SELF-HEATING PROCESS IN MICROWAVE TRANSISTORS

Anthony E. Parker(1) and James G. Rathmell(2)

(1) Department of Electronics, Macquarie University, Sydney AUSTRALIA 2109, mailto: tonyp@ieee.org

(2) School of Electrical and Information Engineering, The University of Sydney, AUSTRALIA 2006, mailto: jimr@ee.usyd.edu.au

ABSTRACT

The temperature of microwave FETs is found to vary significantly at extremely high frequencies even though the fundamental thermal time constant is only a few kHz. This affects third-order intermodulation through a process involving temperature rise due to power dissipation at the fundamental and second harmonic frequencies of signal, which is mixed with the fundamental and second-order products of the device's inherent nonlinearity. The process is strongly influenced by the spacing between the frequencies of signal components, which is exploited in a proposed method for characterizing the frequency dependence of self-heating. The impact on circuit performance is that distortion and intermodulation, which vary with bias and load conditions, additionally vary with self-heating. This additional dependence is overlooked in present circuit models and analysis techniques.

INTRODUCTION

The variation of microwave FET characteristics with operating condition (bias, temperature and frequency) is significant because it is invoked by the complex signals used in communication systems. Self-heating [1] and charge-trapping related to impact ionization and leakage currents are known to contribute to the variation of FET characteristics [2]. Each contribution has an effect that varies with operating conditions and frequency.

The consideration of heating and trapping is usually in terms of the so-called dispersion of device characteristics. This recognizes a difference between dc characteristics, used to determine bias, and the high frequency characteristics, used to simulate circuit operation. Device models usually switch from dc to high frequency over a single time-constant [3]. This is accepted on the assumption that low frequency dispersion effects are insignificant at operating frequencies.

The effect of processes having frequencies far below the operating frequencies, such as baseband impedances, are beginning to be recognized [4], [5]. These processes affect broadband and multi-tone systems because they are stimulated by the relatively slow-changing envelope of the signals. An example is the affect on intermodulation of the impedance of the bias network at the frequency equal to the tone spacing. This affects the level and symmetry of intermodulation products of a two-tone signal [5].

It has been suggested that thermal transient behaviour of active devices is also a low-frequency process that influences high-frequency intermodulation [6]. This paper builds upon this suggestion to characterize the self-heating within a transistor and demonstrate its impact on intermodulation. The heating process in a transistor is examined in terms of the frequency response of temperature rise to power dissipation and the mechanism through which it interacts with the inherent nonlinearity of the transistor. A technique for characterizing the heating process is proposed and the influence on intermodulation versus tone spacing and bias is investigated.

SELF-HEATING PROCESS

Self heating in transistors is caused by the power dissipated in them. In many microwave FET models, this is accounted for by a reduction in current that is proportional to power dissipation [3]. Model implementation is in various forms in

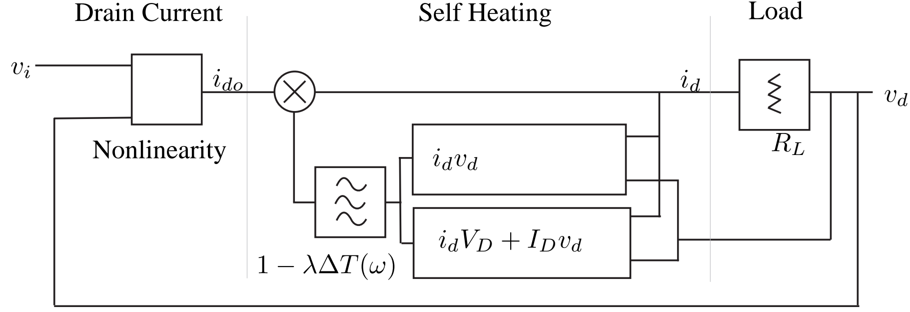

Fig. 1. Schematic block diagram of the self heating process.

various models but all essentially describe drain current, I_d, by:

$$I_d = I_{do} \times (1 - \lambda \Delta T) \tag{1}$$

where I_{do} is the isothermal current (A), λ is the thermal coefficient (1/K), and ΔT is the temperature rise of the device (K) [1]. This model is reasonable for a device operating with velocity saturated drain current, which is the common situation in microwave applications. If the drain current saturation is due to low-field mobility, then an inverse temperature dependency may be more appropriate.

