New Trends for the Nonlinear Measurement and Modeling of High-Power RF Transistors and Amplifiers With Memory Effects

Patrick Roblin, *Member, IEEE*, David E. Root, *Fellow, IEEE*, Jan Verspecht, *Fellow, IEEE*, Youngseo Ko, *Member, IEEE*, and Jean Pierre Teyssier, *Member, IEEE*

(Invited Paper)

Abstract—Power amplifier (PA) behavior is inextricably linked to the characteristics of the transistors underlying the PA design. All transistors exhibit some degree of memory effects, which must therefore be taken into account in the modeling and design of these PAs. In this paper, we will present new trends for the characterization, device modeling, and behavioral modeling of power transistors and amplifiers with strong memory effects. First the impact of thermal and electrical memory effects upon the performance of a transistor will be revealed by comparing continuous wave and pulsed-RF large-signal measurements. Pulsed-RF load–pull from the proper hot bias condition yields a more realistic representation of the peak power response of transistors excited with modulated signals with high peak-to-average power ratio. Next, an advanced device modeling method based on large-signal data from a modern nonlinear vector network analyzer instrument, coupled with modeling approaches based on advanced artificial neural network technology, will be presented. This approach enables the generation of accurate and robust time-domain nonlinear simulation models of modern transistors that exhibit significant memory effects. Finally an extension of the X-parameter (X-parameter is a trademark of Agilent Technologies Inc.) behavioral model to account for model memory effects of RF and microwave components will be presented. The approach can be used to model hard nonlinear behavior and long-term memory effects and is valid for all possible modulation formats for all possible peak-to-average ratios and for a wide range of modulation bandwidths. Both the device and behavioral models have been validated by measurements and are implemented in a commercial nonlinear circuit simulator.

Index Terms—Behavioral model, device modeling, large-signal RF measurements, memory effects.

I. INTRODUCTION

POWER transistor amplifiers express behavior that has as its origin the memory effects of the transistor technology on which they are built. The term memory is used here to de-

scribe effects that vary on timescales much longer than those associated with the RF period. Some causes of memory effects in transistors are dynamic self-heating (always present due to thermodynamic considerations), dynamic trapping (capture and emission) effects, prevalent especially in some compound semiconductor material systems (e.g., GaN and GaAs), and parasitic bipolar transistor effects in semiconductor-on-insulator (SOI) MOSFETs. In microwave components, the biasing circuit is also known to introduce memory effects that must be considered in behavioral modeling.

Self-heating plays a major role in power transistor operation. Conversion of electrical energy to heat creates a change in device temperature that, in turn, affects key electrical device characteristics, such as power gain. The device temperature distribution depends significantly on the class of operation, itself determined by dc-bias conditions, load conditions, and power levels associated with the RF signals. Thermal boundary conditions are also very important.

Trapping phenomena, in the bulk or at the surface of semiconductor material, are prevalent in III–V field-effect transistor (FET) technologies, such as GaAs-based MESFETs and pseudomorphic HEMTs (pHEMTs), and recently introduced high-speed and power technologies based on GaN. Trapping effects of different types are claimed to be responsible for "power slump," "knee collapse," and the phenomena of drain-lag and gate-lag [1], [2].

Memory effects make it much more difficult to estimate, from conventional transistor characterization data, the actual large-signal transistor performance at RF frequencies and large-amplitude stimuli. It is generally impossible to infer, accurately, the power or efficiency of a GaN transistor at large input power from dc and linear S-parameter measurements alone.

Multiple mechanisms producing long-term memory are often present simultaneously in transistors, making it difficult to separately identify and model the independent contributions for a particular device,

Over the past 20 years, pulsed bias and pulsed S-parameter measurement techniques have been deployed to provide information that device modelers could use to separate distinct memory mechanisms for more comprehensive models. The timescales (e.g., 100 ns) for such measurements are shorter than the timescales for variation of a "slow" variable like temperature or states related to trapping. Such "iso-thermal" or "iso-dynamic" data gives a snapshot of the device currents

and linear behavior at conditions set by the quiescent operating point from which the pulses originate. However, all this information is parameterized by the quiescent point. Converting this into a comprehensive dynamic model is a very difficult task, taking many additional steps, and many additional approximations. This is especially true for the details of how the trap states affect both the current, and the output charge storage of the device. Examples of such approaches, based on different assumptions about how the traps modify the drain current, are given in [1] and [7]. Much less work has concerned how the trapping effects modify the model terminal charges, and therefore the "high-frequency" memory of the intrinsic device. Even though traps will not normally respond instantaneously to an RF signal, the average trap occupation will still depend on the dynamic load-line trajectory for large-signal excitations [8], [9] due to cyclostationary effects. Indeed, as was demonstrated in [10], the trap occupation is dependent on the time average of the bias-dependent emission and capture rates along the dynamic load-line trajectory.

Measurements revealing the impact of thermal and electrical memory effects upon the large-signal RF response of a transistor will be presented in Section II. This will be achieved by comparing continuous wave (CW) and pulsed RF large-signal measurements for various hot-bias conditions and duty rates. Pulsed-RF measurements are useful to obtain a more realistic characterization of the RF response of a transistor at peak power when excited with modulated signals with high peak-to-average power ratio (PAPR) excitations.

Section III will present a systematic and accurate approach to generating a detailed time-domain large-signal simulation model directly from large-signal data from a modern nonlinear vector network analyzer (NVNA) instrument. A complete methodology will be used to convert the data directly into nonlinear constitutive relations of current and charge that include detailed dependence on the temperature and the two trap states considered.

Behavioral modeling of microwave components is an alternate approach to device modeling, which is of great interest to the designers of amplifiers that are used in today's wireless communication infrastructure. An important problem faced by these engineers is the difficulty to characterize, describe, and simulate the nonlinear behavior of amplifiers that are stimulated by signals that have a high peak-to-average ratio and that stimulate the amplifier at a power range covering the full operating range. This is problematic for at least two reasons. Firstly, the amplifier behavior may be driven into full saturation and is as such strongly nonlinear. A good overview of existing techniques for the behavioral modeling of microwave amplifiers can be found in [11]. Most of the existing approaches are based on Volterra theory and as such rely on polynomial approximations. Polynomial approximations cannot easily handle hard nonlinear behavior such as saturation. Secondly, the amplifier behavior shows memory effects as previously discussed. Unfortunately approaches that can handle hard nonlinear behavior, such as a simple compression and AM-to-PM characteristic or a more advanced poly-harmonic distortion (PHD) model [12], [13], have no straightforward way of dealing with memory effects.

Section IV of this paper will present an original behavioral model that is not based on polynomials and that can easily handle memory effects. One of the advantages of the new approach, called dynamic X-parameters, is that the model can be extracted by performing a simple set of measurements. Another advantage is that the model remains valid for a wide range of modulation bandwidths, which is typically not the case for classic approaches.

II. PULSED AND CW LARGE-SIGNAL MEASUREMENTS

As was discussed in Section I, various types of low-frequency memory effects arise in transistors due to different physical processes. Self-heating, trapping, and bipolar parasitic transistor effects in MOSFETs are typically the main sources of low-frequency memory effects with electrical time constants running from 100 s to 100 ns. It is of interest to investigate the impact of these memory effects on the performance of power amplifiers (PAs) excited by signals with high PAPR.

Complex modulated signals such as WiMAX and long-term evolution (LTE) exhibit large PAPR reaching 10 dB. However PAPRs larger or equal to 10 dB arise with a very small probability (typically 10^{-5}) and take place for short duration on the order of 100 ns. Thus, most of the time the amplifier is amplifying the modulated signals at the average power. This raises questions about which characterization and modeling techniques enable a realistic simulation of the device performance for excitations with large PAPR.

Load–pull is a phenomenological technique that is often used either in simulation or experimentally to design PAs. In load–pull measurements, the multiharmonic terminations presented by the load and source networks are tuned to optimize the power-added efficiency (PAE) of the PA while maintaining its gain over the dynamic range of the output power. Conventionally, load–pull relies on CW RF signals. However, CW load–pull may not provide a reliable evaluation of the transistor performance if the device exhibits strong memory effects. To demonstrate this, let us first consider self-heating. For amplifiers operating in class B, AB, or class F, the power dissipated by the transistor increases with increasing input power. However, considering that peak power is typically achieved for a very short duration (100 ns), has a low probability of occurrence, and thermal effects are usually relatively slow ($\gg \mu$s), the device will remain operating at the temperature established by the average dissipated power. Therefore, CW load–pull at the peak power may not yield realistic results due to the comparatively much larger self-heating it will induce. Next consider the impact of traps in devices such as GaN HEMTs. Under low-power CW operation, the drain voltage peak remains typically small and the traps are not activated. However, for modulated RF signals, the instantaneous drain voltage reaches intermittently large values (50–60 V). Traps in GaN typically have very short capture time (ns) at high drain voltages (50 V) and very long emission time (s) at lower drain voltages (< 40 V). Therefore, traps charge during the brief periods of peak power and do not have time to discharge during the intervening time. The dominant response of the transistor for the average input power is then directly affected by the
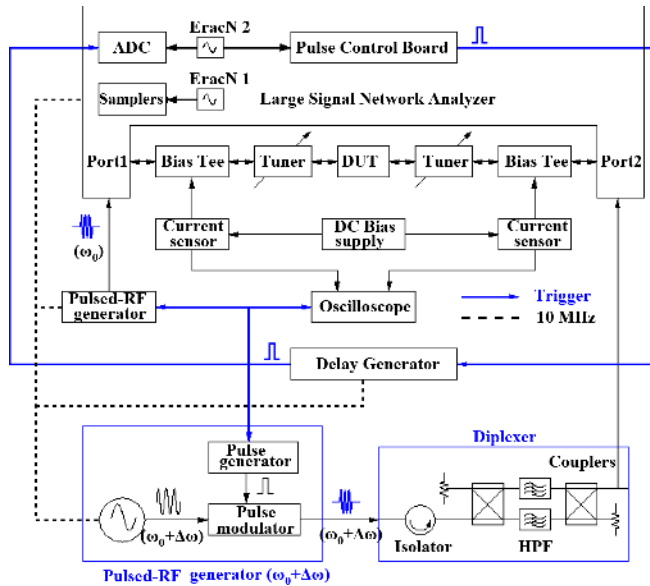
Fig. 1.   Pulsed I–V pulsed RF test bed used to perform both pulsed-RF RTALP and pulsed active load–pull.

