylgallium, (CH₃)₃Ga, (abbreviated TMGa) carried by hydrogen gas and gaseous arsine, AsH₃, are passed over the substrate surface. The adsorbed radicals react to produce GaAs and methane gas. Silane (SiH₄), diluted with hydrogen can be added to the mix to incorporate donor impurities. MOCVD growth rates can be as low as 1 μm/h to as high as 30 μm/h. The main advantage of MOCVD is high throughput and multi-wafer growth capability. Another important advantage is the ability to grow various indium- and phosphorus-containing compounds using phosphine (PH₃). However, in comparison to MBE, it is a greater challenge to achieve the same level of thickness and doping uniformity using MOCVD because of vapor-phase effects, such as source gas-phase depletion, turbulence, and convection.

Another advantage of MBE over MOCVD is the ability to use powerful in situ characterization techniques, such reflection high-energy electron diffraction (RHEED) for monolayer counting and compositional measurement of layers as well as Auger electron spectroscopy (AES), quadrupole mass spectrometry, and automatized ellipisometry [16]. In addition, MBE achieves higher compositional resolution, accuracy, and uniformity across the wafer compared to MOCVD. However, MBE is more complex and expensive, and has lower throughput, while MOCVD is relatively simple and less expensive, and is capable of a higher throughput capacity. Through improvements in the design of MBE systems, significant progress has been made in improving MBE surface morphology and throughput. Production MBE machines have been commercially available for more than a decade, currently costing on the order of one million dollars with a comparable amount required for operation and maintenance per year.

To summarize, MBE provides the ultimate control for growing heterostructures with precise layer composition, doping profile, and sharp layer transition. To date, the best device and circuit results have been demonstrated with MBE grown HEMTs.

5.4 HEMT Materials Evolution—From GaAs to InAs

As previously mentioned, the first successful HEMTs were based on the lattice-matched heterostructure AlGaAs/GaAs. A few years later, In was added to the carrier channel to improve device performance, that is, to increase electron mobility (μₑ), increase the frequency of operation and lower the noise. The InGaAs channel layer inserted between the AlGaAs and GaAs is not lattice matched to either compound, GaAs or AlGaAs, but compressed to match them at their interface; accounting for the device name, pseudomorphic HEMT (PHEMT). The thickness of the InGaAs layer (between 50 to 200 Å [5–20 nm], depending on In concentration) is chosen so that most of the compressional
strain is taken up by the InGaAs layer [19]. To further push device performance, the In concentration was increased from the range of 15–20 percent to 65 percent, the spacer material changed to InAlAs, and InP substrates used to accommodate the larger lattice mismatch [20]. Pure InAs as the carrier channel is currently being investigated as the next logical step to produce the ultimate in HEMT device performance. The advantage of this material is the high electron mobility (30,000 cm²/Vs at 300 K) and velocity ($4 \times 10^7$ cm/s) and a large conduction band offset between InAs and AlSb (1.35 eV) [21,22].

Although the layer number, composition, and thickness vary depending on the desired properties, all HEMT layer structures have the same essential feature, a vertical heterojunction. The heterojunction spatially separates charge carriers from donors and confines them to the channel layer where the electron momentum is quantized in the “vertical” direction but is continuous in the horizontal direction. In the following discussion, the design guidelines for an optimal low noise AlGaAs/GaAs HEMT device are given, then a summary and comparison is presented of some of the essential enhancing features of PHEMTs and lattice-matched InP HEMTs.

5.4.1 Optimized Low-Noise AlGaAs/GaAs HEMT Structure

As shown in Fig. 5-3a, the conventional AlGaAs/GaAs HEMT structure is grown on a GaAs semi-insulating substrate with the following epitaxial layers: an undoped buffer and GaAs channel layer, an undoped AlGaAs spacer layer, a heavily doped (n+) AlGaAs donor or gate-barrier layer, and an n+ GaAs capping or ohmic contact layer. These layers are essential for fabricating and understanding the operation of a HEMT device. Depending on the application (for example, low noise, power, or digital), modifications and refinements to the basic structure are necessary to obtain optimum device performance. Some of the modifications and refinements to fabricate low-noise, high-gain devices are presented in the following paragraphs.

Under normal bias conditions the drain-to-source electric field can inject electrons beyond the 2-DEG channel into the GaAs buffer layer, contributing excess drain current, resulting in gain reduction and degradation of the device noise performance. Introduction of a high band-gap AlGaAs buffer layer before the GaAs buffer suppresses the buffer layer drain-to-source leakage current by creating an energy barrier in the conduction band to reduce electron injection into the buffer, while reducing the velocity of injected electrons [23]. The use of an AlGaAs buffer, however, results in buffer-channel interface roughness that reduces the mobility in the device channel [24]. The interface roughness can be improved by incorporating a thin GaAs smoothing or AlGaAs/GaAs superlattice buffer [25] layer between the buffer and the channel. A superlattice buffer, thin alternating layers of differing materials sharing the same crystalline
HEMT Low-Noise Amplifiers

The thin spacer layer separating electrons from their donors is to reduce the scattering of electrons by the positively charged donors. This is done by placing a thin spacer layer of undoped AlGaAs with a thickness ranging from 20 to 50 Å (2–5 nm) between the AlGaAs donor and the GaAs channel layer to separate the negatively charged 2-DEG from the ionized dopant atoms. At room temperature, a thin spacer layer of approximately 20 Å (2 nm) is preferred for low-noise and power devices due to the reduced parasitic source resistance and the increased transconductance and current density. A thicker spacer, conversely, provides a higher electron mobility with a smaller charge density in the channel. At cryogenic temperatures the noise performance of a HEMT is strongly dependent on the spacer thickness, and a thickness of 40 Å (4 nm) has been determined to be optimum due to the large increase in electron mobility and velocity [26].

In order to eliminate parallel conduction in the AlGaAs donor layer, this layer must be completely depleted by both the AlGaAs/GaAs heterojunction and the Schottky gate. The donor layer is typically uniformly doped with Si at a doping level of approximately $10^{18}$ atoms/cm³. The high doping level makes possible the small spacing between the gate and the carrier channel. A higher doping level results in a higher sheet charge density in the channel, increasing transconductance ($g_m$), unity current gain frequency ($f_T$), and current density, at the expense of a lower breakdown voltage. Fortunately, high sheet charge density and breakdown voltage can be achieved with planar-doping, sometimes also referred to as δ-doping or pulse-doping [27,28]. The planar-doping layer is a monolayer of Si approximately 5 Å (0.5 nm) thick with a doping level of approximately $5 \times 10^{13}$/cm² located just above the spacer. The use of planar-doping also allows a lower doping level in the AlGaAs layer for the gate barrier, increasing the breakdown voltage without sacrificing the channel sheet charge density.

The AlAs mole fraction, $x$, in the AlₓGa₁₋ₓAs donor layer is another important parameter controlling low-noise performance. At room temperature the conduction band discontinuity, $\Delta E_c$, at the AlₓGa₁₋ₓAs/GaAs heterojunction interface is linearly dependent on the AlAs mole fraction given by [29]

$$\Delta E_c = 0.8806x, \text{ for } x < 0.47$$  \hspace{1cm} (5.4-1)

This expression indicates that an increase in the AlAs mole fraction will result in an increase in the total carrier sheet charge density. In addition, a higher AlAs mole fraction reduces scattering of carriers from the carrier channel to the AlGaAs layer and minimizes the excess modulation of the
AlGaAs layer [30]. However, the high AlAs mole fraction reduces the doping efficiency in the channel and for $x > 0.25$ leads to the production of the DX centers [31]. DX centers are deep donor states related to the band structure of AlGaAs and are responsible for the persistent photoconductivity (PPC), threshold voltage shift, and drain I–V collapse at low temperatures [32,33]. Although, DX centers do not pose a problem at room temperature, they severely limit the performance of cryogenic low-noise microwave HEMTs.