In models that attempt to predict thermal dispersion, the temperature rise is proportional to the average power dissipation calculated over a single time-constant. This gives temperature rise a first-order frequency response to instantaneous power dissipation. The result is dc characteristics with current reduction due to self heating and very little thermal response to high frequencies. The predicted transient response is also first order, which follows a exponential behavior over the time constant used for calculating the average power.

However, the measured transient, or pulsed, behaviour of FETs to step changes in power does not follow a single time constant response. In order to gain some insight into the real transient response of temperature to power dissipation, the frequency response of the thermal path can be examined.

It is proposed that the temperature rise be considered as a yet-to-be-determined function of frequency, ω, so that (1) becomes:

$$I_d = I_{do} \times (1 - \lambda \Delta T(\omega)). \tag{2}$$

This is then implemented in a small-signal model shown in the system diagram in Fig. 1. In this system, the output v_d, or drain potential of a FET, is the voltage across a load. The load is a drain current to voltage converter, in the context of this system. The drain current is a nonlinear function of the input v_i and the output signal v_d, which is multiplied by a response to self heating. The nonlinear function is the inherent nonlinearity of the FET transconductance and drain conductances.

The self-heating process involves a transfer function, to implement the frequency dependence in (2), and the power dissipated by the transistor, which for the common-source configuration (Fig. 3) is:

$$P = (V_D + v_d)(I_D + i_d) = (V_D - i_d R_L)(I_D + i_d) = -{i_d}^2 R_L + i_d(V_D - R_L I_D) + V_D I_D \tag{3}$$

where V_D and I_D are the bias potential and current of the transistor, $i_d = I_d - I_D$ is the signal current, and R_L is the load presented to the drain. The two preprocessing blocks in the Self-heating section of Fig. 1 implement the first and second terms of (3).

The interaction of the self-heating process section of Fig. 1 with the inherent nonlinearity section significantly affects the overall nonlinearity of the device. Before exploring this, the frequency response of the temperature rise is examined.

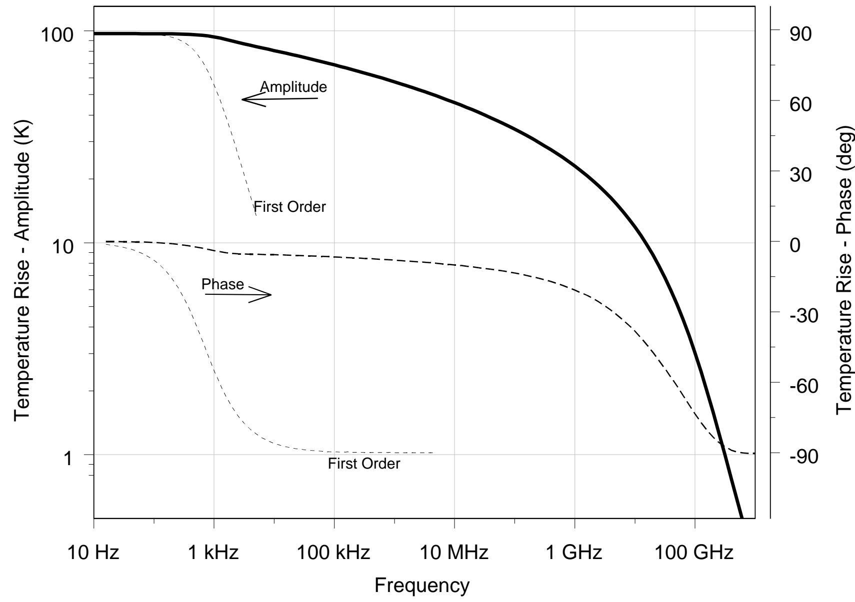

Fig. 2. Simulated frequency response of temperature rise to power dissipation. A first-order response is also shown for comparison.