Fig. 2.   Constant PAE contour obtained from pulsed-RTALP shown here for 0.4-dBm incident power at 1.9 GHz for class-AB operation.

charged traps and associated "knee walkout" effect resulting from these intermittent high power peaks of the modulated RF signal. Thus, the dynamics of charge trapping for modulated RF excitations cannot be captured by CW load–pull simulations and measurements.

An increase of output $1/f$ noise can also be detected due to the activated traps. Note that the state of the traps is not determined by the dc-bias condition, but by the RF load line (cyclostationary effect) [8]–[10]. Other memory effects such as the slow parasitic bipolar in SOI-MOSFET may also be expected to be affected by the dynamic RF load lines.

The impact of memory effects on the transistor performance excited by signals with high PAPR can be detrimental (trap case) or beneficial [lower temperature of operation or reduced parasitic bipolar junction transistor (BJT)] depending on the particular dominant memory process. Since CW load–pull measurements may not provide a realistic characterization approach, other alternatives must be pursued. To more realistically characterize the performance of a transistor for RF signals with high PAPR modulation, pulsed RF large-signal measurements from a well-defined hot-bias operating condition offers such an alternative. The pulsed I–V/RF testbed used in this section is shown in Fig. 1. The pulsed measurements with the large-signal network analyzer (LSNA) rely on the method of multiple recording data acquisition [14], [9] to achieve 0% desensitization independently of the duty rate. The new technique of pulsed-RF real-time active load–pull (RTALP) [15], [16] measurements was used to optimize the output load.

To investigate the use of pulsed-RF large-signal measurements to characterize memory effects, we shall now consider a sub-cell of a power SOI-MOSFET transistor. This device will be initially tested for three different incident power levels of 1, 4, and 11 dBm at 1.9 GHz, for CW and pulsed RF signals of 1% duty rate. The PAE contour plot obtained from pulsed-RTALP is shown in Fig. 2 for an input power level of 0.4 dBm. The optimal load $\Gamma_L = 0.5\angle37$ selected for low-power operation

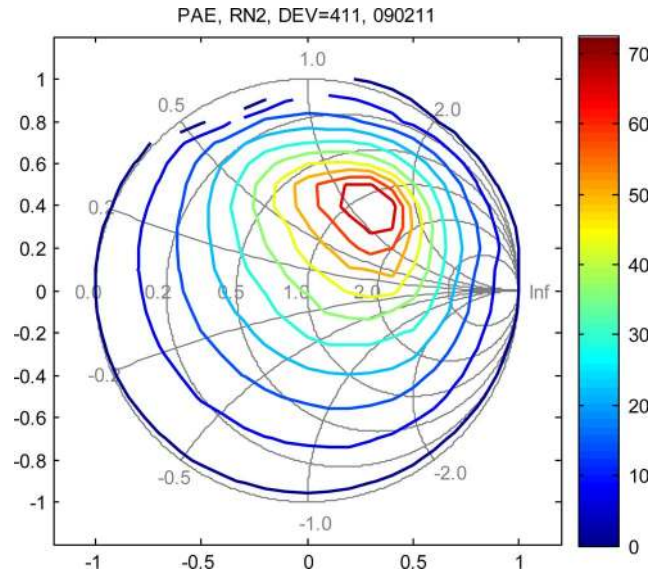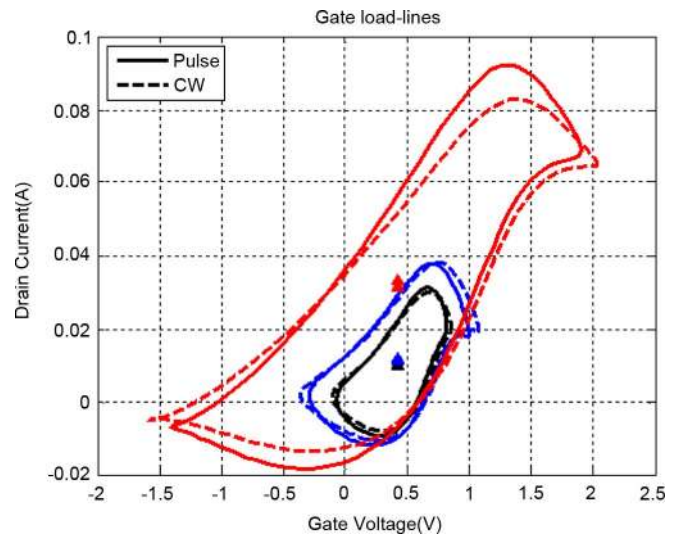Fig. 3.   RF transfer obtained for CW (dashed lines) and pulsed (plain lines) RF excitations with incident power of 1, 4, and 11 dBm and 1% duty rate at 1.9 GHz.

yields a maximum PAE of 66%. Note that a different drain load of $\Gamma_L = 0.22\angle117$ is used at 11-dBm incident power. The transistor is found to operate in quasi-class AB. Load–pull at the second and third harmonic were verified to yield negligible PAE improvements.

The RF transfer characteristic i_D versus v_{GS} and dynamic load lines i_D versus v_{DS} obtained are shown in Figs. 3 and 4 and the performance of the device is summarized in Table I.

The gate voltage swing shown in Fig. 3 is seen to be about the same for the various incident signals. The output drain current is also similar for the 1- and 4-dBm incident power levels. However, for 11-dBm incident power, the drain current swing is much larger. The drain current under pulsed RF reaches a more negative value indicating an increased power generation during the RF cycle associated with the drain charge storage. Fig. 4 also shows a corresponding increased drain voltage swing. It is
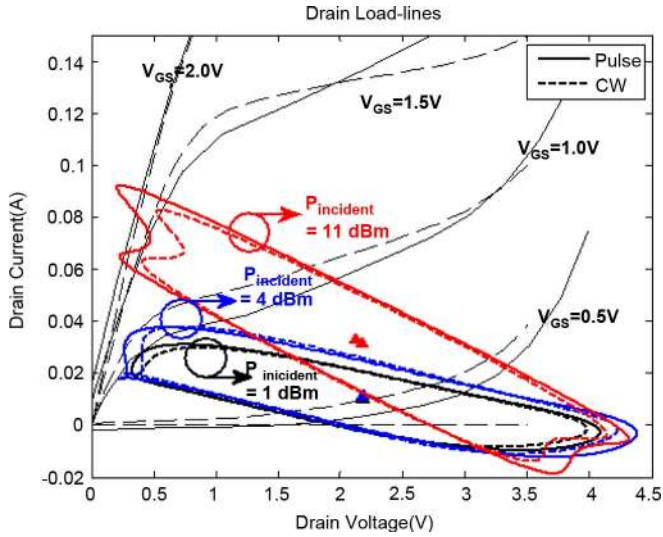
Fig. 4. Drain load lines obtained for CW (dashed lines) and pulsed (plain lines) RF excitations at 1.9 GHz with incident power of 1, 4, and 11 dBm and 1% duty rate.
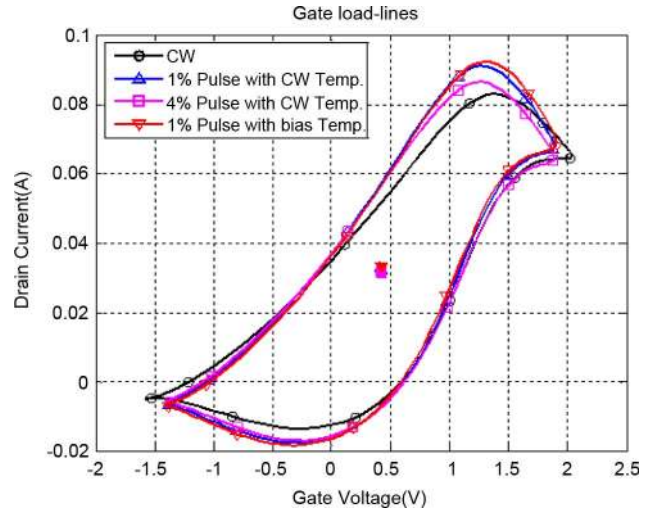

Fig. 5. RF transfer at 1.9 GHz obtained for 11-dBm CW (black line with circles) and pulsed RF excitations for 1% (blue and red lines (in online version) with up and down triangles) and 4% (magenta line (in online version) with square) duty rate and different hot biases.

TABLE I
CW AND PULSED RF TRANSISTOR PERFORMANCE AT 1.9 GHz

Meas.	V_{GS} (V) (OFF/ON)	I_{DS}(mA) (OFF/ON)	P_{in} (mW)	P_{out} (mW)	Gain (dB)	PAE (%)
CW*	0.41	3/9.8	0.3	12.8	16.1	58.0
Pulsed*	0.41/0.41	3/10.0	0.3	14.5	16.7	65.9
CW†	0.41	3/11.7	0.6	16.7	14.6	63.0
Pulsed†	0.41/0.41	3/11.3	0.7	18.0	13.9	70.5
CW#	0.41	3/31.6	2.0	42.8	13.3	59.6
Pulsed#	0.41/0.41	3/33.1	2.5	54.3	13.4	73.7

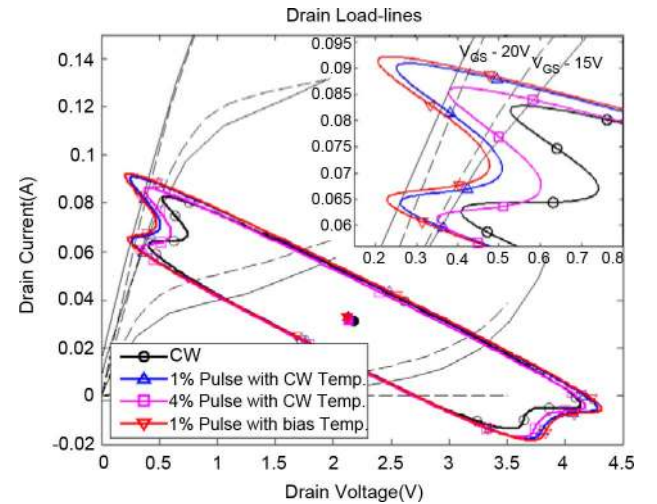*1 dBm, †4 dBm and #11 dBm incident power

Fig. 6. RF dynamic load line at 1.9 GHz obtained for 11-dBm CW (black line with circles) and pulsed RF excitations for 1% (blue and red lines (in online version) with low and upper triangles) and 4% (magenta line (in online version) with square) duty rate and different hot biases.