To provide a good ohmic contact to the charge carrier channel, the GaAs capping layer is approximately 500 Å (50 nm) thick and typically heavily doped with Si at approximately $10^{18}$ atoms/cm$^3$. This reduces the device source resistance and protects the AlGaAs donor layer from surface oxidation and depletion. Higher doping levels and a thicker capping layer would simultaneously reduce the device source resistance and the effectively shorten the source-to-drain spacing, resulting in very high electron velocity, transconductance ($g_m$), and unity current gain frequency ($f_T$) in the device. However, this also significantly reduces the device breakdown voltage and also increases the device drain output conductance and drain-to-gate feedback capacitance. It is also more difficult to obtain uniform gate recess in a HEMT with a thicker capping layer. The capping layer has to be completely recessed under the gate to eliminate parallel conduction in the GaAs layer.

5.4.2 The GaAs Pseudomorphic HEMT—A1GaAs/InGaAs/GaAs PHEMT

In 1986 the GaAs pseudomorphic HEMT was introduced as a high performance alternative to the AlGaAs/GaAs HEMT. Substitution of InGaAs for GaAs as the two-dimensional electron gas channel improves transport properties due to the higher mobility of InGaAs and stronger electron confinement associated with the quantum well at the heterojunction. Thus, injection of electrons back into the AlGaAs from the InGaAs is significantly reduced; thereby improving the transport properties. The larger conduction-band discontinuity at the AlGaAs/InGaAs heterojunction allows a higher sheet charge density and hence a higher current density and transconductance. Additionally, the electron mobility and peak velocity can be further improved by increasing the indium concentration (Fig. 5-5). The room temperature mobility of this type of PHEMT structure is generally 5000 to 7000 cm$^2$/Vs with a 2-DEG concentration of $1–3 \times 10^{12}$ cm$^{-2}$.

Although InGaAs is not lattice matched to either the AlGaAs donor or the GaAs buffer layers (Fig. 5-4), the strain associated with the lattice mismatch can be elastically accommodated within the InGaAs layer. For example, for a PHEMT structure like AlGaAs/In$_x$Ga$_{1-x}$As/GaAs where $x$ is in the range of 0.15 to 0.20, the InGaAs layer must be smaller than the critical thickness ~150 Å (15 nm). Above the critical thickness, lattice dislocations form, while
for a thickness less than ~ 50 Å (5 nm), quantum size effects substantially reduce electron confinement and increase electron scattering [17,34].

In 1990 the state-of-the-art performance for a 0.1-μm gate length device PHEMT was an $f_{\text{max}}$ of 290 GHz [35], an $f_T$ of 130 GHz, and a minimum noise figure of 2.1 dB with an associated gain of 6.3 dB at 94 GHz. Pseudomorphic technology is quite mature, and microwave monolithic integrated circuits (MMICs) based on this type of material are common and routinely exceed this performance [36].

5.4.3 InAlAs/InGaAs on an InP HEMT

A quick check of Fig. 5-4 (higher conduction band discontinuity) and Fig. 5-5 (higher electron velocity) suggest that increasing the In concentration in the carrier channel of the PHEMT will result in further improvements in electron carrier confinement and transport properties. Unfortunately, increasing the indium concentration in $\text{In}_x\text{Ga}_{1-x}\text{As}$ also increases the lattice constant.
Increasing the In concentration to the highest possible value is desirable; however, the higher lattice mismatch strain between $\text{In}_x\text{Ga}_{1-x}\text{As}$ and GaAs cannot be elastically accommodated. Thus, GaAs is not suitable as substrate material for such high indium concentration HEMT structures.

Fortunately, the ternary compounds, $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ and $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ are lattice matched to InP, and they can thus be used as high bandgap and low bandgap material, respectively, in the HEMT structure. Low breakdown voltages due to impact ionization in the InGaAs channel, limit this material to low-noise applications. HEMT devices using this material structure have very good high-gain and low-noise performance at high frequencies. Additionally, this structure also eliminates deep-level trapping sites (DX centers), providing a solution to the persistent photoconductivity effect at cryogenic temperatures. That is, unlike AlGaAs, InAlAs does not suffer from DX center effects and can be doped with Si up to $1 \times 10^{19}$ cm$^{-2}$ [37].

In fact, the mobility in lattice matched InP HEMT structures varies from 8,000 to 12,000 cm$^2$/Vs with a 2-DEG concentration of 2 to $4 \times 10^{12}$ cm$^{-2}$. In 1990 state-of-the-art performance for InP-based HEMTs displayed a maximum operating frequency ($f_{\text{max}}$) of 455 GHz and a maximum available gain of 13.6 dB at 95 GHz for a 0.15-μm gate length HEMT [38]. Another InP-based HEMT had a minimum noise figure of 1.7 dB with an associated gain of 7.7 dB at 93 GHz [39].

Materials with still higher indium concentrations in the channel have superior electron transport characteristics. There are several approaches for the incorporation of more indium into the carrier channel: pseudomorphic InGaAs channel with an In concentration above 0.53 [40], insertion of a thin $\text{In}_x\text{Ga}_{1-x}\text{As}$ ($x > 0.53$) layer in the channel region [41–43], change of material system to InAs/AlGaSb [44], and a graded channel [45,46]. In most of these approaches the thickness of the high-In concentration InGaAs layer must be kept thin to prohibit formation of dislocations. An InP-based pseudomorphic HEMT had a transconductance of 1700 millisiemens/millimeter (mS/mm) with an $f_T$ of 305 GHz for a 65-nm gate length device [47].

An approach that does not limit the channel thickness to the critical layer thickness is to lattice mismatch the substrate to the channel with a buffer that transforms the lattice constant of the substrate to that of a high indium concentration channel. Various methods to realize this buffer have been demonstrated such as graded InGaAlAs [48,49], graded $\text{In}_x\text{Ga}_{1-x}\text{As}$ [50], and step-graded $\text{In}_y\text{Ga}_{1-y}\text{As}$ [51] buffer layers. In [50], a 0.4-μm gate $\text{In}_{0.29}\text{Al}_{0.71}\text{As}/\text{In}_{0.33}\text{Ga}_{0.77}\text{As}$ HEMT on GaAs with an $f_{\text{max}}$ and an $f_T$ of 115 and 45 GHz, respectively, was presented. The transconductance and the current density were 700 mS/mm and 230 mA/mm. In [51], a 0.16-μm
In$_{0.5}$Al$_{0.5}$As/In$_{0.5}$Ga$_{0.5}$As HEMT on GaAs substrate had an $f_{\text{max}}$ and an $f_T$ of 147 and 67 GHz, respectively. The transconductance and current density were 1060 mS/mm and 550 mA/mm.

To date, InP-based HEMTs have been shown to be the best performing three-terminal devices [52], with excellent performance in the microwave and millimeter-wave range. The combination of high gain and low noise has been demonstrated by many devices and circuits having operating frequencies as great as 100 GHz and higher [53–56].

### 5.4.4 InAlAs/InGaAs on GaAs HEMT—Metamorphic HEMT or MHEMT?

A drawback of the InP substrate is that it is a fairly young technology compared to GaAs. InP substrates are more expensive, more fragile, and more difficult to etch. Currently, only 2-inch (in.) and 3-in. (5- and 7.6-cm) high quality InP substrates are available at relatively high cost, whereas 6-in. (15.2-cm) GaAs substrates are readily available. Additionally, wafer thinning and backside processing technologies are more mature for GaAs. To combine the advantages of the GaAs substrate with the advantages of InP based HEMTs, metamorphic InGaAs/InAlAs quantum-well structures or metamorphic high electron mobility transistors (MHEMTs) were developed.

In MHEMTs the lattice constant of GaAs is transformed into the InP lattice constant with an appropriate buffer (typically 1 to 2 $\mu$m thick (1000 to 2000 nm)), on which the lattice-matched InP HEMT layer is grown. MHEMTs using quaternary buffers (such as AlGaAsSb) and a ternary buffer (such as InAlAs on GaAs) have show performance comparable to latticed InP HEMTs at room temperature. The quality of the final heterostructure, and thus the device performance, depends fully on the buffer type and quality [57]. Table 5-3 is a summary of the evolution of the first HEMT structure to current HEMT structure technology.

### 5.5 Device Fabrication

This section outlines the major processing steps in the fabrication of HEMT devices and microwave monolithic integrated circuits (MMICs). These processing steps are mostly based on, and hence quite similar, to GaAs MESFET processing. The major steps following growth of the heterostructure material are surface preparation and cleaning, front-side processing, and backside processing. Front-side processing steps include mesa or device isolation, ohmic contact formation, gate formation, metallization, and device passivation, while back-side processing includes substrate thinning via-hole formation and dicing. The processes are all performed in a clean room environment, varying in grade from Class 10,000 for the least critical steps of
wafer preparation to Class 100 for the most critical steps, such as ohmic contact and gate formation.

