Local self-heating in short gate GaAs PHEMTs

F. Aniel[1,2], S. Lepaul[3], L. Peymayeche[1], A. De Lustrac[1], F. Bouillault[3] and R. Adde[1]

(1) IEF, CNRS URA 22, Bât 220, Université Paris Sud, 91 405 Orsay Cedex, France
(2) France-Telecom-CNET Bagneux, 196 avenue Henri Ravera, BP 107, 92 225 Bagneux, France
(3) LGEP, CNRS URA 0127, Université Paris XI et Paris VI, Plateau du Moulon, 91 192 Gif-Sur-Yvette, France

Abstract : Self heating in ultrashort gate length GaAs Pseudomorphic High Electron Mobility Transistors (PHEMT) is investigated. The experimental results include an original measurement of the average temperature in the gate drain area of the device channel. An original model which accounts for both bidimensional heat transfer and quasi bidimensional energy balance carrier transport is described. Typical simulations in good agreement with the experiments are discussed.

1. Introduction : A channel temperature under the gate of field effect transistors which is well above the chip temperature has detrimental effects on their performances. This paper presents experimental results and an original self-consistent electro-thermal model which gives informations on the spatial lattice temperature distribution in ultrashort gatelength HEMTs and on the time transients due to self-heating. Aponovitch et al. [1] have recently used a bidimensional energy balance model (2D-EBM) to obtain the local lattice temperature in heterojunction devices. However the application of their 2D-EBM model to short gate HEMTs presents severe convergence problems. In order to have both a realistic treatment and a fast numerical efficiency in the simulation of III-V HEMTs, we have developped a simpler approach based on a quasi bidimensional energy balance carrier transport model (Q2D-EBM) coupled self-consistently with a finite element bidimensional heat solver (H2D)[3]. We first describe the main experimental observations attributed to self heating in ultra short GaAs PHEMTs. Then we describe the self-consistent electro-thermal model and finally we discuss typical simulation results.

2. Measurements : The PHEMT vertical structure consists of a 12nm $In_{0.2}Ga_{0.8}As$ channel grown on a 300 nm GaAs buffer, followed by a 3nm $Al_{0.22}Ga_{0.78}As$ spacer, a 4.10^{12} cm^{-2} δ-doping, a 12nm $1.5\ 10^{18}$ cm^{-3} Si doped $Al_{0.22}Ga_{0.78}As$ layer, and a $3.5\ 10^{18}$ cm^{-3} Si doped GaAs layer under the contact metallization.

The measured extrinsic DC output conductance Gd of field effect devices is often smaller than the intrinsic Gd obtained from HF measurements and extraction of the small signal equivalent circuit. This occurs even if the parasitic access resistances (Rs and Rd) have been taken into account in the comparison. The extrinsic Gd has sometimes negative values. Device self-heating and trapping effects may explain such experimental results.

The Fig.1 shows the 300K DC and pulsed I-V characteristics of a 0.1µm PHEMT. The pulsed voltage values are taken only 10ns after the leading edge of the drain current pulse. Although the transient time due to carrier capture processes on deep level centers lies in the same time range as device self-heating, differences between pulsed IV curves and DC ones are currently attributed to self-heating [7]. The Fig.1 shows that at low current levels where self-heating is small, carrier trapping reduces the level of the DC current. Such measurements do not allow to conclude on the relative weight of self heating and of trapping effects on DX centers in ultra small HEMTs.

There is no specific bias condition which could allow to separate the two physical mechanisms which occur at the same time in a device : the bias conditions which enhance self heating may also increase carrier trapping. Moreover there is no way to separate the phenomena using gate length dependence as both self heating and carrier trapping are enhanced by a reduction of the gate length.

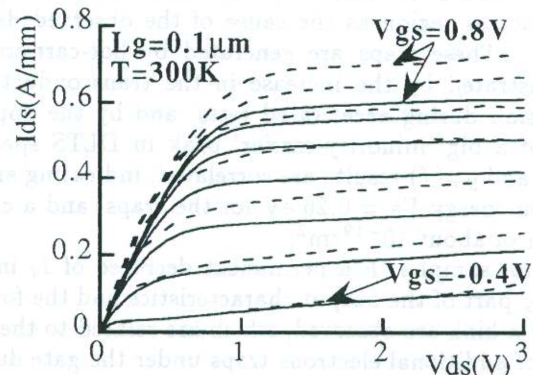

Fig. 1: Static IV (full line) and Pulsed IV (dashed line) characteristic at T_{holder}=300K for a 0.1µm PHEMT.