TABLE II
CW AND PULSED RF TRANSISTOR PERFORMANCE @ 11 dBm

Duty Rate (%)	V_{GS} (V) (OFF/ON)	I_{DS}(mA) (OFF/ON)	P_{in} (mW)	P_{out} (mW)	Gain (dB)	PAE (%)
100	0.41	3/31.6	2.0	42.8	13.3	59.6
1	0.54/0.41	11.8/32.8	2.9	52.5	12.6	71.0
4	0.54/0.41	11.8/31.5	3.1	47.2	11.8	65.6
1	0.41/0.41	3/33.1	2.5	54.3	13.4	73.7

to be noted that the pulsed-RF load line reaches a lower knee voltage than expected from the dc and pulsed I–V at $V_{GS} = 2$ V (gray dashed and plain lines in Fig. 4). As indicated in Table I, the pulsed RF measurements exhibit an improved PAE of 14% compared to the CW measurement while maintaining the same power gain.

To investigate the physical origin of this memory effect, we shall now: 1) vary the dc-bias points of the pulsed-RF measurements or 2) change the duty rate. The results are shown in Figs. 5 and 6 and the transistor performance summarized in Table II. The previous pulsed RF load line (red line (in online version) with top triangles) that was measured with the device at the dc-bias temperature (no RF applied) and the CW load line (black line with diamonds) are included for references. The pulsed RF load lines with a higher hot-biasing current of 11.8 mA (blue line (in online version) with down triangles) that largely overestimates the self-heating in CW operation is seen to remain similar to the pulsed RF load line with the bias-point temperature (red line (in online version) with up triangles). A resulting decrease in PAE of only 2% with this overestimated operating temperature is obtained as shown in Table II. Clearly, self-heating is not the dominant memory effect in this device.

Now increasing the duty rate from 1% to 4% (magenta line (in online version) with squares) is seen to produce a dynamic transfer characteristic and a dynamic load line that is bridging the gap between the ones obtained from CW and pulsed RF in

Figs. 3 and 4. The increased duty rate also yields a reduced PAE increase of 8% instead of 14% in Table II.

In the particular case of the SOI technology considered, the memory effects have a beneficial impact on the device performance. Namely, under the transient peaks of the modulated RF excitations, the transistor performed better than expected from operation at CW peak power while maintaining the same operating temperature. A similar trend was observed on different devices and on different wafers. The precise origin of this memory effect is not known, but several assertions can be made. Self-heating associated with the power dissipation in the device was verified not to be the controlling factor. This is explained by the relatively low thermal resistance of the substrate used. The different gate voltage swing, drain voltage swing, and drain current swing observed in Figs. 3 and 4 indicate that the gate and drain charges under the channel are different under pulsed and CW operations. The lower effective knee voltage observed in pulsed operation compared to CW operation could arise from a reduced effective channel resistance. The more negative drain current observed in Figs. 3 and 4 under pulsed operation that is associated with the improved PAE (negative power dissipation) is also indicative of a more efficient recovery of the charge stored in the channel. Presumably under unrealistic CW operation at peak power, the steady-state impact ionization may lead to charge accumulation adversely modifying the channel characteristics of the transistors.

Note that memory effects do not always have a beneficial impact like in the SOI MOSFET considered here. For example, trapping in GaN HEMTs induced by the transient peaks of the envelope of complex modulated RF excitations usually degrades the device performance, as was discussed. In all the cases, the impact of memory effects on the transistor performance for RF signals with high PAPR needs to be carefully assessed. This is particularly needed for a more realistic design of PAs relying on some type of load modulation for improved power efficiency. In such a case, pulsed-RF load–pull from the proper hot bias condition will yield a more realistic representation of the peak power response of transistors excited with modulated signals with high PAPR. The modeling of memory effects will now be investigated in Sections III and IV.

III. TIME-DOMAIN DEVICE MODELING

A. State Equations for Dynamical Variables

In this section, we are concerned with the development of a detailed time-domain large-signal simulation model directly from large-signal data from a modern NVNA instrument. We assume we can model the transistor by a set of coupled nonlinear ordinary differential equations of circuit theory. This means there is postulated a set of coupled nonlinear equivalent circuits for the electrical, thermal, and other dynamical variables (e.g., trap states) that determine the behavior of the device model.

An intrinsic model of a III–V FET with thermal and trapping dependent memory is presented in Fig. 7. A simple one-pole thermal equivalent circuit is used to compute the junction
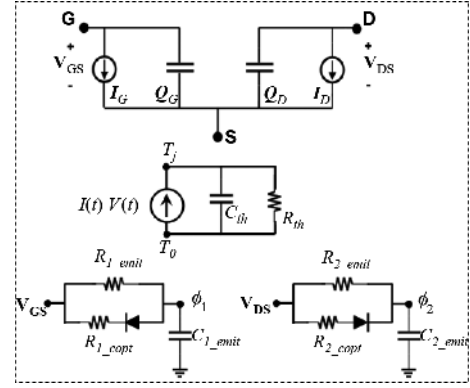

Fig. 7. Nonlinear equivalent circuit of III-V FET model with dynamic self-heating and trapping.

temperature from knowledge of the dissipated electrical power. More poles can be added if necessary to better model the distributed nature of heat propagation. Two species of trapping and emission phenomena are described by the remaining equivalent circuits in Fig. 7. These are of the form proposed in [1] to describe gate-lag and drain-lag phenomena, respectively, common in III–V GaAs and GaN FETs. The key dynamical principle here is that there are asymmetric fast capture processes (in the directions of the diode) and slow emission processes that depend on both the direction and the rate of the applied fields (voltages). The intrinsic electrical model for currents and charges depends nonlinearly on five variables—the two instantaneous intrinsic terminal voltages, the junction temperature, and the two trap state voltages across the respective gate-lag and drain-lag capacitors, respectively.

The equations for the instantaneous drain current and the other time-dependent dynamical variables associated with junction temperature, $T_j(t)$, and traps states for gate-lag and drain-lag, respectively, $\phi_1(t)$ and $\phi_2(t)$, are given in (1)–(4). Gate current equations can be found in [3]

$$I_D(t) = I_D(V_{GS}(t), V_{DS}(t), T_j(t), \phi_1(t), \phi_2(t))$$
$$+ \frac{d}{dt} Q_D(V_{GS}(t), V_{DS}(t), T_j(t), \phi_1(t), \phi_2(t)) \quad (1)$$

$$\frac{dT_j}{dt} = \frac{T_0 - T_j(t)}{\tau_{th}} + \frac{1}{C_{th}} \langle I(t)V(t) \rangle \quad (2)$$

$$\frac{d\phi_1}{dt} = f_1(V_{GS}(t) - \phi_1(t)) + \frac{V_{GS}(t) - \phi_1(t)}{\tau_{1\_emit}} \quad (3)$$

$$\frac{d\phi_2}{dt} = f_2(V_{DS}(t) - \phi_2(t)) + \frac{V_{DS}(t) - \phi_2(t)}{\tau_{2\_emit}}. \quad (4)$$

Equations (2)–(4) are state equations—first-order differential equations for the evolution of the key dynamical (state) variables that are arguments of the electrical constitutive relations appearing in (1). The functions f_1 and f_2 appearing in (3) and (4), respectively, are diode-like nonlinearities that account for preferential trapping rates when the instantaneous gate (drain) voltage becomes more negative (positive) than the values of ϕ_1 and ϕ_2. The parameters τ_1 and τ_2 are characteristic emission times, typically assumed to be very long compared to the RF time scales.
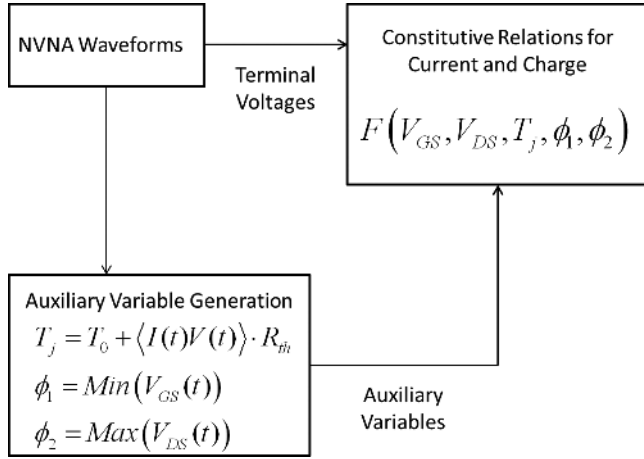
Fig. 8. Model identification process with explicit functional formulas for state variable values.

Fig. 9. Parametric plot of measured dynamic load lines used for large-signal model generation: (*top*) $V_{GS}(t)$ versus $V_{DS}(t)$ and (*bottom*) $I_{DS}(t)$ versus $V_{DS}(t)$.

B. State Variable Identification From Large-Signal Waveforms

The model identification process is then to define the detailed nonlinear constitutive relations defining the current and charge function, I_D, and Q_D, respectively, in (1), as nonlinear functions of all five independent variables.

Remarkably, it is possible to identify these complicated multivariate dependencies directly from measured device large-signal waveform data under purely steady-state conditions. That is, by exciting the device with CW large-signal sinusoidal inputs at each port, varying the input and output port power, the relative phase between input and output port excitations, dc-bias conditions, and backside (or case) temperatures, it is possible to sample the detailed dependence of the current and charge functions everywhere in the operating space of the five independent variables [3].

To understand why this is true, we note that at typical RF frequencies, the temperature cannot follow the RF signal. At steady state, therefore, the junction temperature assumes a fixed value equal to the increase over ambient temperature given by the average dissipated power over an RF cycle, multiplied by the thermal resistance (a parameter that can be extracted by other techniques [5]). For each steady-state periodic large-signal RF load line, the junction temperature T_j can be calculated by the simple formula in Fig. 8.

For trap capture rates much faster than the RF signal, and for emission rates much slower than the RF signal, the trap circuits operate as peak detectors. The value, under steady-state large-signal conditions for these trap states, therefore, becomes the minimum (maximum) of the excursion of the $V_{GS}(t)$ ($V_{DS}(t)$) waveforms over a given period.