5.5.1 Wafer Preparation and Cleaning

To maintain optimum fabrication conditions and insure high device yields, cleaning operations are performed before all major steps during device processing. Relatively benign solvents, acids, bases, and rinses are used to remove contaminants such as organic materials, metals, and oxides.

For example, organic solvents effectively remove oil, grease, wax, photo resist, and electron-beam resist without affecting the HEMT device and circuit materials. The most common cleaning method is to immerse and agitate the wafer in the heated solvent. Solvents are then removed with alcohol that is, in turn, rinsed off with filtered, de-ionized water.

Acids are used for the wet etch removal of III–V semiconductor material as well as the removal of metal and oxide contaminants. The presence of thin interfacial dielectric layers, such as oxides, cause poor ohmic contact and Schottky barrier formation. These oxides can be dissolved with wet bases or removed with plasma etching techniques. Following cleaning and rinsing, the wafer is carefully dried to avoid leaving solvents or water stains.

<table>
<thead>
<tr>
<th>Layer</th>
<th>HEMT (x&lt;0.25)</th>
<th>GaAs PHEMT</th>
<th>InP HEMT</th>
<th>InAs HEMT</th>
<th>MHEMT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cap</td>
<td>Heavily doped</td>
<td>Heavily doped</td>
<td>Heavily doped</td>
<td>InAs</td>
<td>Heavily doped</td>
</tr>
<tr>
<td></td>
<td>GaAs</td>
<td>GaAs</td>
<td>InGaAs</td>
<td>In0.6Al0.6As</td>
<td>GaAs</td>
</tr>
<tr>
<td>Donor Schottky</td>
<td>Heavily doped</td>
<td>Heavily doped</td>
<td>InyAl1−yAs</td>
<td>AlSb</td>
<td>InyAl1−yAs</td>
</tr>
<tr>
<td></td>
<td>AlxGa1−xAs</td>
<td>AlxGa1−xAs</td>
<td>Si δ doping</td>
<td>Si δ doping</td>
<td>Si δ doping</td>
</tr>
<tr>
<td>Spacer</td>
<td>AlxGa1−xAs</td>
<td>AlxGa1−xAs</td>
<td>InyAl1−yAs</td>
<td>AlSb</td>
<td>InyAl1−yAs</td>
</tr>
<tr>
<td>Channel</td>
<td>GaAs</td>
<td>GaAs</td>
<td>InyGa1−xAs</td>
<td>InAs</td>
<td>InyGa1−xAs</td>
</tr>
<tr>
<td>Buffer-1</td>
<td>GaAs</td>
<td>InP</td>
<td>AlSb</td>
<td>Al0.5Ga0.5</td>
<td>InAlAs</td>
</tr>
<tr>
<td>Buffer-2</td>
<td>InAlAs</td>
<td>Al0.7Ga0.3As</td>
<td>AlSb</td>
<td>Al0.5Ga0.5</td>
<td>InAlAs</td>
</tr>
<tr>
<td>Buffer-3</td>
<td>AlSb</td>
<td>GaAs</td>
<td>AlSb</td>
<td>In0.5Ga0.5</td>
<td>InAlAs</td>
</tr>
<tr>
<td>Buffer-4</td>
<td>GaAs</td>
<td>InP</td>
<td>GaAs</td>
<td>GaAs</td>
<td>GaAs</td>
</tr>
</tbody>
</table>

Table 5-3. Summary of HEMT structures.
5.5.2 “Hybrid” Lithography

To achieve high speed, low noise temperature, and high power-added efficiency at high frequencies, HEMTs require very short gate lengths. The intrinsic maximum frequency of operation, $f_{\text{max}}$, is the figure of merit used to evaluate HEMT performance which is given by

$$f_{\text{max}} = \frac{v_e}{2\pi L_g}$$ (5.5-1)

where $v_e$ is the electron velocity and $L_g$ is gate length. Fabricating very short gates is a challenge that requires well-developed lithography and pattern-transfer techniques.

Masks for the lithography of HEMTs are fabricated using a combination of electron beam lithography (EBL) [58] and photolithography. Although direct-write EBL is a low throughput exposure process, it facilitates accurate definition and alignment of sub-micron geometries while providing flexibility and fast turnaround for design iterations. EBL is widely and routinely used to produce gate dimensions of less than 0.25 $\mu$m [59–61]. To improve the wafer exposure throughput, optical lithography is used for the coarse features (>1 $\mu$m or 1000 nm), and the direct EBL is used only for the very short gates. This hybrid lithography has the advantage of both the high throughput of optical lithography and the high resolution and accuracy of EBL. The hybrid lithography scheme is also used for HEMT-based MMICs requiring sub-micron gates.

5.5.2.1 Frontside Processing—Device Isolation or Mesa Formation. The devices are isolated from each other by selective etching [62] of doped layers down to the buffer layer or all the way to the substrate. The isolation process involves a number of steps that include resist deposition, photolithographic exposure, development, semiconductor etching, and resist strip and cleaning. The etching results in formation of isolated islands of conducting epitaxial layers or mesas that are surrounded by semi-insulating buffer or substrate material. Device isolation is checked with a simple direct current measurement. This process also reduces the parasitic capacitances and back-gating, and it provides an insulating surface for MMIC passive components.

5.5.2.2 Ohmic Contact Formation. Since HEMTs are large-current and small-voltage devices, the saturation voltage and transconductance are very sensitive to the contact resistance. It is essential that extremely low contact resistances be formed to the 2-DEG that is situated approximately 300–1000 Å (30–100 nm) below the surface to fully utilize the HEMT’s potential.
After mesa formation, the ohmic contact areas are patterned using lithographic techniques and then metallized. Ohmic contact areas are either alloyed or non-alloyed [63]. The goal is to dope the surface of the semiconductor sufficiently high to assure that the dominant conduction mechanism between the contact metal and the semiconductor is field emission [64].

For non-alloyed contacts, a metal can be deposited directly on the ohmic area. To obtain good non-alloyed ohmic contacts, a very heavily doped Ge layer or low-band-gap material such as InGaAs is usually used for the capping layer.

The most commonly used elements for alloyed contacts are a judicious combination of gold (Au), germanium (Ge), and nickel (Ni). These are evaporated onto the patterned HEMT wafer to form ohmics for the source and drain electrodes. After removing the unwanted metal through a lift-off process (whereby the unwanted metal is lifted off by dissolving the underlying resist), the wafer is then thermally alloyed in an inert nitrogen (N2) atmosphere at a temperature between 320 and 450 deg C to form low-resistance ohmic contacts. A rapid thermal annealing (RTA) or furnace annealing technique is typically used for ohmic alloy [65]. The optimum ohmic surface morphology depends on the metallization composition, thicknesses, and alloy cycle. A good ohmic morphology also provides clean, sharply defined ohmic contact edges in the device channel so that the gate can be placed close to the source to minimize the source resistance.

5.5.2.3 Gate Formation. In addition to the short gate length, a small gate resistance is essential to the fabrication of HEMTs for high-gain, low-noise, and high-power applications. The most widely used gate cross-sectional structure is the T-shaped or mushroom-shaped gate formed using a multi-layer resist technique with E-beam lithography [66–69]. In this structure, the small footprint or bottom of the T defines the gate length, and the wider top of the T provides a low resistance.

A trilayer resist system, PMMA/P(MMA,MAA)/PMMA (PMMA is polymethylmethacrylate, and the other materials are copolymers of PMMA) is used to define T-gates [70]. The least sensitive resist is first deposited on the wafer. Then a sensitive resist is deposited, and finally, a thin, relatively insensitive resist is used to define a good lift-off mask. In addition, the trilayer resist system gives good control of the recess slot width.

Following gate lithography and resist development, the exposed HEMT channel area is recessed to achieve the desired channel current and threshold voltage prior to the gate metallization. The recess etching is performed using either a wet chemical etch or a reactive ion etching (RIE) [71,72] technique. The depth to which the gate is recessed is a critical parameter to the HEMT performance. The etching is discontinued when a target source to drain current
is achieved. Figure 5-6 shows the T-shaped resist cavity using the trilayer resist system.