FREQUENCY RESPONSE OF SELF HEATING

To understand the nature of the self-heating transfer function, a physical model of the heating process was used. The temperature at a point in a device can be determined by the thermal-rate equation, which is:

$$\nabla^2 T + \frac{p}{\kappa} = \frac{1}{\alpha}\frac{dT}{dt} \tag{4}$$

where κ is the thermal conductivity [46 W/mK for GaAs], p is power input per unit volume [W/m^3], and α is the thermal diffusivity [4.4×10^{-5} m^2/s for GaAs] [7]. Nonlinearity of thermal conductivity and diffusivity are not considered here.

For this investigation, the thermal-rate equation was solved by a finite-difference technique, which divides the geometry of the device into three-dimensional cells with volume $\Delta x\,\Delta y\,\Delta z$. The circuit simulator SPICE was used to calculate the solution, with an electrical analogy of: resistance for thermal resistivity of each cell ($r_x = \Delta x/(\kappa \Delta y \Delta z)$, etc), capacitance for heat capacity of each cell ($\Delta x\,\Delta y\,\Delta z \kappa/\alpha)$), voltage for temperature, and current into each cell for power into that cell ($\Delta x\,\Delta y\,\Delta z p$); as in [8]. The use of SPICE affords rapid calculation of the frequency response of a large array of cells with the AC analysis.

Fig. 2 shows the frequency response of a structure representing a FET in the middle of a 1mm square GaAs wafer that is 100μm thick. Power of 50 mW is assumed to be delivered to a 50μm by 200Å region that is 250Å deep, which represents a region under the drain end of a FET. The system was divided into 40,000 cells distributed in a logarithmic fashion, with 300 cells covering the under-drain region. The underside of the wafer is connected to an infinite heat sink at ambient temperature (a voltage source in the electrical analogy).

The simulation does not include the effects of metal contacts, heat flow to air, or the precise details of device geometry. It does, however, provide insight into the nature of the heating of a confined source of power sitting on a large wafer structure. Several similar structures were examined and a few points noted:

- There is a low-frequency roll-off at around 1 to 10 kHz. This frequency is set by the macro dimensions of the thermal path — that is, by the heat capacity and thermal resistivity of the bulk device.
- There is a faster roll-off at very high frequencies. The frequency of this is set by the dimensions of the under-drain power dissipation region, and is likely to be more than 1 to 10 GHz for typical devices.

- The thermal response is significant over many decades of frequency between the low-frequency and final roll-off points. The response in this middle region is much less than 20dB/decade expected of a first-order system. This means that temperature response at 1 GHz maybe only 30 dB below that at 1 kHz, rather than $>$100 dB predicted by a first-order thermal model.
- Changing the structure to a one-dimensional thermal path from the FET to the underside of the wafer increases the middle region roll-off to 10dB/decade. It is expected that a real device will have a roll-off at a lesser slope.
- The phase response is –4.5° for each db/decade of roll-off slope. That is –90° in the 20dB/decade region, or less in proportion in regions of slower roll-off.

The salient features of the thermal response can be captured in the following expression for thermal impedance, which is a transfer function with an order n that is less than unity:

$$R_{th}(\omega) \quad = \quad \frac{R_{tho}}{(1+j\omega/\omega_o)^n(1+j\omega/\omega_c)^{(1-n)}} \tag{5}$$

where, R_{tho} is the steady-state thermal impedance, ω_c is the upper roll-off frequency, ω_o is the lower roll-off frequency, and n is the order of the response in the intervening frequencies. The term n gives a roll-off of $20 \times n$ dB/decade and a corresponding phase shift of $-90 \times n$ degrees.