Thus, the large-signal RF CW excitation sets the values of the slow dynamical variables, but the measurement allows for their computation using simple formulas. The present methodology associates a set of three auxiliary values for T_j, ϕ_1, and ϕ_2 with each waveform measurement, one waveform per power, dc bias, backside temperature, and complex load. That is, the junction temperature value and values for the trap states are *functionals* of the particular waveform. The model identification flow is given by Fig. 8.
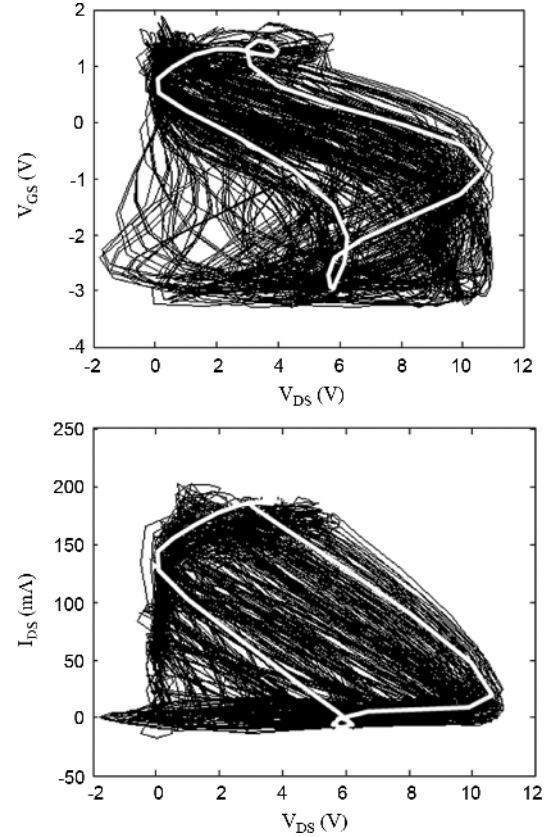
It is possible to engineer a complete set of dynamic load lines by properly varying the load (through passive tuning or active injection at port 2), RF power, dc-bias conditions, and backside temperatures to cover the entire large-signal operating range of the device.

An example of measured waveforms, in $V_{GS} - V_{DS}$ space and $I_D - V_{DS}$ space is given in Fig. 9. It is important to note that this region extends far beyond the conventional region of dc operation of a device. A great advantage of NVNA data is that the extreme regions of the device operation can be characterized with much less degradation of the transistor. This is because the instantaneous voltages only enter the high-stress regions for sub-nanosecond periods as the device is stimulated with signals at 1 GHz or higher frequency. For the same reason, less energy is dissipated in the device at high instantaneous power regions than under dc conditions. The larger domain of device operation means the need for the final model to extrapolate during large-signal simulation is dramatically reduced or even eliminated completely. Actual nonlinear data obtained under realistic operating conditions means the modeling process does not have to "extrapolate" from linear and dc data to predict nonlinear RF behavior, as do the approaches based on small-signal data parameterized by dc-bias conditions. Moreover, the NVNA data provides detailed waveforms for comprehensive nonlinear model validation without the need for additional instruments, such as spectrum analyzers, that only give the magnitude of the generated spectrum—the NVNA measures the magnitudes and the phases of the distortion products.

C. Constitutive Relation Approximation

The remaining task is to smoothly approximate the device-specific nonlinear constitutive relations from the sampled values of the waveforms and dynamical variables obtained by the procedure outlined in Section III-B. In principle, any of a wide variety of fitting techniques can be used to fit the currents and charges as nonlinear functions of each of the five independent variables. In the work reported here, advanced artificial neural network (ANN) training techniques were implemented [3]. ANNs have a great many advantages over other methods [6]. Unlike polynomials, ANNs have infinitely many nonzero partial derivatives, useful for high-order distortion simulation at relatively small amplitudes, and ANNs do not diverge beyond the training region, a property useful for robust convergence in nonlinear simulators. ANNs can be trained easily on the scattered data in the multivariate space of independent variables. ANN-based nonlinear constitutive relations can be easily compiled into standard nonlinear circuit simulators.

Previously, specific and simplified assumptions about how the trap state values affect the shape of the current and charge-storage characteristics had to be assumed, typically by modifying the intrinsic terminal voltages or parameters in the constitutive relations such as threshold voltage [1], drain resistance [2], or adding simple self back-gating equivalent circuits [17].

There is great insight that can be obtained by looking at the constructed constitutive relations based on large-signal steady-state waveforms. Two examples of generated intrinsic constitutive relations for different sets of trap states are shown in Fig. 10. The model current constitutive relations corresponding to extreme trap states [see Fig. 10 (top)] bears a striking resemblance to pulsed bias characterization from quiescent bias points associated with the trap state biases [4]. The advantage of the NVNA approach is that the model characteristics are inferred from DUT responses to signals typically three or more orders of magnitude faster than what can be measured with most pulsed systems that are typically limited from 0.1 to 1 μm.

The complete model solves for the trap states, junction temperature, and currents self-consistently during simulation. When embedded back into the parasitic model, final comparison can be made to measured data. Fig. 11 shows the validation with measured dc I–V curves. Note how different the static nonisothermal I–V curves in Fig. 11 are from the intrinsic model constitutive relations under the conditions of Fig. 10 (top) and (bottom).

Although not shown here, the dependence of the key constitutive relations on five state variables provides sufficient degrees of freedom to fit the bias dependence of the small-signal model over the entire bias space at both dc and high frequencies [3]. That is, frequency dispersion phenomena are predicted accurately under small- and large-signal conditions as a consequence of incorporating properly both dynamic trapping and electrothermal effects. Models with just electrothermal effects are not capable of such good fits to both dc and high-frequency behavior at all biases.

Fig. 12 shows the nonlinear validation results for the advanced FET model for power-dependent gain and bias current versus power. The distinctive car-shaped gain compression
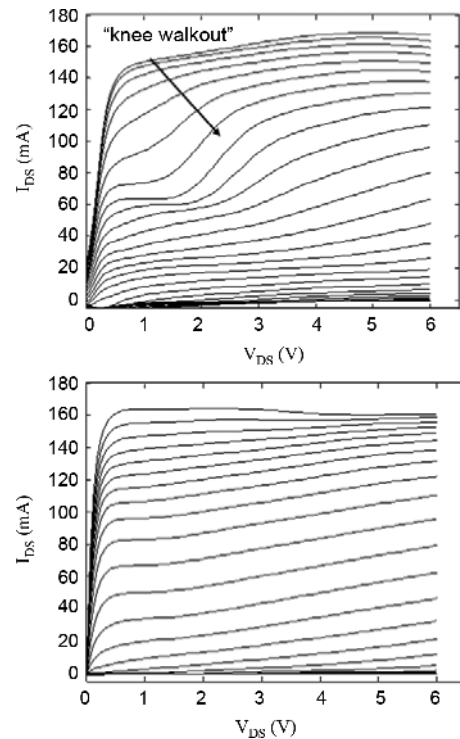

Fig. 10. I_D intrinsic model constitutive relation. (*top*) For large fixed trap state values indicating knee collapse. (*bottom*) When trap states move with dc bias.
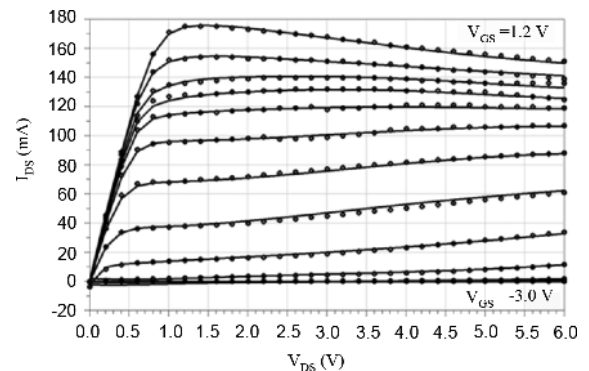

Fig. 11. Simulation of full model (solid line) and measured (symbols) comparison of extrinsic dc I–V curves.

characteristic and significant nonmonotonic dependence of the bias current with power is a result of the dynamics of drain-lag and the detailed constitutive relation obtained with the ANN training.

Fig. 13 shows the model validation of distortion versus power for this device, validating both the dynamical description and accuracy and robustness of the ANN approach to modeling the complicated constitutive relations.

Fig. 14 shows the model simulation and NVNA waveform validation measurement of a steady-state dynamic load line at a complex output impedance not used in the ANN training. Also shown is the set of dc-bias conditions demonstrating how much
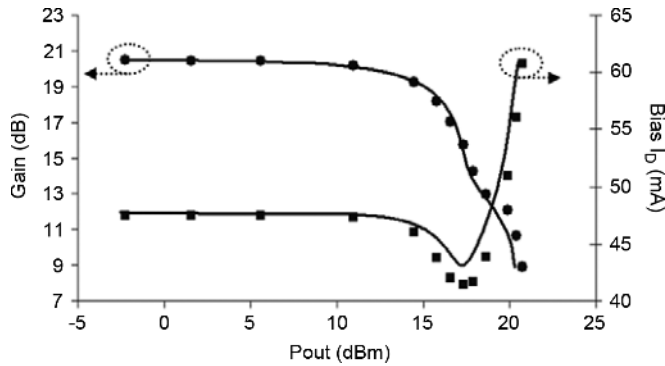
Fig. 12. Modeled and measured gain versus P_{out}, and dc-bias current versus P_{out}.
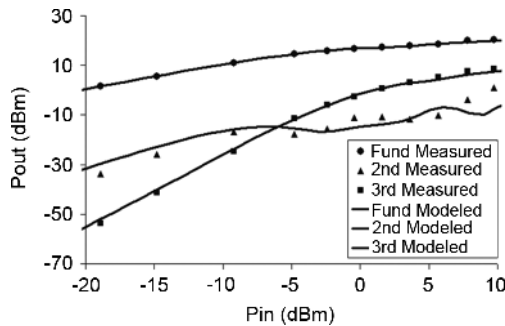

Fig. 13. Measured and modeled fundamental and harmonic output power versus input power.
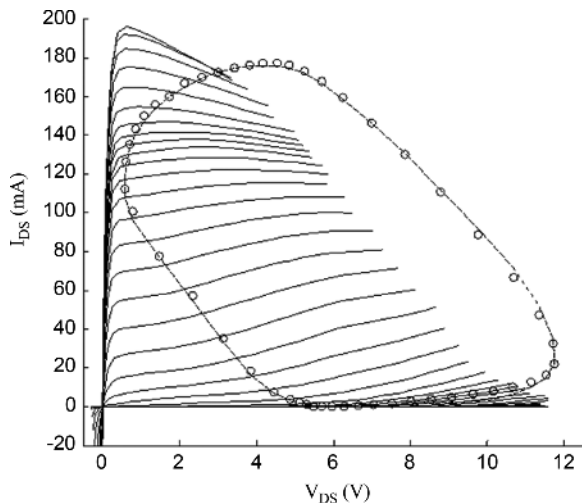

Fig. 14. Simulated (dashed line) and independently measured large-signal dynamic load line (circles) extending beyond dc *I–V* data (plain line).

more of the device operating range is accessible by NVNA characterization methods.