A current gate analysis and an electroluminescence spectroscopy [2] of a 0.1µm HEMT at high drain current level reveal that the substrate temperature in the gate-drain region can rise by more than 150K above the 300K chip temperature. The local temperature is determined from the energy shift of the GaAs recombination line in the electroluminescence spectrum [2].

The Fig.2 shows the quasi linear increase of this measured average lattice temperature under the channel versus electric power in the device. The electro-thermal simulation of the same device is described in part 4. It must be noted that if impact ionization occurs at relatively moderate source-drain biases in 0,1µm gate length HEMT (Vds >3V), such biases are larger than the conditions of best microwave performances, and are far from the low noise bias range. The apparent linearity of the temperature variation versus electric power dissipated in the device is quite surprising due to the complex non linear mechanisms involved in carrier transport and thermal processes which lead to self heating. This result is interesting because it reinforces the empirical

model describing the device as a thermal resistor. We show here its applicability to ultra short gate HEMTs.

Finally similar experiments at cryogenic temperatures (30K-150K range) did not show such significant self-heating.

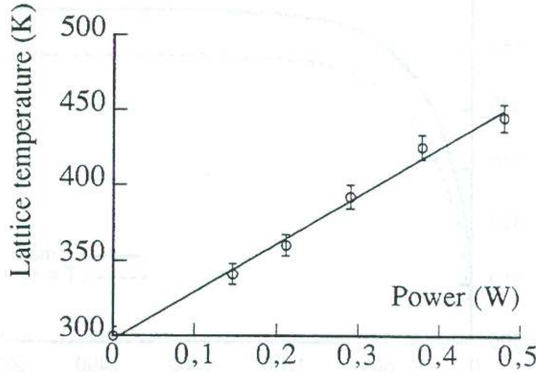

Fig. 2 : Average lattice temperature under the channel versus device electrical power measured by electroluminescence

3. Simulations : In an attempt to separate the contributions due to self-heating and to carrier trapping in the I-V characteristics, and to quantify temperature evolutions in ultra small HEMTs at low bias we have developped a self-consistent electro-thermal simulator. As we have also studied HEMTs at low temperature, it is useful to investigate the evolution of self-heating versus temperature in order to be sure of the temperature of the device in the cryostat [5] during DC and microwave studies. This is also important for DLTS analysis which requires a very precise knowledge of the temperature in the device.

Our model is based on a quasi bidimensional energy balance carrier transport model (Q2D-EBM) coupled with a finite element bidimensional heat solver (H2D)[3].

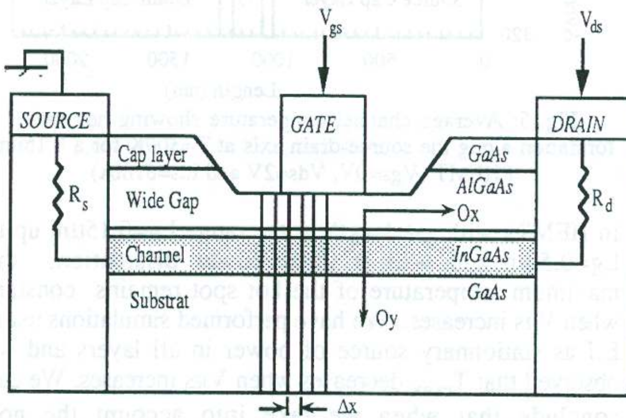

Fig.3: Schematic diagram of a simulated PHEMT.

Transport Solver Q2D :

A Q2D transport model is a well known tool [4] to investigate the behavior of III-V field effect transistors as long as the so called aspect ratio is large. The Q2D model consists of two parts. Poisson's equation is solved under quasi-bidimensional hypothesis.

$$\frac{\partial^2 V(x,y)}{\partial y^2} = -\frac{\rho(y)}{\varepsilon(y)} - \frac{\partial^2 V(x,y)}{\partial x^2}$$

Considering the vertical part of the structure localized between x and x+dx (Fig.3), the electric field gradient in direction Ox (Fig.3) corresponds to a variation of a carrier charge density defined as

$$N_L(y) = \frac{\varepsilon(x,y)}{q} \frac{\partial^2 V(x,y)}{\partial x^2}$$

Poisson's equation becomes

$$\frac{\partial^2 V(x,y)}{\partial y^2} = -\frac{\rho'(y)}{\varepsilon(y)}$$

with $\rho'(y) = (N_D^+ - N_A^- - n(y) + p(y) - N_L(y))$. $N_D^+ - N_A^-$ are the total density of ionized dopants. The electron density n and the hole density p are expressed through Fermi-Dirac statistics. The carrier densities in each layer are tabulated versus both the gate bias Vg and the lateral density $N_L(y)$.