In the frequency domain, temperature rise is the product of the thermal impedance and the low-signal power dissipation. The latter is transformed to the frequency domain to give:

$$\Delta T(\omega) = R_{th}(\omega) \, \times \, \mathcal{F}(P). \tag{6}$$

INTERMODULATION

The overall nonlinearity of a transistor that is affected by the self-heating process can be assessed by examining its distortion characteristics. To do this, a simple model of the inherent nonlinearity, which produces i_{do}, is incorporated in the system shown in Fig. 1. A suitable description is the following two-dimensional Taylor Series that is truncated to third order and that ignores mixing terms:

$$I_{do} = I_{DO} + g_m v_g + \frac{1}{2} g'_m {v_g}^2 + \frac{1}{3} g''_m {v_g}^3 + G_d v_d + \frac{1}{2} G'_d {v_d}^2 + \frac{1}{3} G''_d {v_d}^3 \tag{7}$$

where g_m, g'_m, and g''_m are the transconductance and its derivatives with respect to v_g, and G_d, G'_d, and G''_d are the drain conductance and its derivatives with respect to v_d, and $I_{DO} = I_D/(1 - \lambda R_{tho} I_D V_D)$ is the isothermal bias current.

The system in Fig. 1 was implemented in a harmonic balance circuit simulator with equations (2), (3), (5), (6) and (7). Various scenarios with a two-tone input and constant load impedance were investigated to determine the dependence of harmonic and intermodulation distortion products on tone frequency and tone frequency spacing. The third-order intermodulation products provided the simplest behavioural pattern, so was further investigated.

An algebraic expression for the third-order intermodulation product can be determined by Volterra analysis of the system [9]. Assuming that the input is two tones with amplitude V_i and frequencies ω_a and ω_b and that the load R_L is independent of frequency, then the low-signal drain current at frequency $2\omega_a - \omega_b$ is:

$$\begin{aligned} i_d(2\omega_a - \omega_b) \quad \approx \quad V_i^3 \Bigg\{ & \frac{1}{4} \left[\frac{g''_m}{1+R_L G_d} - \frac{g_m^3 G''_d R_L^3}{(1+R_L G_d)^3} \right] \\ & - \lambda (R_{th}(\omega_a) + R_{th}(\omega_b)) \frac{g_m}{1+R_L G_d} (V_D - I_D R_L) \times \frac{1}{8} \left[\frac{g'_m}{1+R_L G_d} + \frac{g_m^2 G'_d R_L^2}{(1+R_L G_d)^2} \right] \\ & + \frac{\lambda}{2} \left(R_{th}(\omega_a - \omega_b) + R_{th}(2\omega_a) \right) \frac{g_m^2 R_L}{(1+R_L G_d)^2} \times \frac{1}{2} \frac{g_m}{1+R_L G_d} \Bigg\}. \end{aligned} \tag{8}$$

The derivation of (8) starts with the assumption that the output at the fundamental frequencies is a function of linear terms only and is not affected by the thermal process. This allows elimination of v_d to produce an explicit expression for i_d in terms of v_g, which is evaluated in the frequency domain using the Volterra technique. Thus (8) ignores high-order contributions and subtle feedback effects. However, it does identify the following three primary contributions to third-order intermodulation:

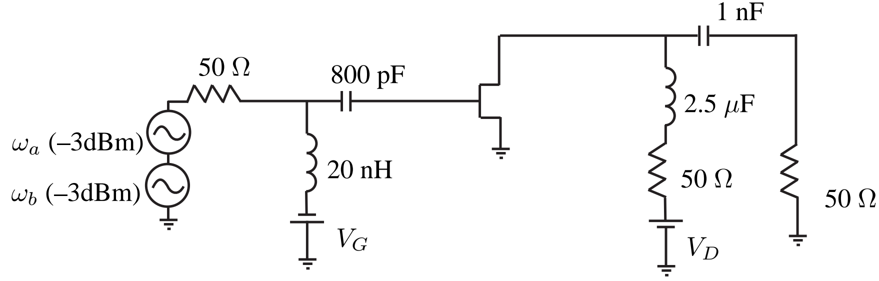

Fig. 3. Set up for measuring third-order intermodulation. The right-most 50Ω resistor is the spectrum analyser. The two-tone source is generated by two signal sources and a resistive combiner.

- The first term in (8) is generated by the inherent third-order nonlinearity of (7). This is bias dependent and, in all FETs, is zero at a point near pinch-off.
- The second term in (8) is generated by the mixing action of the multiplier block in Fig. 1, which is the $\times$ operator in (8). The signals being mixed are the linear terms of the power dissipation (3) that feed back via the thermal resistance, and the products of the second-order nonlinearity of (7). The linear term in (3) is dominated by the fundamental frequencies, so the thermal resistance is a function of these frequencies.
- The third term in (8) is also generated by the mixing action of the multiplier block in Fig. 1. The signals being mixed are the second-order terms of the power dissipation (3) fed via the thermal resistance, and the first-order product of (7). The second-order terms in (3) are dominated by a second harmonic and the difference frequency, so the thermal resistance is a function of these. (Note that the second harmonic at $2\omega_b$ appears in the $2\omega_b - \omega_a$ product.)

It is possible to control the three contributions to third-order intermodulation by choice of tone frequencies and bias. A bias exists in FETs such that the inherent nonlinearity is zero, which eliminates the first contributor. Then the contribution involving $R_{th}(\omega_a - \omega_b)$ can be made dominant by using high frequency tones. This leads to an extraction procedure for $R_{th}(\omega)$ from a plot of $i_d(2\omega_a - \omega_b)$ versus $(\omega_a - \omega_b)$. It is also possible to eliminate the second contributor by setting $V_D = I_D R_L$, though this may be more difficult in practice.

MEASUREMENT

The two-tone measurement set up in Fig. 3 was used to investigate the extraction of $R_{th}(\omega)$. The drain bias is delivered via an RLC network that presents 50 Ω to the drain at all frequencies. This is the case when $\sqrt{L/C} = 50$. The bias network limits the view of the spectrum analyser to frequencies above 10 MHz and imposes a 50Ω load line on the dc characteristics.

A constant impedance load must be presented to the drain of the transistor, so that baseband impedance variations do not contribute to the intermodulation [5]. In particular, any variation in R_L at the difference frequency $(\omega_a - \omega_b)$ will be visible in a plot of $i_d(2\omega_a - \omega_b)$ versus difference frequency. With the bias network in the measurement set up, this baseband impedance contribution will be independent of difference frequency. However, the thermal process should still contribute to variation in intermodulation.

A high-power MESFET (with 800mA saturated drain current and –3.3V pinch-off potential) was investigated with the set up in Fig. 3. The third-order intermodulation is shown in Fig. 4. These results are for a region near pinch-off (5% I_{DSS}) where the inherent third-order nonlinearity is expected to be zero. The higher tone was at 200 MHz which ensured that the third-order products were visible to the spectrum analyser. This frequency is also significantly lower than the f_t of the device.

All contributors to (8) are present in the results of Fig. 4, to a greater or lesser extent depending on the gate bias. However, for a fixed bias and fixed ω_a, only the terms involving $R_{th}(\omega_b)$ and $R_{th}(\omega_a - \omega_b)$ contribute to variation with difference frequency (c.f. (8)). The upper curves in Fig. 4 exhibit the expected behaviour of $R_{th}(\omega_a - \omega_b)$ added to the other constant contributors. When the constant contributors are at a minimum, the thermal transfer function is fully visible. Fitting (5) to this gives an estimated thermal response of –7.6dB/decade with a knee frequency of 1.6 kHz.