IV. DYNAMIC BEHAVIORAL MODELING

Behavioral modeling of microwave components provides an alternate modeling approach that bypasses the need for detailed device and circuit modeling of the circuits involved.

Unfortunately as was discussed in Section I, behavioral models that can handle hard nonlinear behavior, like AM-to-PM characteristic or the PHD model, have no straightforward way

of dealing with memory effects. As explained in [20], memory effects can be classified as either short term or long term. Short-term memory effects are caused by physical dynamics that occur at the timescale of the carrier. This is, for example, the case with frequency dispersion in filters. Long-term memory effects are caused by physical dynamics happening at a much longer timescale. This is, for example, the case for dynamic self-heating or self-biasing effects. In this section, a new behavioral model that can easily handle long-term memory effects for a wide range of modulation bandwidths is presented.

A. Introduction of the Model Equations

Unlike existing approaches that deal with long-term memory, the approach we present is not derived from Volterra theory, but is derived from simple first principles. The result is a model that describes hard nonlinear behavior, as well as long-term memory effects. In all of the following, a simple matched system will be considered, having only one input signal $A(t)$, the incident fundamental, and one output signal, $B(t)$, typically the amplified signal. Note that all signals are considered to be complex envelope representations of a modulated carrier. All of the concepts can be extended toward mismatched conditions by including the incident and scattered wave signals, $A_2(t)$ and $B_1(t)$, at both signal ports, and can be extended towards harmonic effects like the models described in [12] by including the harmonics of the carrier signal. The above extensions are outside the scope of this paper. It will be assumed that the input signal can be represented as a modulated carrier with a fixed frequency. Our model is a so-called low-pass equivalent model that only processes the envelope information signal [11]. The new model equation, which will be derived later, is

$$B(t) = \left(F_{\mathrm{CW}}\left(|A(t)|\right) + \int_0^\infty G\left(|A(t)|, |A(t-u)|, u\right) du \right) \cdot \exp\left(j\phi\left(A(t)\right)\right). \quad (5)$$

The basic idea of the new approach is that $B(t)$ can be written as the superposition of a static part, represented by the nonlinear function $F_{\mathrm{cw}}(.)$, and a dynamic part, represented by a simple integral over time of a nonlinear function $G(.)$. Note that $F_{\mathrm{cw}}(.)$, as well as the integrated function $G(.)$, are a function of the instantaneous amplitude of the input signal $A(t)$. The dependency on the phase of the input signal is represented as a separate multiplicative vector $\exp\left(j\phi\left(A(t)\right)\right)$. To simplify the mathematical notation, introduce $\Phi(t) = \exp\left(j\phi\left(A(t)\right)\right)$.

As shown in [12], this phase dependency can be explained by the principle of time-delay invariance. The static part behaves like a classic PHD model. The dynamic part is original as it represents the long-term memory effects. These are described as the integral effect of a general nonlinear function of the instantaneous amplitude of the input signal $A(t)$, the past values of the input signal $A(t-u)$, and how long ago that past value occurred (variable u). The model (5) can be derived by using the notion of hidden variables [18].

B. Derivation of the Model Equation

To derive (5), one starts with writing the simple static PHD-model equation, as described in [12], which is equivalent to a

simple compression and AM-to-PM characteristic. This results in

$$B(t) = F(|A(t)|) \; \Phi(t). \qquad (6)$$

In (6), the function $F(.)$ represents the mapping from the amplitude of the input signal to the corresponding output signal amplitude and phase, whereby the dependency on the phase of the input signal is represented as a separate multiplicative vector.

As described in [18], memory effects can be introduced by making use of one or more hidden variables. The idea is that, in a system with memory, the mapping from the input signal to the output signal is no longer a function of the input signal amplitude only, but is also a function of an arbitrary number N of *a priori* unknown hidden variables, denoted $h_1(t), h_2(t), \ldots, h_N(t)$. These variables represent time varying physical quantities inside the component, such as temperatures, bias voltages or currents, trapping states, ... that influence the mapping from the input signal to the output signal. With the introduction of the hidden variables, (6) becomes

$$B(t) = F(|A(t)|, h_1(t), h_2(t), \ldots, h_N(t)) \cdot \Phi(t). \quad (7)$$

To make (7) useful in the context of a black-box modeling approach, one further needs to make an assumption regarding the relationship between the hidden variables and the input signal. Note that one could use *a priori* information on the physics of the device-under-test (DUT) to find this relationship. The black-box assumption on the relationship between the input signal and the hidden variables is mathematically expressed as

$$\forall i : h_i(t) = \int_0^\infty P_i(|A(t-u)|) \; k_i(u) \; du. \qquad (8)$$

Equation (8) expresses that the ith hidden variable is generated by a linear filter operation, characterized by its impulse response $k_i(.)$, which operates on a nonlinear function $P_i(.)$ of the input signal amplitude $|A(.)|.P_i(.)$, can be interpreted as a source term that describes how the input signal is related to the excitation of a particular hidden variable; in general, this is a nonlinear relationship. A good example could be, e.g., that $P_1(.)$ describes the power dissipation as a function of the input signal, whereby $h_1(.)$ is the temperature. The impulse response $k_i(.)$ describes the actual dynamics of a hidden variable, e.g., $k_1(.)$ could describe a thermal relaxation. Note that the model as described in [13] is actually a special case of the above equations whereby there is only one hidden variable with

$$h_1(t) = \int_0^\infty |A(t-u)|^2 \; e^{-u/R} \frac{du}{R} \text{ or } P_1(x) = x^2$$

and

$$k_1(x) = \frac{1}{R} e^{-x/R}.$$

In [13], the equations are physically related to the assumption that the amplitude squared relates to the bias current flowing through a first-order linear bias circuit, which causes a variation

of the bias voltage, which modulates the PHD model. One can easily imagine many other physical effects that are described by (8), such as trapping effects and self-heating, whereby the dissipated power is a nonlinear function of $A(t)$ and causes temperature changes, which are linearly related to the dissipated power, and which modulate the PHD model.

The link with (5) is then made by assuming that one can linearize (7) in the hidden variables $h_i(t)$. In order to linearize, we need to choose an operating point. In the following, we choose to linearize around the steady-state solution for the hidden variables that corresponds to the instantaneous amplitude $|A(t)|$. In other words, we linearize around the steady-state solution that the system would reach assuming that we keep the input amplitude constant for all future time instants, the amplitude being equal to the instantaneous input amplitude. This linearization implicitly assumes that the deviations of the hidden variables from their steady-state solution are always small enough not to violate the superposition principle. These steady-state solutions are a function of the input signal amplitude and are noted in the following by $h_i^{\mathrm{CW}}(.)$. For example, the value of $h_i^{\mathrm{CW}}(X)$ is equal to the asymptotic value of the hidden variable $h_i(t)$ when one applies a constant input signal amplitude equal to X. Note that the argument of $h_i^{\mathrm{CW}}(.)$ is a signal amplitude, whereas the argument of the hidden variable $h_i(t)$ is time. The functions $h_i(t)$ are easily calculated by simply substituting $A(t-u)$ by a constant X in (7)

$$\forall i : h_i^{\mathrm{CW}}(X) = \int_0^\infty P_i(X) \; k_i(u) du = W_i \, P_i(X) \quad (9)$$

with

$$W_i = \int_0^\infty k_i(u) \; du \qquad (10)$$

Equation (7) can then be rewritten as

$$\begin{aligned} B(t) = F(|A(t)|, W_1 P_1(|A(t)|) \\ + \Delta_1(t), W_2 P_2(|A(t)|) + \Delta_2(t), \ldots) \cdot \Phi(t) \end{aligned} \quad (11)$$

with

$$\forall i : \Delta_i(t) = h_i(t) - W_i P_i(|A(t)|). \qquad (12)$$

In (12), the new variable $\Delta_i(t)$ represents the deviation of the hidden variable $h_i(t)$ from its steady-state solution corresponding to the instantaneous input signal amplitude. It is now further assumed that (7) can be linearized in these deviations from steady-state $\Delta_i(t)$. This results in

$$B(t) = \left(F_{\mathrm{CW}}(|A(t)|) + \sum_{i=1}^N D_i(|A(t)|) \Delta_i(t) \right) \cdot \Phi(t) \tag{13}$$

with

$$\begin{aligned} F_{\mathrm{CW}}(|A(t)|) \\ = F(|A(t)|, W_1 P_1(|A(t)|), W_2 P_2(|A(t)|), \ldots) \end{aligned} \quad (14)$$

and

$$D_i(x) = \left. \frac{\partial F}{\partial h_i} \right|_{(x, \, W_1 P_1(x), \, W_2 P_2(x), \ldots)}. \qquad (15)$$

The functions $D_i(.)$ represent the sensitivity of the output signal to variations of the ith hidden variable. For example, if $h_1(.)$ represents the temperature, $D_1(.)$ represents how sensitive the output signal is to temperature changes. Note that $D_i(.)$ is a general nonlinear function of the instantaneous input signal amplitude. It is perfectly possible for example that the output signal is highly sensitive to temperature changes for small input signals (corresponding to a temperature-dependent small-signal gain), but not at all for high input signals whereby the output is fully saturated.