After the recess, the wafer is then metallized, and the lift-off process is performed to form the metal gates. The metal used to create a Schottky barrier must adhere to the semiconductor and possess thermal stability. The gate metal is typically composed of several metal layers to decrease resistivity. The most commonly used gate metallization layers are titanium/platinum/gold (Ti/Pt/Au), with titanium/palladium/gold (Ti/Pd/Au), titanium/molybdenum/gold (Ti/Mo/Au) and chromium/palladium/gold (Cr/Pd/Au) as possible alternative compositions. Because the gates are very small, a scanning electron microscope (SEM) is used before and after gate metallization to measure the gate dimension and inspect for defects produced during the gate-formation process.

Submicron T-gates fabricated using this technique have demonstrated excellent mechanical stability, and they also exhibit extremely low gate resistances. However, for 0.15-μm or less gates, the T-gate resistance rapidly increases [73] and becomes significant compared to the source resistance. Thus for extremely short gate lengths, a trade-off between the gate resistance and the gate length must be made. An overview of the gate formation process for FETs can be found in Weitzel’s review article [74]. Figure 5-7 shows the lifted-off T-shaped gate using the PMMA/P(MMA,MAA)/PMMA trilayer resist system.

To increase the contact conductivity, to simplify bonding to the device and to add MMIC components, such as, inductors, capacitors, and transmission lines, it is necessary to add more metal. Addition of thicker metallization and higher levels of metallization requires additional processing steps such as, resist deposition, photolithographic exposure, development, thick metallization evaporation, lift-off, etc. Au is usually used because of its good conductivity in combination with another metal like Ti or Cr. The Ti or Cr layer, usually 200–1000 Å (20–100 nm) thick, is applied first to provide good adhesion to the III-V semiconductor.
5.5.2.4 Dielectric Deposition and HEMT Passivation. The device channel area is susceptible to surface damage, chemical and mechanical. Long-term degradation can occur through oxidation or particulate contamination and/or damage during handling and probing. Dielectric films such as polyimide, silicon nitride ($\text{Si}_3\text{N}_4$) and silicon dioxide ($\text{SiO}_2$) are commonly used to seal and protect the surface, keeping humidity, chemicals, gases, and particles away from sensitive areas of the device. Device protection or passivation (as it is named), requires a continuous, uniform, low-loss dielectric film. The polyimide film can be spun on the wafer, while $\text{Si}_3\text{N}_4$ and $\text{SiO}_2$ can be deposited using plasma-enhanced chemical vapor deposition (CVD).

For capacitors in a HEMT MMIC, a certain dielectric film thickness is needed to achieve the desired capacitance and to improve device and circuit reliability. A thick passivation improves reliability by eliminating pin-holes but reduces device performance by introducing extra parasitic input and feedback capacitance between the gate and drain. A thick passivation degrades the device noise figure, gain, and possibly power-added efficiency, especially at millimeter-wave frequencies. The final dielectric film thickness for HEMTs or MMICs is determined by a trade-off among the device and circuit reliability, performance, and capacitor requirements for the application.

5.5.2.5 Backside Processing. The last steps in HEMT fabrication are wafer thinning, via-hole formation, and dicing. The substrate is thinned to reduce its thermal impedance, to improve its mechanical handling, and to facilitate
transmission line and via-hole formation. A final substrate thickness of 3 to 4 mils (0.0762–0.1016 mm) is typically used for microwave low-noise HEMTs and MMICs. Uniform wafer thinning is the key to a high-yield via-hole process. Via-holes provide low-inductance source grounding which is critically important for high-frequency power HEMTs and low-noise and power MMICs.

The wafer thinning requires mounting of the wafer, usually with the wax frontside down, on a carrier such as glass, quartz, or silicon. Thinning the wafer is accomplished by mechanically lapping the backside of the wafer with a slurry of water and grit, usually silicon carbide (SiC), between the wafer and a flat plate, usually glass. To obtain a smoother surface, the wafer can either be further lapped with a finer grit or chemically polished.

After the wafer has been thinned and polished, the backside is patterned to open holes corresponding to the desired via locations. The via-hole pattern is defined with photolithography using infrared light for aligning the backside pattern to the frontside pattern. (Most III–V semiconductors are transparent to infrared light.) Via-holes can be formed with a wet-chemical etch or with RIE techniques. Compared to the wet-chemical process, the RIE via-hole process is less sensitive to the uniformity of the final substrate thickness and also provides smaller vias with controlled etch profiles. Figure 5-8 shows typical via-holes etched through a 4-mil-thick (10-mm) substrate in a HEMT using the wet-chemical approach. After the via-hole formation, the backside of the wafer is metallized. The chips are then separated through a wafer sawing or a scribe-and-break technique. Figure 5-9 shows a typical HEMT chip after scribe-and-break. Chips with mechanical damage or surface contamination are screened out using both optical microscope and SEM inspections. The HEMT chips are finally electrically tested for dc and RF performance.

To summarize, in order to obtain uniform, high-yield, and reproducible HEMTs, the following are the most critical fabrication issues:

1) Uniform and low-defect density HEMT epilayers
2) Formation of uniform submicron gates on large-size wafers
3) Uniformity of the HEMT gate recess etch.

Each of these areas represents a significant technical challenge and requires a substantial investment in equipment and technology development to insure future HEMT device performance improvements.
Fig. 5-8. Via-holes etched through a 4-mil-thick (10-µm) GaAs substrate using a wet etch.

Fig. 5-9. Typical HEMT chip after scribe and break. (Note: all chips are inscribed with data identifying position on the wafer, gate number, and gate width.)
5.6 HEMT Noise Modeling

To develop ultra-low noise-microwave amplifiers for cryogenic applications, one must have a reliable, reproducible source of state-of-the-art cryogenic devices and the capacity to accurately characterize and model them at the wafer or device level at cryogenic temperatures. Currently, small-signal, semi-empirical, circuit models have proven to be the best way to simulate both the noise and scattering parameters of low-noise HEMTs and FETs. Although this approach is quite successful for the circuit design of cryogenic, low-noise amplifiers, it provides only modest feedback data required for iterative HEMT materials and device optimization.

5.6.1 Noisy Linear Two Port Model

For purposes of circuit modeling and device characterization, a noisy linear two-port device can be represented as a noiseless linear two-port device with the noise sources at the input and/or output [75,76]. Depending upon the utility of the representation, the internal (voltage and/or current) noise sources can be placed at the input or output port of the noiseless network. Figure 5-10 shows a convenient representation that leads to four noise parameters \( T_{\text{min}} \), \( R_{\text{opt}} \), \( X_{\text{opt}} \), and \( R_n \) that can be determined from the measurement of noise temperature as a function of input match, \( Z_g \). It consists of a series noise voltage \( (e_n) \) and shunt noise current \( (i_n) \) sources at the input [77]. In this representation, the two-port device’s noise parameters are given by the equivalent noise resistance \( (R_n) \), noise conduction \( (g_n) \), and correlation coefficient \( (r) \),

\[
R_n = \frac{\langle e_n e_n^* \rangle}{4kT_0B} \quad (5.6-1)
\]

\[
R_n = \frac{\langle e_n e_n^* \rangle}{4kT_0B} \quad (5.6-2)
\]

Fig. 5-10. ABCD representation of noisy two-port using voltage and current noise sources at the input.
and

\[ r = \frac{\langle e_n^* i_n \rangle}{\sqrt{\langle e_n^* e_n \rangle \langle i_n^* i_n \rangle}} \]  \hspace{1cm} (5.6-3)

where \( T_0 = 290 \) K, \( k \) is Boltzmann’s constant, and \( B \) is the noise bandwidth.