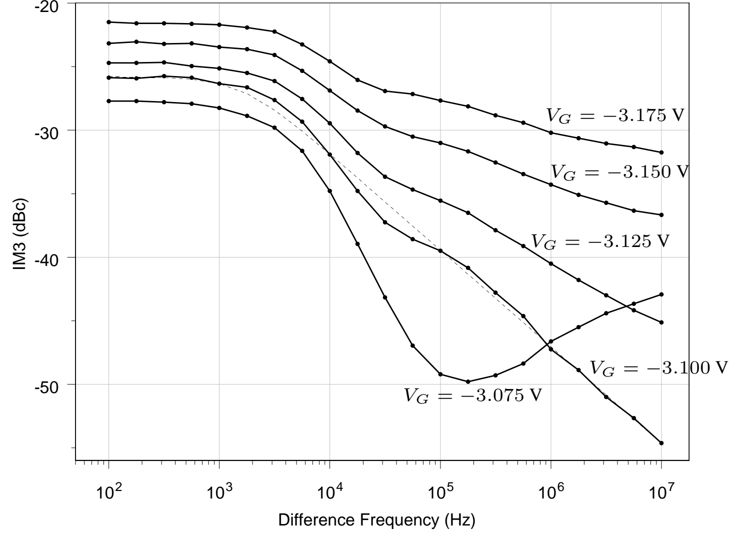

Fig. 4. Measured third-order intermodulation versus difference frequency using the set-up in Fig. 3. The device is a power MESFET with 800mA I_{DSO} and –3.3V pinch off potential. The higher tone was at 200 MHz and $V_D = 3$ V. The dashed line represents a –7.6dB/decade estimated thermal response with a knee frequency of 1.6 kHz.

The lower curve in Fig. 4 exhibits a contribution from $R_{th}(\omega_b)$. Note that ω_b is reducing as the difference frequency is increasing, so $R_{th}(\omega_b)$ also increases. This is more significant in the results for higher gate bias (not shown here).

SIMULATION

To support the interpretation of the measured data above, a simple large-signal model was used to simulate similar results. The system in Fig. 1 was simulated with the following expression for I_{do} and equations (2), (3), (5), and (6):

$$I_{do} = \beta \left(\sigma \ln \left[e^{(V_G + v_g + \gamma(V_D + v_d) - V_{TO})/\sigma} + 1 \right] \right)^2 \tag{9}$$

where V_{TO} is the pinch-off potential, σ is potential difference for transition to sub-threshold mode, and γ is the drain-feedback factor [3]. This expression was used to provide a convenient and realistic nonlinearity that varies with bias.

The results of a representative simulation are shown in Fig. 5. The parameters used were $\beta = 0.10$ A/V^2, $\gamma = 0.01$, $\lambda R_{tho} = 0.2$ W^{-1}, $n = 0.38$, $\omega_o = 2\pi \times 1600$ rad/s, $\sigma = 0.2$ V, $V_D = 3$ V, $V_{TO} = -3.575$ V. The simulation shows that the gate bias with minimum intermodulation at high difference frequency follows the R_{th} function very closely.

Thus a procedure for extraction R_{th} is straight forward. A constant impedance load is required for the drain. A suitable gate bias is found by searching for minimum intermodulation with a large tone spacing. Then a measurement of intermodulation versus tone spacing gives the frequency dependence of $R_{th}(\omega)$, which is offset on a decibel scale. The value of λR_{tho} can be extracted from dc and pulsed data once the frequency dependence (and hence transient response) is known.

DISCUSSION

The impact of the thermal process at other biases was explored by simulation. Fig. 6 shows the third-order intermodulation for the same parameters used previously except for β, which was reduced to 0.01 A/V^2, and V_D, which was increased to 8 V. These reduce change in V_D on the 50Ω load line to avoid the saturation knee.