In the structure of (13), one can distinguish a static part, represented by $F_{\mathrm{CW}}(.)$ and a dynamic part, represented by the summation over the hidden variables index "i." The static part, which can be derived from $F(.)$ by using (14), corresponds to a classic static PHD model. Note that the subscript "CW" is used in $F_{\mathrm{CW}}(.)$ since this function corresponds to the response of the DUT to a single-tone CW excitation. Substitution of $h_i(t)$ in (12) using (8) and subsequently substitution of $\Delta_i(t)$ in (13) using (12) results in

$$
\begin{aligned}
B(t) = {} & F_{\mathrm{CW}}(|A(t)|) \cdot \Phi(t) + \sum_{i=1}^{N} D_i(|A(t)|) \\
& \cdot \left(\int_0^{\infty} P_i(|A(t-u)|)\, k_i(u)\, du - W_i P_i(|A(t)|) \right) \Phi(t).
\end{aligned}
$$
(16)

Using (10) and changing the order of summation and integration, one can write

$$
\begin{aligned}
B(t) = {} & F_{\mathrm{CW}}(|A(t)|) \cdot \Phi(t) \\
& + \int_0^{\infty} \left\{ \sum_{i=1}^{N} D_i(|A(t)|) \right. \\
& \qquad \left. \cdot \left(P_i(|A(t-u)|) - P_i(|A(t)|) \right) k_i(u) \right\} du \\
& \cdot \Phi(t).
\end{aligned}
$$
(17)

In general, one may neither know what the hidden variables are, nor how many there are or how they interact with the system. Nevertheless, one can always define the multivariate function $G(x, y, u)$ by

$$
G(x, y, u) = \sum_{i=1}^{N} D_i(x)\left(P_i(y) - P_i(x)\right) k_i(u)
$$
(18)

such that

$$
\begin{aligned}
& B(t) \\
& = \left(F_{\mathrm{CW}}(|A(t)|) + \int_0^{\infty} G(|A(t)|, |A(t-u)|, u)\, du \right) \Phi(t)
\end{aligned}
$$
(19)

which is identical to (5). As such, we have been able to derive (5) from a hidden variables approach, which was the goal of this paragraph.

Equation (18) reveals that the multivariate function $G(x, y, u)$ is not arbitrary since the following relationship will always be valid:

$$
\forall x, u : G(x, x, u) = 0.
$$
(20)

As will be shown next, this property is key in the experimental determination of the function $G(.)$.

C. Model Identification

An important question is how one can determine the model functions $F_{\mathrm{CW}}(.)$ and $G(.)$ from experiments. As will be shown in the following, the model can completely be identified from measuring a simple set of large-signal step responses. The set of input step signals should be such that one switches, at time zero, from one value to another value, covering the whole range of possible input signal values.

Consider the application of a step input signal, whereby $A(t) = A_1$ for $t < 0$ and $A(t) = A_2$ for $t >= 0$. The solution for such a large-signal step response $B(t)$ will be noted as $B_{LS}(A_1, A_2, t)$ and is then given by

$$
\begin{aligned}
\forall t < 0 : {} & B_{LS}(A_1, A_2, t) \\
& = F_{\mathrm{CW}}(|A_1|)\, \exp(j\varphi(A_1)) \\
\forall t \geq 0 : {} & B_{LS}(A_1, A_2, t) \\
& = F_{\mathrm{CW}}(|A_2|)\, \exp(j\varphi(A_2))
\end{aligned}
$$
(21)

$$
\begin{aligned}
& + \int_0^{t} G(|A_2|, |A_2|, u)\, du \cdot \exp(j\varphi(A_2)) \\
& + \int_t^{\infty} G(|A_2|, |A_1|, u)\, du \cdot \exp(j\varphi(A_2)).
\end{aligned}
$$
(22)

Equations (20) and (22) can be rewritten as

$$
\begin{aligned}
\forall t \geq 0 : {} & B_{LS}(A_1, A_2, t) \\
& = F_{\mathrm{CW}}(|A_2|)\, \exp(j\varphi(A_2)) \\
& + \int_t^{\infty} G(|A_2|, |A_1|, u)\, du \cdot \exp(j\varphi(A_2)).
\end{aligned}
$$
(23)

Taking the derivative versus "t" at both sides of the above equation results in

$$
\begin{aligned}
\forall t \geq 0 : {} & \frac{dB_{LS}(A_1, A_2, t)}{dt} \\
& = -G(|A_2|, |A_1|, t) \cdot \exp(j\varphi(A_2))
\end{aligned}
$$
(24)

$$
\begin{aligned}
\forall t \geq 0 : {} & G(|A_2|, |A_1|, t) \\
& = -\frac{dB_{LS}(A_1, A_2, t)}{dt} \cdot \exp(-j\varphi(A_2)).
\end{aligned}
$$
(25)

Note that, as a consequence of the inherent causality of (5), $G(x, y, t)$ only needs to be defined for positive values of "t." Equation (25) has important consequences. First of all, it allows for a straightforward measurement of the function $G(x, y, t)$ by taking the inverse of the derivative of the step response when starting with input amplitude "y" and switching to input amplitude "x" at time 0.

A second important consequence is that there is a one-to-one mapping between the model and the step responses. Given that one can measure all possible large-signal step responses, one could apply the principle of Ockham's razor and raise the question "what is the simplest nonlinear dynamic model that is able to generate the measured large signal step responses?" The answer is the model as described by (5). This is definitely a significant advantage of the new modeling approach when compared to existing black-box memory modeling methods.

D. Simplification in Case of Fast Varying Input Envelopes

As explained in [19], an interesting case occurs when the input envelope varies fast relative to the timescale of the $G(.,.,.)$ kernel and whereby the input envelope has ergodic properties (this means that the statistical characteristics are constant over time). The model will then be seemingly static, but with a static characteristic that is determined by the probability density function of the input envelope.

Consider (5),

$$B(t) = \left(F_{CW}(|A(t)|) + \int_0^\infty G(|A(t)|, |A(t-u)|, u) \, du \right) \cdot \exp(j\phi(A(t))). \quad (26)$$

For a fast variation of $A(.)$, one can write the dynamic memory term on the right-hand side of (26) as follows [19]:

$$\int_0^\infty G(|A(t)|, |A(t-u)|, u) \, du$$
$$\approx \int_0^\infty \int_0^\infty G(|A(t)|, a, u) \, p(a) \, da \, du \quad (27)$$

with $p(a)$ equal to the probability density function of the input envelope amplitude, which is now considered to be a stochastic variable denoted by "a." Changing the order of integration on the right-hand side of (27) leads to

$$\int_0^\infty G(|A(t)|, |A(t-u)|, u) \, du$$
$$\approx \int_0^\infty \left(\int_0^\infty G(|A(t)|, a, u) \, du \right) p(a) \, da. \quad (28)$$

Next we will calculate the inner integral of the right-hand side of (28) by using identity (25)

$$\int_0^\infty G(|A(t)|, a, u) \, du = -\int_0^\infty \frac{dB_{LS}(a, |A(t)|, u)}{du} du \quad (29)$$

whereby one notes that the integral operator at the right-hand side annihilates the derivative operator. The result is

$$\int_0^\infty G(|A(t)|, a, u) \, du$$
$$= -B_{LS}(a, |A(t)|, \infty) + B_{LS}(a, |A(t)|, 0). \quad (30)$$

As $B_{LS}(a, |A(t)|, \infty) = F_{CW}(|A(t)|)$, subsequent substitution of (30) in (28) and in (26) results in

$$B(t) \approx \left(\int_0^\infty B_{LS}(a, |A(t)|, 0) p(a) \, da \right) \cdot \exp(j\varphi(A(t))). \quad (31)$$

From (31), one can conclude that the relationship between the input envelope $A(t)$ and the output envelope $B(t)$ will be given by a seemingly static characteristic. This static characteristic is equal to a weighted average of the large-signal step responses evaluated at time zero, using as a weight function the probability density function of the envelope input values. Since this simplified dynamic X-parameter model is typically valid for a wideband modulation, which is equivalent to a fast varying envelope, it is called a wideband X-parameter model.

E. Experimental Model Identification and Validation

The dynamic X-parameter model and the wideband X-parameter model were first experimentally validated using a Mini-Circuits ZFL−11AD+ packaged microwave amplifier. Two measurement platforms were used for the extraction of the model, as well as for the validation of the model. A first measurement platform was based on a customized PNA-X, a second measurement platform was based on a PXA digital spectrum analyzer in combination with an ESG or MXG digital synthesizer, all from Agilent Technologies Inc., Santa Rosa, CA. Note that the model extraction is based on using large-signal steps as excitation signals, whereas the model validation is based on using two-tone signals, as well as a periodically modulated signals with WCDMA characteristics.

The model extraction, as well as the model validation was also performed using a carrier frequency equal to 1.75 GHz. Measured results of the model extraction are represented in Figs. 15–17.

For the large-signal step experiments, the values of A_1 and A_2 are both swept from close to zero to $0.2V_p$. Note that the maximum input value of $0.2V_p$ corresponds to about 5 dB of compression. Fig. 15 shows the large signal output steps for a fixed value of $0.14V_p$ for A_1 and for sweeping A_2. Note the rather slow frequency ringing in the step response.

Both the static X-parameter kernel $F_{CW}(.)$ and the memory kernel $G(.)$ are extracted from the above-mentioned set of large-signal step measurements. The $F_{CW}(.)$ corresponds to a classic AM–AM AM–PM characteristic. The new memory kernel $G(.)$ is hard to visualize, as it is defined as a complex number defined on a square prism (two sides correspond to the amplitude ranges of A_1 and A_2, the third side corresponds to the time duration of the memory). Fig. 16 shows the amplitude of the memory kernel $G(.)$ at a fixed time instant (equal to 50 ns) as a function of A_1 and A_2. Fig. 17 shows the time evolution of the $G(.)$ kernel for a small A_1 and a large A_2. The upper graphic has a time range of 60 μs, where as the bottom graphic zooms in on the first 5 μs only.