The noise temperature \( (T_n) \) of the two-port device is driven by a generator impedance \( (Z_g) \) and is given by the expression

\[ T_n = T_{\text{min}} + \frac{T_0 g_n}{R_g |Z_g - Z_{\text{opt}}|^2} \]  \hspace{1cm} (5.6-4)

where \( Z_{\text{opt}} \) is the optimal generator impedance that yields a minimum noise temperature and \( Z_g = R_g + jX_g \) is the generator impedance (where \( R_g \) is the real component and \( jX_g \) is the imaginary component). The relationship between the first set of noise parameters and those in the above expression is given by the following equations

\[ X_{\text{opt}} = \frac{\text{Im}(C)}{g_n} \]  \hspace{1cm} (5.6-5)

\[ R_{\text{opt}} = \sqrt{R_n / g_n - X_{\text{opt}}^2} \]  \hspace{1cm} (5.6-6)

and

\[ T_{\text{min}} = 2T_0 \left[ g_n R_{\text{opt}} + \text{Re}(C) \right] \]  \hspace{1cm} (5.6-7)

where

\[ C = r \sqrt{R_n g_n} \]  \hspace{1cm} (5.6-8)

In principle, the above noise parameters (\( Z_{\text{opt}}, T_{\text{min}}, \) and \( R_n \)) for FET and HEMT devices can be determined by measuring the noise temperature for four or more different known source impedances at a given frequency. In practice, however, since there are errors associated with the source impedance and the noise temperature measurements, additional measurements are usually taken to improve the statistics.
Next, the above equations are rewritten in terms of parameters (A, B, C, and D) that linearize the noise figure or noise temperature expression [78]. The noise parameters are then computed by linear regression analysis of the data. The noise parameters, along with the scattering parameters, can then be utilized for the optimum design of an amplifier circuit.

5.6.2 Semi-Empirical Small Signal Noise Models

In principle, the measured small signal noise and scattering parameters at relevant bias settings as well dc characteristics can be tabulated and utilized for the low-noise amplifier design. The main advantage of a tabular model is that no equivalent circuit needs to be extracted or optimized. However, this approach has some clear disadvantages. First, it may require a large amount of memory for multiple bias settings. Secondly, tabulated data cannot be extrapolated to higher frequencies where measurements may not be possible, and thirdly, tabular models cannot be scaled.

In contrast, semi-empirical models require some experimentally determined fitting factors but do not have these limitations. The goal of semi-empirical circuit models is to find the minimum number of idealized, frequency independent circuit components that reasonably represent complex physical processes so that the equivalent circuit accurately models noise and scattering parameters.

Analytical models that consider fundamental semiconductor steady-state transport properties usually only treat thermal noise within the channel, and thus only model the intrinsic HEMT. These models are progressively more complex treatments of van der Ziel’s original work [79,80]. The numerical noise model approach taken by Cappy et al. [81] takes into account electron dynamics and adequately explains noise temperature results at room temperature. Joshin’s analytical one-dimensional electron transport noise model [82] also reasonably explains noise temperature results at room temperature and suggests that drain noise current is nearly canceled by induced-gate-noise current due to the asymmetric distribution of noise generation along the HEMT channel. However, the dependence of the measured noise temperature on device parasitics (pad capacitances, inductances, and resistances) and input circuit impedance complicates the full evaluation of numerical and analytical models.

5.6.2.1 PRC Model. The PRC model, based on the work of Pucel et al., is a current source model, where the current sources are connected to the input and output ports of the intrinsic transistor [83,84]. The current sources are correlated and the correlation is imaginary. Figure 5-11 shows the intrinsic HEMT equivalent circuit model with the PRC model current noise sources. The advantages of the PRC model are its close connection to the physical processes
in the device and simplification of the model extraction process. The model is named after the parameter names $P$, $R$, and $C$ and are given below along with the minimum noise figure, $F_{\text{min}}$:

\[
P = \frac{\langle i_d i_d^* \rangle}{4kT_0 B g_m} \tag{5.6-9}
\]

\[
R = \frac{\langle i_g i_g^* \rangle g_m}{4kT_0 B (\omega C_{gs})^2} \tag{5.6-10}
\]

\[
C = -j \frac{\langle i_g i_d^* \rangle}{\sqrt{\langle i_g i_g^* \rangle \langle i_d i_d^* \rangle}} \tag{5.6-11}
\]

and

\[
F_{\text{min}} = 1 + 2 \sqrt{P R (1 - C^2)} \frac{f}{f_T} + 2 g_m R_i P \left( 1 - C \sqrt{\frac{P}{R}} \right) \left( \frac{f}{f_T} \right)^2 \tag{5.6-12}
\]

5.6.2.2 Fukui Model. The most well established model for device and circuit optimization is the semi-empirical one developed by Fukui [85]. In the Fukui model, the noise parameters are simple frequency dependent functions of the equivalent small-signal intrinsic circuit elements (transconductance, $g_m$, gate-to-source capacitance, $C_{gs}$, and source and gate resistances, $R_s$ and $R_g$). These circuit elements are (in turn) analytic functions of the device’s geometrical (e.g., gate and channel dimensions) and material (e.g., doping
concentration and active channel thickness) parameters. The semi-empirical approach of Fukui yields the following expressions for the noise parameters

\[ T_{\text{min}} = \frac{k_1 \omega T_0 C_{gs}}{2\pi} \sqrt{\frac{R_g + R_s}{g_m}} \] (5.6-13)

\[ R_n = \frac{k_2}{g_m} \] (5.6-14)

\[ R_{\text{opt}} = k_3 \left( \frac{1}{4g_m} + R_g + R_s \right) \] (5.6-15)

and

\[ X_{\text{opt}} = \frac{2\pi k_4}{\omega C_{gs}} \] (5.6-16)

where \( k_1, k_2, k_3, \) and \( k_4 \) are fitting factors that are determined experimentally. Although the Fukui model is widely used by device designers and served to guide the development of the first cryogenic HEMT device for the Voyager 2 encounter with Neptune, the model is equivalent-circuit dependent and provides little insight into the physics of noise in HEMTs.

**5.6.2.3 Pospieszalski Noise Model.** The Pospieszalski model uses small signal circuit elements that yield closed-form expressions for the noise parameters [86]. This model introduces frequency-independent equivalent temperatures \( T_g \) and \( T_d \) for the intrinsic gate resistance \( R_i \) and drain resistance \( R_{ds} \), respectively. The noise processes are modeled by \( R_{ds} \), and \( T_d \) is the only free parameter, while \( T_g \) is taken to be the ambient temperature of the device. The equivalent noise model for the intrinsic HEMT is shown in Fig. 5-12 while the noise correlation relations are given below [87]

\[ R_{ds} = \frac{\langle i_d i_d^* \rangle}{4kT_d B} \] (5.6-17)

\[ R_i = \frac{\langle i_g i_g^* \rangle}{4kT_g B} \] (5.6-18)
The noise parameters for the above intrinsic circuit are given by the following expressions

\[
R_{\text{opt}} = \sqrt{\frac{T_g R_i R_{ds} g_m}{T_d \omega C_{gs}}} + R_i^2 \quad (5.6-20)
\]

\[
X_{\text{opt}} = \frac{1}{\omega C_{gs}} \quad (5.6-21)
\]

\[
T_{\text{min}} = 2 \frac{\omega C_{gs}}{g_m} \sqrt{\frac{T_d T_g R_i}{R_{ds}}} + \left( \frac{T_d R_i \omega C_{gs}}{R_{ds} g_m} \right)^2 + 2 \frac{T_d R_i}{R_{ds}} \left( \frac{\omega C_{gs}}{g_m} \right)^2 \quad (5.6-22)
\]

and

\[
R_n = \frac{T_{\text{min}}}{T_0} R_i + \frac{T_d}{T_0} \frac{1}{R_{ds} g_m} \left( 1 + \left( \frac{\omega C_{gs} R_i}{g_m} \right)^2 \right) \quad (5.6-23)
\]

The utility of this model is that it allows prediction of the noise parameters for a broad frequency range from a single frequency noise-parameter measurement at a given temperature. Although the Pospieszalski model only considers thermal noise sources and does not take into account the correlated noise between the gate and the drain, it is an accurate model for high-quality devices operated at low-noise bias. Additionally, for devices operated at higher drain currents, \( T_g \).
is taken to be higher than the ambient temperature; and gate leakage current can be modeled with a resistor across $C_{gs}$ and $R_i$ at an elevated temperature [88].

### 5.6.2.4 Monte Carlo HEMT Noise Model

Although considerable research has been conducted on the noise performance and theory of HEMTs, a noise model that is useful for cryogenic-device optimization and circuit design is not yet available. However, Monte Carlo techniques using microscopic theory have recently been successfully applied to calculate and predict the noise parameters of HEMTs at room temperature [89].