The local minimum in intermodulation near pinch-off is visible on the right-hand edge of the surface in Fig. 6, which is the region discussed in the previous section.

At higher gate bias, there is significant dependence on frequency spacing, which is caused by the self-heating process.

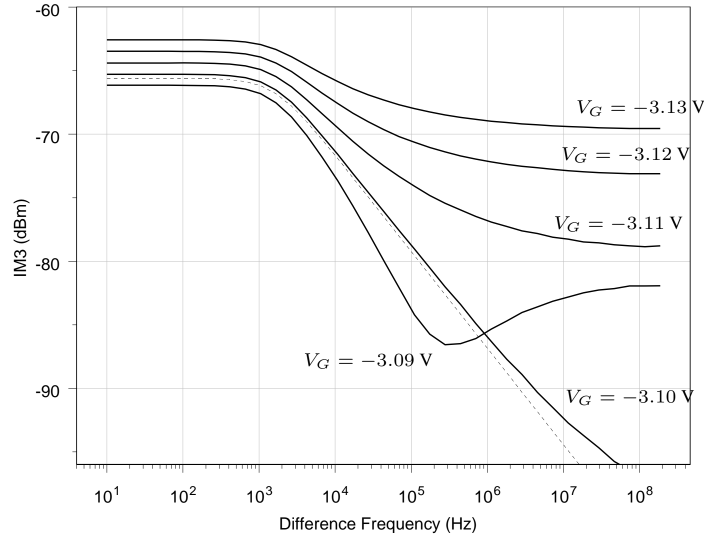

Fig. 5. Simulated third-order intermodulation versus difference frequency. The dashed line represents the –7.6dB/decade thermal response that was included in the model.

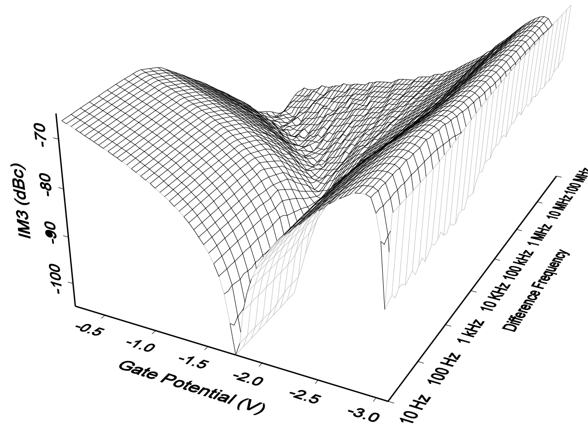

Fig. 6. Simulated third-order intermodulation versus difference frequency and gate bias. Equations (2), (3), (5), (6), and (9) were used with $\beta = 0.01$ A/V^2, $f_a = 500$ MHz, $\gamma = 0.01$, $\lambda R_{tho} = 0.2$ W^{-1}, $n = 0.38$, $\omega_o = 2\pi \times 1600$ rad/s, $\sigma = 0.2$ V, $V_D = 8$ V, $V_{TO} = -3.575$ V.

The large variation in intermodulation versus bias is highly dependent on the frequency spacing of the two-tone input. A significant contributor to this dependence is the term involving $V_D - I_D R_L$ in (8). Fig. 7 shows the marked affect of difference frequency on the intermodulation level versus gate bias and its dependence on drain bias.

CONCLUSION

The results above confirm the existence of a self-heating process in FETs that, to date, has been largely ignored. That is, present models tend to include heating for the evaluation of bias condition only.

The self-heating process is significant at all frequencies. Although dominant at dc, its effect is reduced by only a few tens of dB at microwave frequencies.