Validation results are shown in Figs. 18 and 19. The first experiment is based on two-tone signals with a tone separation of 19.2 kHz, repeated for four different power levels. The measured input signals are also used to stimulate the extracted dy-
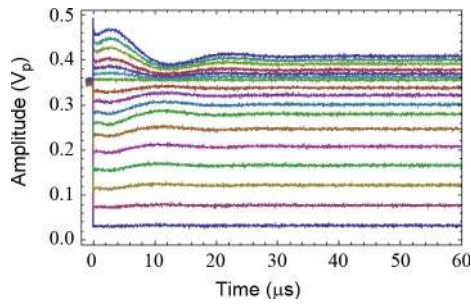
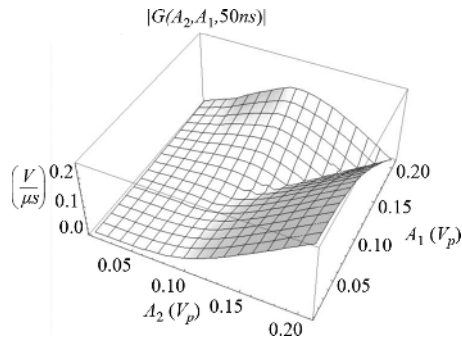Fig. 15. Measured large-signal output steps for a fixed large value of A_1 (A_2 swept).

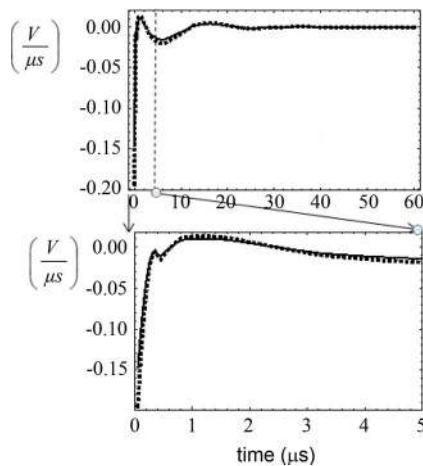Fig. 16. Memory kernel at fixed $t = 50$ ns for Mini-Circuit ZFL11AD+ amplifier.

Fig. 17. Memory kernel (V/μs) versus time (μs) for a step transition from low (0.01 V_p) to high (0.2 V_p) input amplitude. (solid = real part; dashed = imaginary part).

namic X-parameter model inside the ADS2009 envelope simulator. The simulated output signals, as well as the measured output signals, are represented in Fig. 18. One notes the skewing of the output waveform, due to long-term memory effects, and the accuracy of the model to predict this behavior.

An interesting way to look at this data is the dynamic compression characteristic shown in Fig. 19. It is a plot of the instantaneous output amplitude versus the instantaneous input amplitude for the four two-tone experiments. The modeled and the measured results are shown. As one notes, there is a significant looping effect for the higher power level, a clear manifestation of long-term memory effects, which is accurately modeled by
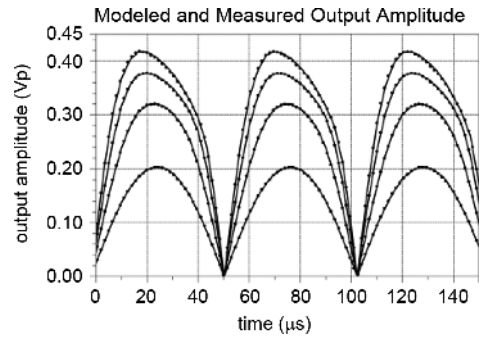

Fig. 18. Measured and modeled amplitude of the output envelope (19.2-kHz tone spacing) (solid = measured, dots = modeled).
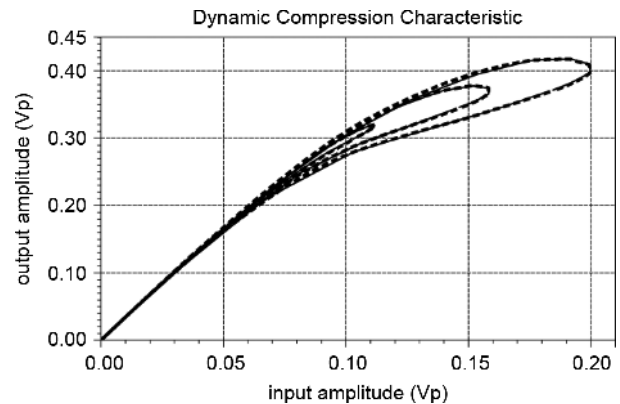

Fig. 19. Measured and modeled dynamic compression characteristic (measured: solid, modeled: dashed).
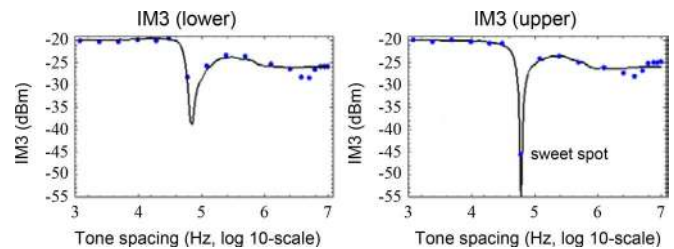

Fig. 20. Measured and modeled IM3 amplitude for two-tone experiment (solid = modeled, dots = measured).

the dynamic X-parameter model. Whereas the tone spacing in Fig. 18 was fixed at 19.2 kHz, it is interesting to check what happens if one sweeps the tone spacing across a wide range, in our case, almost four decades from 1.2 kHz to 10 MHz. Fig. 20 shows the measured and modeled third-order-intermodulation (IM3) product amplitudes (both upper and lower) as a function of the frequency spacing, at a −10-dBm fixed input amplitude level per tone. Note that the model predicts the resonance around 60-kHz tone spacing very well. The model also predicts the asymmetry in the lower and upper IM3 characteristics.

More advanced modulated signals were also applied, as shown in Fig. 21. In this figure, we plot the modeled and measured spectral regrowth when stimulating the amplifier with a signal with WCDMA characteristics (matching modulation bandwidth and statistics). To show the improvement of the dynamic X-parameter model over a static X-parameter approach, the prediction based on a static (or CW) X-parameter
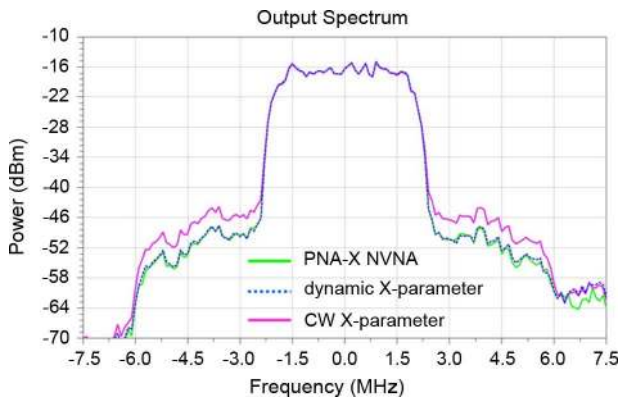
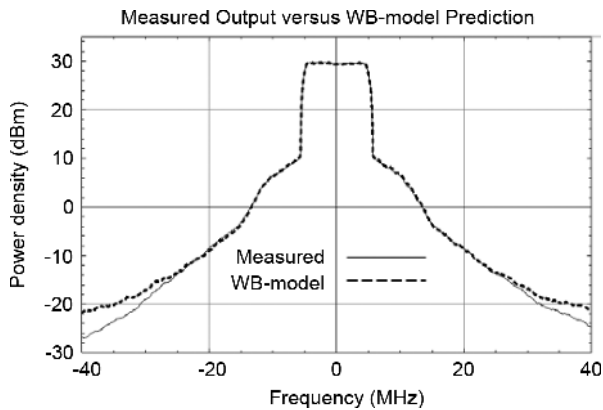Fig. 21. Spectral regrowth using a signal with WCDMA characteristics.

Fig. 22. Spectral regrowth prediction using wideband X-parameters (25-W GaN amplifier of CREE @ 2-GHz carrier, LTE signal is used).

approach is included. One notes that the dynamic X-parameter predicts the spectral regrowth with great accuracy with an error less than 1 dB, whereas the CW X-parameter systematically overestimates the spectral regrowth by about 5 dB.

Results on the validation of the wideband X-parameter modeling approach, applied to the ZFL − 11AD+, can be found in [19]. The wideband X-parameter model was also validated using a CREE CMP2560025 F amplifier. This is a 25-W 2500–6000-MHz GaN monolithic microwave integrated circuit (MMIC) PA. A setup using an MXG digital synthesizer and an PXA digital spectrum analyzer were used for extracting the wideband X-parameter kernel and for performing the validation measurements using an LTE signal at a 2-GHz carrier. The result is shown in Fig. 22. One notes that the spectral regrowth is accurately predicted, with an error less than 0.3 dB across a modulation bandwidth of about 40 MHz.

V. CONCLUSIONS

Low-frequency memory effects associated with self-heating, traps, parasitic BJT, and biasing circuits can greatly impact the large-signal RF performance of transistors and amplifiers excited by modulated RF signals. In this paper, the authors have presented new synergetic approaches for the: 1) characterization; 2) device modeling; and 3) behavioral modeling of power transistors and amplifiers with strong memory effects.

In the first part of this paper (Section II), the use of CW and pulsed RF large-signal measurements was investigated to quantify the impact of thermal and electrical memory effects upon the large-signal RF performance of transistors. Pulsed-RF load–pull measurements from realistic average hot bias condition yield a more realistic representation of the peak power response of transistors excited with modulated signals with high PAPR than can be obtained by CW load–pull measurements. For the SOI-MOSFET transistor considered, an improved efficiency of 14% was observed under pulsed operation compared to CW operation yielding a peak PAE of 73.7%. The results obtained further demonstrated that slow memory effects do not respond solely to the average bias or temperature, but can also depend on the RF load lines. Similar effects have been observed in GaN were intermittent high drain voltages can charges traps and induce knee voltage walkout degrading in that case the average performance of the transistor. The understanding of these processes is critical for both the design and simulation of amplifiers involving transistors exhibiting memory effects.

In the second part of the paper (Section III), an RF device model capable of accounting for memory effects under large-signal operation was presented. The methodology used is a systematic, general, and powerful approach for characterizing and identifying an advanced self-heating and trap-dependent nonlinear simulation model directly from large-signal waveform data from an NVNA. The detailed nonlinear model constitutive relations for drain current and charge as functions of instantaneous terminal voltages, dynamic junction temperature, and two species of trap states, are identified and constructed using ANN modeling technology. The process is more procedural, more general, and more accurate than previous methods based on pulsed I–V/S-parameter data and simplified constitutive relations based on particular assumed trapping mechanisms. The model is implemented in a commercial nonlinear circuit simulator and is usable in all analysis modes (e.g., transient, harmonic balance, circuit envelope, small-signal, etc.). The model has been validated extensively for dc, linear, and large-signal conditions.