The most valuable feature of Monte Carlo techniques is that these techniques enable investigation of the physical origins of noise in semiconductor devices and allow differentiation between noise temperature and electron temperature [90]. The essential feature of the Monte Carlo noise temperature model is to follow the evolutionary motion of the charge carriers in time domain while taking into account all of the important microscopic scattering mechanisms associated with the semiconductor material (ionized impurities, inter-valley transitions, phonons, alloy, electron-electron, etc.) [91]. The Monte Carlo simulator calculates (1) instantaneous velocities and energies, (2) mean velocity and mean energy from the instantaneous values, (3) instantaneous velocity fluctuations from (1) and (2), and then (4) the spectral density of velocity fluctuations. The intrinsic noise temperature, $T_n$, is calculated from the spectral density of velocity fluctuations, $S_v(f)$, and electron temperature, $T_e$, is calculated from the average energy using the equipartition principle as shown below

$$T_n(f) = \frac{qS_v(f)}{4k\mu_D(f)} \quad (5.6-24)$$

and

$$T_e = \frac{2}{3k} \langle E \rangle \quad (5.6-25)$$

$$S_v(f) = 4\int_0^\infty C(t)\cos(2\pi ft)dt \quad (5.6-26)$$

$$C(t) = \langle \delta v(t')\delta v(t+t) \rangle \quad (5.6-27)$$

where $\mu_D$ is the differential mobility, $C(t)$ is the autocorrelation function of velocity fluctuations and $\delta v(t) = v(t) - \langle v \rangle$. In fact, Pantoja et al. [90]
demonstrated very good agreement between measured and modeled noise temperature \((T_n)\) at 300 K and 77 K for GaAs at 10 GHz.

Furthermore, Mateos et al. [89] performed a complete Monte Carlo analysis on a low-noise 0.1-\(\mu\)m T-gate AlInAs/GaInAs HEMT. The modeled dc and rf properties showed exceptionally good agreement with measurements. The model included effects such as degeneracy, surface charges, T-shape of the gate, presence of dielectrics, and contact resistances. Moreover, the extrinsic parameters of the device were added to the intrinsic small-signal equivalent circuit, allowing a realistic calculation of the dc characteristics and the noise and scattering parameters. Although, the simulations take hours on a personal computer, the reliability allowed by this Monte Carlo simulator will enable faster and cheaper optimization of HEMT devices.

### 5.7 LNA Development

The development and demonstration of cryogenic, InP HEMT-based front-end low-noise amplifiers for the DSN also require accurate LNA component characterization and modeling from 1 to 100 GHz at physical temperatures down to 12 K. The characterization and modeling starts with the individual HEMT chip, RF and DC bias components, proceeds to the multi-stage HEMT LNA module, and it culminates with the complete cryogenic front-end receiver package for the antenna.

#### 5.7.1 Device Characterization—Cryogenic Probe Station

The development of a complete on-wafer cryogenic microwave measurement system has been primarily driven by the need for

1) Greater understanding of the device physics in advanced high speed transistor technologies

2) Continued advancement of cryogenic LNAs with noise temperatures less than five times the quantum limit \((T_n < 5hf / k)\) for ground-based and space-based applications

3) Hybrid and monolithic microwave integrated-circuit (MMIC) semiconductor-superconductor circuits

The cryogenic microwave system uses coplanar waveguide probes in a vacuum station coupled to a vector network analyzer for scattering parameter measurements, and a noise meter and noise test set with a noise system for microwave noise parameter measurements. The cryogenic probe measurement system, shown schematically in Fig. 5-13, contains ports for RF cables, thermometers, vacuum pumps, dry nitrogen back-fill lines, coplanar probes with manipulators, and a closed-cycle refrigerator cold head. The probe body rests on a copper block attached to a fiberglass post. The fiberglass reduces the
thermal load, and copper braiding from the cold head thermally anchors the probe to the 12-K cold station, assuring sample temperatures of 12 to 20 K. The mechanical and thermal stability of the wafer stage is established by supporting it on fiberglass posts above the cold head and thermally anchoring it to the cold station with flexible copper braids.

The most important feature of this design is the incorporation of a closed-cycle helium refrigeration system. The first successful designs of on-wafer cryogenic systems used open-cycle cooling to reduce start-up costs and avoid mechanical vibrations. However, for long term use at the rate of one cool down per week, a closed-cycle system is significantly less expensive.

Decoupling and damping of the vibrations from the cold head to the probe station are accomplished with a two-dimensional bellows and vibration mounts. This system allows small-signal microwave measurements from dc to 40 GHz over a physical temperature range of 16 to 300 K. Since the microwave hardware is insulated by vacuum, there is no frost buildup or large thermal gradients, resulting in a system that is accurate, reliable, and flexible (active and passive discrete devices, as well as MMICs, can be measured). Figure 5-14 is a photograph of the cryogenic probe test chamber showing the input (at left) and the output (at right) coplanar microwave probes with a calibration standard and test HEMT devices epoxied to an alumina substrate (near the center).
5.7.2 Device Characterization—Cryogenic Probe Station Calibration

The key to accurate on-wafer microwave and millimeter-wave measurements is proper establishment of the electrical reference plane. The reference plane can be determined with either the line-reflect-match (LRM) or the short-open-load-thru (SOLT) calibration method utilizing an impedance standard substrate (ISS), which is available from Cascade Microtech [92]. The LRM method requires fewer standards, and the reflect standards need not be well known. In addition, experience has shown that the LRM calibration is slightly better in accuracy than is the SOLT at cryogenic temperatures. In the SOLT method, the short standard introduces uncertainty in the reference plane location because of sensitivity to probe tip placement. The LRM method obviates this problem by replacing the short with an open and by having the probe tip held approximately 10 mils (0.25 mm) above the substrate during the calibration sequence.

The measurement accuracy is also directly related to the calibration conditions. Thermal gradients across the gold-plated ceramic probe tips and
HEMT Low-Noise Amplifiers

Coax-to-coplanar transitions [93] alter the electrical characteristics of the measurement lines. The process of cooling the sample in the laboratory can also produce changes in the calibration. The network analyzer typically requires a new calibration if the ambient laboratory temperature varies by greater than 1 deg C. The proximity of the cooled probe system to the network analyzer can change the ambient environment (both temperature and humidity). The combined results of these effects are appreciable errors in cryogenic temperature measurements. For example, room-temperature calibrations have been shown to introduce as much as a 20-percent [94] error in the cryogenic measurement. In addition, measured results have been reported with room-temperature calibrations with moding effects (deviations from one-pole roll-off) [93–96].

Since the microwave probe offers a significant thermal load to the device under test (DUT), it is evident that the chuck and device temperatures are different. For example, with a chuck temperature of 20 K and the probes contacting the DUT, a DUT temperature as high as 50 K has been observed. This large temperature differential affects calibration and skews the interpretation of data collected at different device temperatures.

The solution to maintaining calibration integrity and achieving low sample temperatures is to thermally anchor the probe body and perform cryogenic calibrations. By thermally anchoring the probe to the cold head at 12 K, the thermal load to the DUT is minimized. The remaining microwave hardware (connectors, cables, and input to the automatic network analyzer (ANA)) are thermally isolated via vacuum and stainless steel hardware. The thermally anchored probe can be calibrated at specific temperatures during the measurement cycle. This eliminates the problem of moding and allows accurate correlation of DUT temperature and measured characteristics [97,98].

For this initial investigation of on-wafer noise parameter measurements at cryogenic temperatures, only the probe tips are cooled while the impedance state generator and solid-state noise source (both commercially available) are kept at room temperature (several wavelengths away from the DUT). In this configuration, the input losses introduce noise comparable to or greater than the noise of the DUT and reduce the range of available impedance states. For example, in the frequency range of 2 to 18 GHz for cryogenic temperatures, the worst-case noise temperature error is ±25 K, while device noise temperatures are typically under 10 K. Although this configuration does not provide accurate single-frequency noise parameter measurements, it does provide for fast and efficient broadband (2- to 18-GHz) on-wafer measurements [99].

The most accurate and repeatable method of measuring noise parameters at cryogenic temperatures is to place the impedance generator within a wavelength of the DUT input. The equivalent noise temperature of the noise source must also be comparable to the DUT noise temperature. This approach
would, however, require development of a cryogenic noise generator and noise source.