In a second step, the transport equation coupled with the Ox-Poisson's equation is solved. The stationary transport equations are taken as in [3]:

$$\frac{\partial(nv)}{\partial x} = 0 \; ; \; \frac{\partial v}{\partial x} = \frac{qE_x}{m_{eff}(\xi).v} - \frac{1}{\tau_m(\xi)} \; ; \; \frac{\partial \xi}{\partial x} = qE_x - \frac{\xi - \xi_0}{\tau_\xi(\xi).v}$$

m_{eff} : average effective mass

τ_m : momentum relaxation time

τ_ξ: energy relaxation time

E_x : lateral electric field

ξ: average energy of carriers

ξ_0: average energy of carriers at thermal equilibrum

v : average velocity of carriers

The local potential V_c in the channel is calculated by solving the differential equation : $\frac{\partial V_c}{\partial x} = E_x$.

All the parameters used or calculated with Q2D are temperature dependent. The relaxation times are calculated by Monte Carlo simulations and are energy dependent.

Heat solver H2D

The heat equation is : $\rho c_p \frac{\partial T}{\partial t} = \nabla(K \nabla T) + H$

where c_p is heat capacity, ρ is the volumic mass, K the thermal conductivity, and H the heat source. In this model, K, ρ, c_p are temperature dependent.

In the layers where carrier transport may be considered as stationary, H is defined as the electric field times the current density. But in the channel under the short gate, as the electric field is increased, the carriers gain more energy from the field than they lose due to collisions. The distribution function describing the electrons is defined by a carrier temperature T_e, in general higher than the lattice temperature T_L. The part of energy contributing to self-heating corresponds to the energy lost through the collision term $\frac{3}{2} n k_B \frac{(T_e - T_L)}{\tau_\xi}$, where T_L is the equilibrium lattice temperature and k_B is the Boltzmann constant.

We selected as in [1] the boundary conditions and a thermal resistor with a value $R_{th} = 6.7\ 10^{-4}$ Kcm^2/W to connect the bottom of the device to a 300K thermal sink.

Apanovitch et al [1] have provided an algorithm to extend a drift-diffusion treatment to 2D energy balance. However their convergence algorithm is slow. Actually, our Q2D solver requires 1.6 CPU second for a 2500 node mesh on a SPARC IPX workstation. This advantage in speed requires some reasonable approximations .

Since the heat transient time(a few nanoseconds) is larger than the transport relaxation time (a few picoseconds), the time independent transport assumption of Q2D is legitimate. In the Q2D model, the diffusion current, carrier trapping and impact ionisation are neglected and a single valley conduction band is considered. Although a Fermi-Dirac statistics is used, the degeneracy factor does not appear in transport equations. Full 2D Monte Carlo simulations of ultra short devices [6] show that electron injection increases noticeably. Then the Q2D approximation leads probably to a moderate overestimation of the local lattice temperature especially in the drain-gate access area, where the electric field rises to high values and is no more horizontal.

The transport equations are discretized in Q2D by dividing the source-drain distance in vertical slices of thickness dx (Fig1) where the lateral electric field is assumed to be homogenous. The lattice temperature differs from one slice to another. Q2D calculates the physical quantities (electric field, energy, velocity...) from which the heat source term H is determined . Since the heat equation is discretized using a finite element mesh, a routine transforms the heat source term on a finite difference grid into finite element data for H2D. Then, H2D determines the lattice temperature at each node of the finite element meshing. Then the temperature data are injected in the Q2D model for a new run with the determination of new temperature dependent physical quantities. The process reaches self consistent electric/thermal values after a few iterations.

4.Results :

The numerical simulations of HEMTs are in good agreement with experimental results [2] for various topology, doping, gate length and temperature. The aim of the simulator is to supply a fast thermal cartography of HEMTs with a relatively good accuracy (20%). In order to illustrate the evolution of self-heating versus temperature, we present the thermal behavior of a single recessed 0.15µm PHEMT at 300K (V_{gs}=0V, V_{ds}=2V, I_{ds}=61mA).

The Fig.4 shows the calculated time dependence of the hot spot temperature transient with T_{holder}=300K while the Fig.5 corresponds to the average temperature spatial distribution along Ox in the channel when the temperature is stationary (t>2µs).

For fixed bias conditions, the average lattice temperature under the gate rises as expected when the gate length is reduced, due to the increase of device current even if electrons are far less confined for ultra short devices.