In the third part of this paper (Section IV), dynamic X-parameters and a simplified version called wideband X-parameters were introduced as novel ways to build behavioral models for RF PAs that include long-term memory effects. The kernels of these models can be extracted by using a set of large-signal step response measurements, which are readily straightforward to perform.

The resulting measurement-based models can be implemented in a commercial complex envelope simulator and accurately predict the behavior of the amplifier, including long-term memory effects. Unlike any existing black-box modeling approach, the dynamic X-parameter model is valid for a wide range of signals: for high power and low power, for slow as well as fast modulation, for pulsed signals, for two-tone signals, and for wideband modulated signals like, e.g., WCDMA and LTE. The model remains accurate for all possible modulation formats, independent of the probability density function or derived figures-of-merit (like, e.g., PAPR) of the input signal.

The various measurement and modeling techniques presented in this paper for transistors and amplifiers with strong memory effects should benefit the PA community in multiple ways. The

pulsed-RF load–pull characterization can be directly used for amplifier design as in the case of load modulation [21] besides being used for device model extraction or verification. The enhanced transistor model that accounts for memory effects provides more realistic simulations for PA design. Finally, the novel PA behavioral model can greatly benefit the linearization of PAs exhibiting memory effects.

ACKNOWLEDGMENT

REFERENCES

[1] O. Jardel, F. DeGroote, T. Reveyrand, J. C. Jacquet, C. Charbonniaud, J. P. Teyssier, D. Floriot, and R. Quere, "An electrothermal model for AlGaN/GaN power HEMTs including trapping effects to improve large-signal simulation results on high VSWR," *IEEE Trans. Microw. Theory Tech.*, vol. 55, no. 12, pp. 2660–2669, Dec. 2007.

[2] A. M. Conway and P. M. Asbeck, "Virtual gate large-signal model of GaN HFETs," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2007, pp. 605–608.

[3] J. Xu, J. Horn, M. Iwamoto, and D. E. Root, "Large-signal FET model with multiple time scale dynamics from nonlinear vector network analyzer data," in *IEEE MTT-S Int. Microw. Symp. Dig.*, May 2010, pp. 417–420.

[4] A. E. Parker and D. E. Root, "Pulse measurements quantify dispersion in pHEMTs," in *URSI Int. Signals, Syst., Electron. Symp.*, Pisa, Italy, Sep. 1998, pp. 444–449.

[5] M. Iwamoto, J. Xu, and D. E. Root, "DC and thermal modeling: III-V FETs and HBTs," in *Nonlinear Transistor Model Parameter Extraction Techniques*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[6] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. New York: Prentice-Hall, 1999.

[7] S. A. Albahrani, J. G. Rathmell, and A. E. Parker, "Characterizing drain current dispersion in GaN HEMTs with a new trap model," in *Proc. 39th Eur. Microw. Conf.*, Rome, Italy, Oct. 2009, pp. 1692–1695.

[8] I. Suh, P. Roblin, Y. Ko, C.-K. Yang, A. Malonis, A. Arehart, S. Ringel, C. Poblenz, Y. P. Speck, and U. Mishra, "Additive phase noise measurements of AlGaN/GaN HEMTs using a large signal network analyzer and tunable monochromatic light source," in *74th ARFTG Conf. Dig.*, Boulder, CO, Dec. 2009, pp. 1–5.

[9] P. Roblin, *Nonlinear RF Circuits and Nonlinear Vector Analyzer*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[10] A. S. Roy and C. C. Enz, "Analytical modeling of large-signal cyclo-stationary low-frequency noise with arbitrary periodic input," *IEEE Trans. Electron Devices*, vol. 54, no. 9, pp. 2537–2545, Sep. 2007.

[11] J. C. Pedro and S. A. Maas, "A comparative overview of microwave and wireless power-amplifier behavioral modeling approaches," *IEEE Trans. Microw. Theory Tech.*, vol. 53, no. 4, pp. 1150–1163, Apr. 2005.

[12] J. Verspecht and D. Root, "Polyharmonic distortion modeling," *IEEE Microw. Mag.*, vol. 7, no. 3, pp. 44–57, Jun. 2006.

[13] J. Verspecht, D. Gunyan, J. Horn, J. Xu, A. Cognata, and D. Root, "Multi-port, and dynamic memory enhancements to PHD nonlinear behavioral models from large-signal measurements and simulations," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2007, pp. 969–972.

[14] F. De Groote, O. Jardel, T. Reveyrand, J. -P. Teyssier, and R. Quéré, "Very small duty cycles for pulsed time domain transistor characterization," *37th Eur. Microw. Conf.*, vol. 4, pp. 112–117, Jun. 2008.

[15] F. De Groote, P. Roblin, J. P. Teyssier, C. Yang, S. Doo, and M. V. Bossche, "Pulsed multi-tone measurements for time domain load pull characterizations of power transistors," in *73rd ARFTG Conf. Dig.*, Boston, MA, May 2009, pp. 1–4.

[16] P. Roblin, P. Y. S. Ko, C. K. Yang, I. Suh, and S. J. Doo, "NVNA techniques for pulsed RF measurements," *IEEE Microw. Mag.*, vol. 13, no. 2, pp. 65–76, Apr. 2011.

[17] Z. Ouarch, J. M. Collantes, J. P. Teyssier, and R. Quere, "Measurement based nonlinear electrothermal modeling of GaAs FET with dynamical trapping effects," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 1998, pp. 599–602.

[18] J. Verspecht, J. Horn, L. Betts, C. Gillease, D. E. Root, R. Pollard, and D. Gunyan, "Extension of X-parameters to include long-term dynamic memory effects," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2009, pp. 741–744.

[19] J. Verspecht, J. Horn, and D. E. Root, "A simplified extension of X-parameters to describe memory effects for wideband modulated signals," in *75th ARFTG Conf. Rec.*, Jun. 2010, pp. 1–6.

[20] A. Soury and E. Ngoya, "Using sub-systems behavioral modeling to speed-up RFIC design optimizations and verifications," in *INMMIC Conf. Rec.*, CITY, Spain, 2008, pp. 165–168.

[21] H. Jang, Y. Ko, P. Roblin, C. Yang, and H. Park, "Pulsed load–pull based optimal load-modulation PA design methodology for average efficiency enhancement," in *78th ARFTG Conf. Dig.*, Phoenix, AZ, Dec. 2011, pp. 1–6.

Patrick Roblin (M'85) was born in Paris, France, in September 1958. He received the Maitrise de Physics degree from the Louis Pasteur University, Strasbourg, France, in 1980, and the D.Sc. degrees in electrical engineering from Washington University, St. Louis, MO, in 1984.

In 1984, he joined the Department of Electrical Engineering, The Ohio State University (OSU), Columbus, where he is currently a Professor. He is the founder of the Non-Linear RF Research Laboratory, OSU. At OSU, he has developed two educational RF/microwave laboratories and associated curriculum for training both undergraduate and graduate students. He authored *Nonlinear RF Circuits and Nonlinear Vector Network Analyzers* (Cambridge Univ. Press, 2011). He coauthored the textbook *High-Speed Heterostructure Devices* (Cambridge Univ. Press, 2002). His current research interests include the measurement, modeling, design, and linearization of nonlinear RF devices and circuits such as oscillators, mixers, and PAs.

David E. Root (M'89–SM'01–F'02) received the B.S. degrees in physics and mathematics in 1978, and the Ph.D. degree in physics in 1986, from the Massachusetts Institute of Technology (MIT), Cambridge.

In 1985, he joined the Hewlett-Packard Company (now Agilent Technologies), where he has held technical, management, and strategic positions. He is currently an Agilent Research Fellow with the Measurement Research Laboratories, Agilent Technologies, Santa Rosa, CA. His current responsibilities include nonlinear behavioral and device modeling, large-signal simulation, and nonlinear measurements for new technical capabilities and business opportunities for Agilent Technologies. In Fall 2005, he was a Visiting Scholar and Lecturer with the University of California at San Diego, La Jolla. He has authored or coauthored over 100 peer-reviewed technical journal papers, international conference and workshop papers, short courses, books, and book chapters. He coauthored and coedited *Nonlinear Transistor Model Parameter Extraction Techniques* (Cambridge Univ. Press, 2011) and *Fundamentals of Nonlinear Behavioral Modeling for RF and Microwave Design* (Artech House, 2005). He serves as reviewer for several technical journals and international conferences.

Dr. Root was an IEEE Microwave Theory and Techniques Society (IEEE MTT-S) Distinguished Microwave Lecturer (2006–2008). He is past chair of the IEEE Working Committee on Computer-Aided Design (CAD) (MTT-1). Since 1995, he has been a member of the Technical Program Committee of the IEEE MTT-S International Microwave Symposium (IMS). He was a corecipient of the 2007 IEEE ARFTG Technology Award.

Jan Verspecht (M'94–SM'05–F'07) was born in Merchtem, Belgium, on December 12, 1967. He received the Electrical Engineering and Ph.D. degrees from Vrije Universiteit Brussel (VUB), Brussels, Belgium, in 1990 and 1995, respectively.

From 1990 to 1999, he was a Research Engineer with the Hewlett-Packard Company. From 1999 to 2002, he became a Technical Leader with Agilent Technologies Inc., Santa Rosa, CA. In 2003, he founded the company Jan Verspecht b.v.b.a., where he held the position of Chief Consultant until 2011. In 2008, he cofounded the company Verspecht-Teyssier-DeGroote s.a.s., where he was responsible for business development until 2011. In 2012, he rejoined Agilent Technologies Inc., as a Master Research Engineer. He is a pioneer of and key contributor to the development of NVNA technology. He has authored over 30 conference papers, 14 refereed journal papers, one book chapter, and the ARFTG short course on "Large-Signal Network Analysis." He holds nine patents. He is an inventor of X-parameters. His research interests include the large-signal characterization and behavioral modeling of RF, microwave, and digital components.

Dr. Verspecht was the recipient of the 2002 ARFTG Technology Award and the 2009 Best IEEE Microwave Theory and Techniques Society (IEEE MTT-S) International Microwave Symposium (IMS) Oral Presentation Award.

Youngseo Ko, photograph and biography not available at time of publication.

Jean Pierre Teyssier, photograph and biography not available at time of publication.