5.7.3 Device Characterization Measurements and Models

The InP HEMTs developed by TRW for the DSN are grown by molecular beam epitaxy on 3-in. (7.6-cm) semi-insulating InP wafers. The cross-section of this device is shown in Fig. 5-15. First, to inhibit impurity diffusion into the active region and to improve carrier confinement, two buffer layers (one of InAlAs and the other InP) are grown on the semi-insulating InP wafer. Then the active undoped In$_{0.65}$Ga$_{0.35}$As (65% indium concentration) layer is grown followed by another spacer layer of InAlAs to further improve carrier confinement and to reduce donor ion scattering. The donor Si atoms are added in a single atomic layer in the next undoped InAlAs Schottky layer. The last two layers are n-type InGaAs cap layer and a heavily doped InGaAs layer to provide ohmic contacts. The devices are then passivated with a thin SiN layer.

A key step in the process of developing InP HEMT hybrid and MMIC LNA modules is the accurate, broadband, cryogenic characterization and modeling of active and passive circuit components. From 0.05 to 40 GHz, device and component measurements covering room to cryogenic temperatures are made with a cryogenic, coplanar waveguide probe station. To enhance and
extrapolate component models up to 100 GHz, a 2½ dimensional electromagnetic (2½ DEM) simulator was used.

There are three main methods for small-signal parameter extraction:
1) Complete parameter extraction from calculations on the measured S-parameters [100]
2) Parameter extraction based upon iterative computer optimization routines
3) A combination of (1) and (2) where as many elements as possible are quickly identified, providing constraints to the software optimization routine that serves as the validity check.

All of the above methods can be applied to the standard HEMT small-circuit model, the hybrid π-circuit topology [101]. The small-signal elements in Fig. 5-16 are broken down into the intrinsic and extrinsic elements. The intrinsic elements are the transconductance ($g_m$), output resistance, ($R_{ds}$), gate-source capacitance, ($C_{gs}$), gate-drain capacitance, ($C_{gd}$), drain-source capacitance, ($C_{ds}$), gate-source resistance, ($R_i$), and delay time ($\tau$). These elements are bias dependent and important to the understanding of device behavior. The extrinsic elements are independent of bias and include the three terminal inductances ($L_g$, $L_d$, and $L_s$), the contact resistances for the three terminals ($R_g$, $R_d$, and $R_s$), and the parasitic pad capacitances ($C_{pg}$ and $C_{pd}$).

The simplest and most straight-forward method is the complete parameter extraction method. It has been shown that, with a sequence of microwave and dc measurements, each term can be uniquely determined [100]. A so-called

![Fig. 5-16. Pospieszalski HEMT circuit model showing both extrinsic and intrinsic circuit elements.](image)
shell technique can be applied to remove the inductances, the pad parasitics, and the contact resistances. For this technique, the S-parameters of the device are measured for a variety of drain-to-source voltage conditions. By using the simplifying assumptions that \((\omega C_{gs} R_i)^2 < 0.01\) and that \(\omega \tau << 1\), one can determine the parasitic elements uniquely, leaving only the intrinsic circuit to be determined. The drawbacks to this method are that several measurements must be made under different bias conditions and simplifying assumptions must be made, increasing the uncertainty of the extracted parameters.

The second method of extraction is to simply fit the measured S-parameters to the equivalent circuit by iteratively solving for the individual elements. The disadvantage to this technique is that, with such a large parameter space, it is very difficult to uniquely identify each element. This frequently produces unrealistic values for a number of the circuit elements. The most desirable and accurate method is to apply the techniques from the complete parameter extraction method coupled with small-signal circuit optimization routines that serve as a validity check. The important parasitics are estimated by measuring a simple test structure and by independent dc measurements.

Although a test structure for the parasitic elements was not available, the method used is based on the methods developed by [102–104]. Extraction starts with determination of the parasitic pad capacitances with the HEMT biased under ColdFET \((V_{ds} = V_{gs} = 0)\) and pinch off conditions (i.e., \(V_{ds} = 0\) and \(V_{gs} < V_{\text{pinch-off}}\)). Next, the device is biased under ColdFET and full channel conditions (i.e., \(V_{ds} = 0\) and \(V_{gs} > 0\)). The gate resistance, sheet resistivity, gate width, and gate length are used to estimate the initial values of parasitic resistances. This is sufficient to extract the rest of the extrinsic components. The intrinsic components of the active HEMT are calculated by fitting the measured S-parameters to the HEMT model. Finally, starting values of \(T_g\) and \(T_d\) are estimated from measured room temperature device noise parameters and then recalculated based on room temperature and cryogenic LNA module noise and gain measurements [105].

The widely used Pospieszalski HEMT noise model [86], shown in Fig. 5-16, possesses physically realizable elements, and it is symmetric and sufficiently broadband for this work. Figure 5-17 shows the agreement between the measured and modeled scattering parameters at 18 K with (+) illumination and without (−) illumination at low noise bias, \(V_{ds} = 0.8\) v, and \(I_{ds} = 2.0\) mA, and \(V_{gs} = 0.08\) v.
Fig. 5-17. Models versus measured cryogenic (a) gain and (b) reflection of 32-GHz HEMT at low noise bias with a frequency of 1 to 35 GHz (Note: Model and measurement are in excellent agreement).
5.7.4 Passive Component Characterization Measurements and Models

From 1 to 40 GHz, the 2½ DEM simulator was used to verify our cryogenic measurements on passive components and to refine and extend our user-defined cryogenic models employed in MMICAD (Optotek, Ltd.) to 100 GHz. Although measurements to 40 GHz are sufficient for the 8.4-GHz LNA module modeling and design, it is inadequate for the 32-GHz LNA module modeling and design. The maximum frequency of operation, $f_{\text{max}}$, for the 32-GHz device approaches 150 GHz. Thus, at 32 GHz, especially in regards to stability, it is critical to have accurate component models up to 100 GHz. Figure 5-18 shows the measured and modeled response of a 1-kΩ bias circuit resistor at cryogenic temperatures.

5.8 LNA Modeling and Characterization

The design approach and fabrication process for both the 8.4-GHz and the 32-GHz LNA modules are essentially the same with some practical differences.

Fig. 5-18. Measured and modeled S-Parameters of 1-kΩ Bias Resistor at 18K.
At 8.4 GHz only three stages are needed, since there is sufficient gain per stage, while the 32-GHz module requires four stages. The 8.4-GHz module uses TRW four-finger 300-μm gate width InP HEMTs in all three stages, while the 32-GHz module uses TRW four-finger 80-μm gate width InP HEMTs in all four stages. Additionally, WR-28 wave guide input and outputs are used at 32 GHz instead of coaxial k-connectors to further reduce RF losses.

The LNA module design goal is to minimize the noise temperature at the DSN band of operation while at the same time maintaining unconditional stability both inside and outside the module’s bandwidth. The 8.4-GHz LNA is designed to be unconditionally stable from 0 to 40 GHz, while the 32-GHz LNA is designed to be unconditionally stable from 0 to 100 GHz.

Since all of the HEMT devices used for the LNA designs are unstable (i.e., μ-factor <1 [106] over their usable gain bandwidth), the first step is to stabilize the device at the LNA module band of operation without significantly increasing the device noise temperature. For example, for the 8.4-GHz LNA module design, the first-stage device is first stabilized near 10 GHz by a judicious choice of gate, drain, and source bond wire lengths. Next, the device gate and drain bias networks are used to load the device and control the stability below 10 GHz and from 10 to 40 GHz, respectively. The rest of the stages are similarly optimized. The loaded biased devices then serve as the fundamental circuit building block. A similar procedure is used for the 32-GHz LNA module design. Figure 5-19 shows how the device bond wires and bias network affect and help control the device stability, while Fig. 5-20 shows the trade-off, an associated increase in noise temperature for the 8.4-GHz first-stage device.

![Figure 5-19. Bondwire and bias circuit stability (μ-factor) affects for 8.4-GHz HEMT at physical temperature of 16K.](image-url)
Then, each of the stages is iteratively matched to its optimum source and load impedance. At 8.4 GHz the first and second stage are optimized for noise performance while the third stage is optimized for gain and output match. Following initial optimization, the LNA model is loaded at the input and the output with the string of passive microwave components that will be implemented for field use, and the bond wire lengths re-optimized.