However if we compare the maximum temperature reached in the channnel for a given electrical power , the increase of temperature is sligthly smaller for short gate HEMTs (Lg=0.15um) than for longer ones (Lg=0.5um), due to the increase of non stationnary transport when reducing gate length. As there are slight differences of topology between small and long gate HEMT which may influence the thermal behavior, more investigations are needed here.

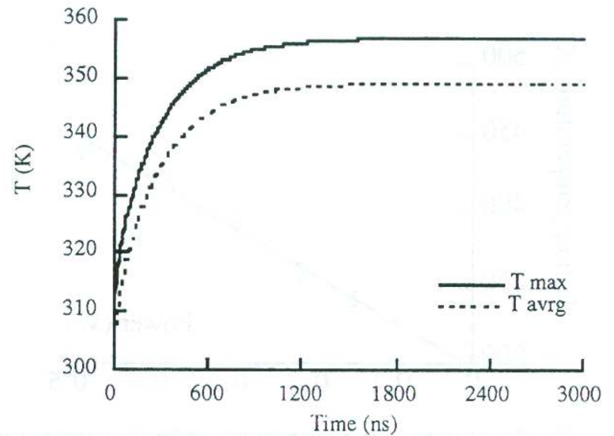

Fig.4: Hot spot temperature (Tmax) and average temperature (Tavrg) in the channel versus time at T=300K for a 0.15µm PHEMT at the bias point V_{gs}=0V, V_{ds}=2V, Ids=61mA.

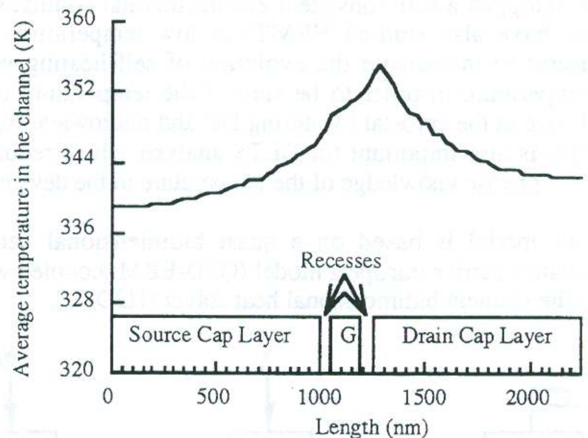

Fig. 5: Average channel temperature showing hot spot formation along the source-drain axis at T=300K for a 0.15um PHEMT (Vgs=0V, Vds=2V and Ids=61mA)

In HEMTs with gatelength in the range Lg=0,15µm up to Lg=0.5um and with a fixed power dissipation, the maximum temperature of the hot spot remains constant when Vgs increases. We have performed simulations using E.J as stationnary source of power in all layers and we observed that T_{max} decreases when Vgs increases. We can conclude that when we take into account the non stationnary nature of transport in small HEMTs, the decrease of the hot carrier maximum energy is approximatively compensated by the larger number of carriers involved in transport.

We have also simulated the thermal behavior of 0.15µm PHEMT under the same bias conditions at 300K and 100K. The 2D electro-thermal simulation at 300K of a AlGaAs/InGaAs/GaAs 0.15µm PHEMT (Vgs=0.4V,

Vds=5V) shows in Fig. 3 a hot spot located on the drain side of the gate which maximum temperature of 402K is comparable to our experimental results which represent an average on the gate-drain region.

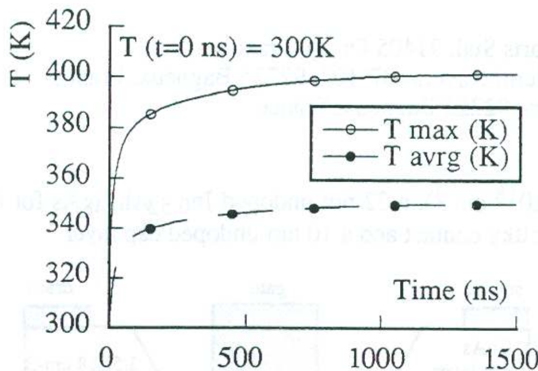

Fig. 6 : Hot spot temperature transient versus time at T=300K : 0.15mm PHEMT at V_{gs}=0.4V and V_{ds}=5V.