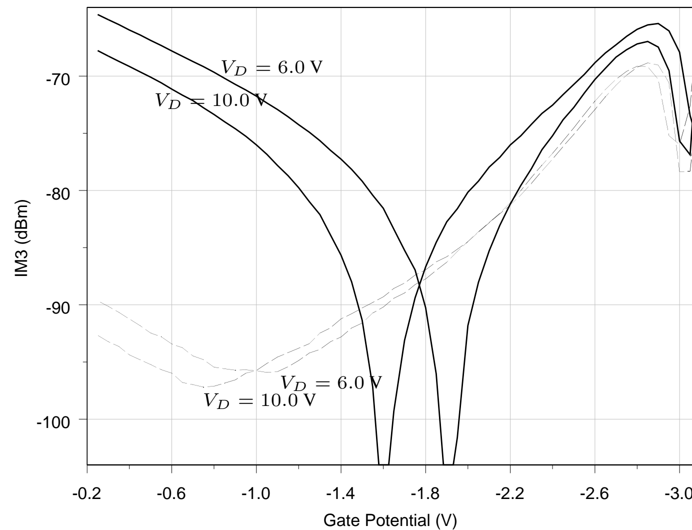

Fig. 7. Simulated third-order intermodulation versus gate bias for a difference frequencies of 10 Hz (—) and 1 MHz (– –). Parameters are the same as those in Fig. 6.

At the operating frequencies of the transistor, the self-heating process is significant enough to alter the bias versus intermodulation characteristics. This implies that the extraction of low-signal Taylor Series coefficients will vary with the frequency spacing used. A significant impact is to be expected on the modelling of broadband systems where the signals are composed of many tones and frequency spacings. The solution to this is to incorporate the self-heating process in simulation models, so that the response to complicated signals can be accurately predicted.

The inherent nonlinearity in broadband models can be de-embedded from measured distortion characteristics once the thermal response is known. Work is proceeding to develop a suitable extraction process for these comprehensive models. Note that it is also necessary to incorporate the effects of electron trapping and impact ionization, which will each exhibit a frequency dependence that must be determined.

References

[1] A. E. Parker, J. Rathmell, and J. Scott, “Pulsed measurements,” in *The Modern Microwave and RF Handbook* (M. Golio, ed.), ch. 4.1, USA: CRC Press, pp. 4–68 – 4–95, 2000.

[2] A. E Parker and J. G. Rathmell, “Bias and Frequency Dependence of FET Characteristics,” *IEEE Journal on Microwave Theory and Techniques*, vol. 51, no. 2, pp. 588–592, Feb. 2003.

[3] A. E. Parker and D. J. Skellern, “A realistic large-signal MESFET model for SPICE,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 45, pp. 1563–1571, Sept. 1997.

[4] N. B. De Carvalho and J. C. Pedro, “A comprehensive explanation of distortion sideband asymmetries,” *IEEE Trans. Microwave Theory Tech.*, vol. 50, pp. 2090-2101, Sept. 2002.

[5] J. Brinkhoff and A. E Parker, “Effect of Baseband Impedance on FET Intermodulation,” *IEEE Journal on Microwave Theory and Techniques*, vol. 51, no. 3, pp. 1045-1051, Mar. 2003.

[6] S. David, W. Batty, A. J. Panks, R. G. Johnson, and C. M. Snowden, “Thermal Transients in Microwave Active Devics and Their Influence on Intermodulation Distortion,” *IEEE International Microwave Symposium*, vol. 1, pp. 431–434, Phoenix, USA, 20-25 May 2001.

[7] S. M. Sze, *Semiconductor Sensors*, USA: John Wiley & Sons Inc., 1994.

[8] K. M. McNally, A. E. Parker, D. L. Heintzelman, B. S. Sorg, J. M. Dawes, T. J. Pfefer, and A. J. Welch, “Dynamic Optical Thermal Modeling of Laser Tissue Soldering with a Scanning Source,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 5, no. 4, pp. 1072–1082, July/Aug. 1999.

[9] S. A Maas, *Nonlinear Microwave Circuits*, USA: IEEE Press, 1997.