A photograph of the 8.4-GHz LNA module is shown in Fig. 5-21. The module carrier is gold-plated brass, the input, inter-stage and output matching circuits are etched on Cuflon, and the dc blocking and bias circuitry use surface-mount thin-film resistors and capacitors.

The completed LNA modules are then characterized in a cryogenic testbed at a physical temperature of 12 K using the cold attenuator method [107]. A photograph of the 32-GHz testbed is shown in Fig. 5-22. In this technique, a 20-dB attenuator connected to the LNA module input is cooled along with the module. The cooled attenuator serves as the cold noise source when the hot-noise source, noise diode, is turned off and eliminates impedance-match errors associated with the noise diode on–off states. The noise and gain are automatically measured using a commercial noise diode and noise figure meter. At 8.4 GHz, the gain measurement error is approximately ±0.1 dB, and the noise error is ±0.3 K, while at 32 GHz the errors are about five times higher. Figure 5-23 shows the measured and modeled noise and gain performance of the 8.4-GHz LNA, while Fig. 5-24 shows a similar plot for the 32-GHz LNA module. The module is subsequently cooled without the attenuator, and the output is monitored for oscillations or instabilities with a spectrum analyzer as the input impedance is varied.
Fig. 5-21. Photograph of three-stage 8.4 GHz InP LNA module (JPL part number JPL 2000-10).

Fig. 5-22. Photograph of the 32-GHz LNA cryogenic testbed.
Fig. 5-23. Measured and modeled noise and gain of an 8.4-GHz LNA module (LNA module 1 noise/gain at \( T_{\text{physical}} = 12.5 \) K model versus measurement).

Fig. 5-24. Measured and modeled noise temperature and gain of a 32-GHz LNA module.
5.9 Subsystem Measurements

Additional passive microwave components (filters, isolators, adapters, couplers, and polarizers) are required to implement an LNA module in a DSN antenna due to the deleterious effects of radio-frequency interference (RFI), the need for calibration signals, and the need for a redundant receive capability. Since significant effort is devoted to minimizing the LNA module’s noise temperature, an equivalent effort must be expended to minimize the noise temperature contribution of these components to meet the DSN low-noise receiver specifications.

A useful expression to determine the noise temperature contribution of passive two-port networks can be derived from the noise temperature function of the noise wave matrix representation for a passive network [108]. When the network is placed at the input of an LNA, the effective input-noise temperature, $T_e$, of the cascaded pair is given by the following expression [109]:

$$T_e = \frac{[(L-1)+\Gamma L^2]T_L + LT_{LNA}}{(1-\Gamma L^2)}$$  \hspace{1cm} (5.9-1)

where $L$ = loss ratio, $\Gamma_L$ = reflection coefficient, and $T_L$ = physical temperature of the passive network, while $T_{LNA}$ = LNA module noise temperature. Hence, in order to minimize the noise temperature contribution of the passive network, it must be well matched, be low-loss, and be kept at the lowest physical temperature possible.

Following complete characterization, the LNA module is integrated and measured with the necessary filters, isolators, and adapters. This cascaded network is placed in a larger testbed and characterized once again using the cold attenuator method described above.

A photograph of the cascaded LNA network is shown inside the testbed in Fig. 5-25. Noise temperature and gain measurement results are shown in Fig. 5-26.

Finally, the LNA module and components are integrated into the closed cycle refrigerator (CCR) package and characterized using the DSN’s ambient load/cold sky method [110]. A photograph of the LNA CCR package is shown in Fig. 5-27. (Current systems use a CCR built by Sumitomo Heavy Industries Ltd., which provides 1.5 W of cooling capacity at a physical temperature of 4.2 K.) This method uses the sky as the cold noise source and ambient load as the hot noise source placed over a calibrated feed horn [111] to determine the LNA CCR-package noise temperature. Noise temperature measurements (right-hand circularly polarized (RCP) and left-hand circularly polarized (LCP)) of the DSN LNA CCR package are shown in Fig. 5-28. These measurements are referenced to the room temperature waveguide input flange.
Fig. 5-25. Photograph of the cascaded 8.4-GHz LNA assembly in testbed.

Fig. 5-26. Cascaded LNA assembly noise and gain (model versus measured) at 9.6-K physical temperature.
Once implemented in the field these front-end receivers demonstrated operating system noise temperatures of 17 K at 8.4 GHz on a 70-m Cassegrain antenna and 39.4 K at 32 GHz on a 34-m beam waveguide antenna, both at zenith [112]. (To date, three more 8.4-GHz, two more 32-GHz comparable
front-end receivers, and six X/X/Ka (X-transmit, X-receive, and Ka-receive) HEMT/CCRs systems have been delivered to the DSN.) At 8.4 GHz, the $T_{op}$ for these units has been within 1 K of the predicted value, while at 32-GHz it has been within 2.5 K. This close agreement between predicted and measured performance is a testament to the measurement and modeling accuracy required to successfully develop these ultra-low noise cryogenic InP-HEMT-based CCR receive systems.

5.10 Conclusion

The work reported here is the by-product of a program at JPL to develop cryogenic InP HEMTs and MMICs for both ground-based and spaceborne radiometers and receivers. This work is thoroughly investigating the device parameters that will yield the best InP HEMTs. This study investigated a variety of indium concentrations, dopant concentrations and profiles, spacer and buffer layer thicknesses and compositions, and device geometry variations (shorter length gates and/or multiple gates, sources, and drains). In summary, the device types successfully fabricated to date are of five different indium concentrations (53, 60, 65, 70, and 80 percent) with two different gate dimensions (0.1 and 0.07 μm) [113].

In order for the noise temperature of a cryogenic, 32-GHz HEMT LNA module to drop to 3 K, the HEMT maximum frequency of operation, $f_{\text{max}}$, should exceed 500 GHz. The $f_{\text{max}}$ of the InP device used for the 32-GHz modules was 150 GHz. The best TRW InP devices that are currently being tested have $f_{\text{max}}$'s of 250 to 300 GHz. TRW device research is aimed at pushing $f_{\text{max}}$ beyond 300 GHz.

Although the emphasis of this chapter is technical, it must be pointed out that the partnership among the Jet Propulsion Laboratory (JPL), the Georgia Institute of Technology (GIT), and TRW is really the corner stone of this work. This partnership was sustained for several years and provided the right technical mix to develop all of the key components required for the successful development of 8.4- and 32-GHz state-of-the-art, ultra-low-noise cryogenic, InP HEMT-based CCR receive systems for the DSN. In summary, the key contributions of each of the partners were cryogenic, LNA module design and characterization (JPL), cryogenic, on-wafer noise parameter measurements and HEMT modeling (GIT), and state-of-the-art cryogenic InP HEMTs (TRW). Organizations seeking to upgrade their HEMT state-of-the-art need to have this complete a suite of capabilities and a comparable long-term commitment to the development process.
References


[40] U. K. Mishra, A. S. Brown, and F. E. Rosenbaum, “DC and RF Performance of 0.1 μm Gate Length Al0.48In0.52As/Ga0.38In0.62As HEMT with Optimized Transport Parameters,” IEEE Technology Digest of International Electron Devices Meeting (IEDM), San Francisco, California, pp. 180–183, December 1988.


[47] L. D. Nguyen, “650 A Self-Aligned Gate Pseudomorphic Al0.43In0.57As/Ga0.20In0.80As High Electron Mobility Transistors,” *IEEE Electron Device Letters*, vol. 13, pp. 143–145, March 1992.


[50] P. Win, Y. Druelle, P. Legry, S. Lepilliet, A. Cappy, Y. Cordier, and J. Favre, “Microwave Performance of 0.4 μm Gate Metamorphic In0.20Al0.71As/In0.30Ga0.70As HEMT on GaAs Substrate,” *Electronics Letters*, vol. 29, pp. 169–170, 1993.


[57] K. van der Zanden, M. Behet, and G. Borghs, “Comparison of Metamorphic InGaAs/InAlAs HEMT’s on GaAs with InP Based LM HEMT’s,” *Proceedings of the GaAs International Conference on Gallium Arsenide Manufacturing Technology (MANTECH)* (Vancouver, Canada), pp. 1–4, April 1999.