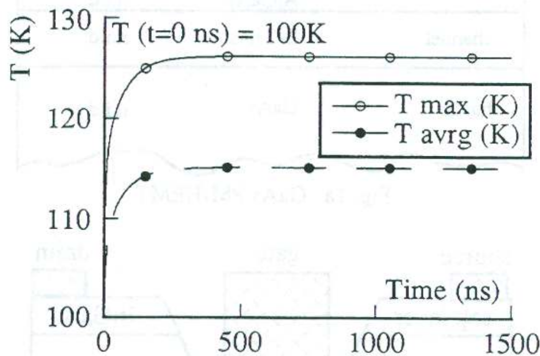

Fig. 7 : Hot spot temperature transient versus time at T=100K : 0.15mm PHEMT at V_{gs}=0.4V and V_{ds}=5V.

The Fig.6 shows the calculated time dependence of the hot spot temperature transient (Tmax) and the average temperature under the gate in the quantum well (Tavrg) for T_{holder}=300K. Fig.7 corresponds to T_{holder}=100K. The absolute increase of temperature is much smaller at T=100K than at T=300K and the transient time is shorter at low temperature while the device dissipated power is larger (in the same bias conditions I_{ds} increases at low temperature thanks to the reduction of the interaction rate of electrons with the lattice). This limited heating (10-25K) of the PHEMT at low temperature and its shorter transient time is due to the increase of the GaAs thermal conductivity from 300K down to 20K, even if the specific heat decreases at low temperature.

This phenomenon is in good agreement with the significant reduction of HEMT noise level measured at low temperature. The relatively small increase of lattice temperature given by the solver explains why we did not detect experimentally by electroluminescence an appreciable rise of temperature. These results at low temperature are a good point for the reliability of cryogenic devices. The recess width, cap layer depth and passivation layer are crucial parameters both for transport and heating of ultra short gate length HEMT.

The main surprising result of the simulations is the nearly linear increase of maximum temperature versus power dissipated in the device for ultra short gate length HEMT. All equations described above are non linear, but the dependency of the hot spot temperature versus power dissipated in the device is linear. Besides the simulated curve is in very good agreement with the experimental results given by electroluminescence. We can conclude that the phenomenological approach using a thermal resistance to model the lattice temperature evolution in short gate HEMTs is legitimated by our simulations in non stationnary transport conditions.

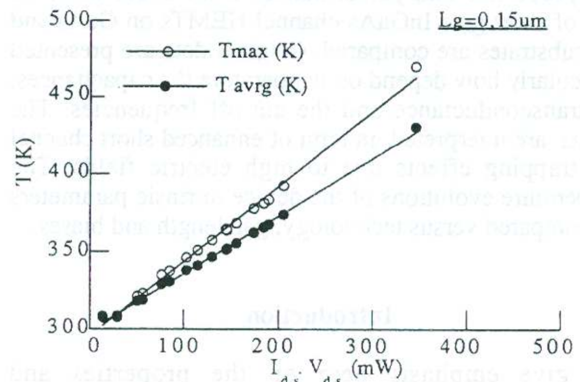

Fig.8 : Maximum temperature and average temperature in the channel of a 0.15um gate length PHEMT versus device electrical power (Tholder=300K).

Conclusion : Self-consistent electrothermal modeling at 300K of short gate HEMTs shows a significant hot spot in good agreement with experimental data. The temperature rise is much smaller at cryogenic temperature so that the gain in performance upon cooling is not lost. The simulation gives the transient time due to thermal effects. At room temperature, a better topology of ultra short gate length PHEMTs should be developped in order to reduce self-heating.

***Acknowlegment** : The authors are grateful to P. Dollfus (IEF) for supplying transport parameters from Monte Carlo simulations and Y. Jin (L2M/CNRS, Bagneux) for realizing the devices.*

References

[1] Y. Apanovitch, P. Blakey, R. Cottle, E. Lyumkis, B. Polsky, A. Shur, A. Tcherniaev, IEEE. Trans. on Electron Devices, Vol.42, N°5,1995

[2] F. Aniel, P. Boucaud, A. Sylvestre, P. Crozat, F. Julien, R. Adde, Y. Jin, J. Appl. Phys, 77 (5), 1995.

[3] S. Lepaul, F. Aniel, L. Peymayeche, A. De Lustrac, F. Bouillault and R. Adde Electron. Letters, in Press

[4] B. Carnez, A. Cappy, A. Kaszynski, E. Constant, G. Salmer, J. Appl. Phys., Vol.51, Jan.1980

[5] F. Aniel, A. Sylvestre, Y. Jin, P. Crozat, A. De Lustrac, G. Vernet, R. Adde same conference 4A7

[6] I. Kizilyalli, M. Artaki, N. Shah, and A. Chandra, IEEE Trans. on Electron Devices, Vol 40, N°2, 1993

[7] J.S. Atherton, C. M. Snowden,J.R. Richardson, IEEE MTT-S Digest, 1993