A Scalable Small-Signal and Noise Model for High-Electron-Mobility Transistors Working Down to Cryogenic Temperatures

Felix Heinz[ID], Fabian Thome[ID], Dirk Schwantuschke, Arnulf Leuther[ID], and Oliver Ambacher[ID]

Abstract— This article reports on a scalable, temperature-dependent small-signal and noise model of a 50-nm metamorphic high-electron-mobility transistor (HEMT) technology. The model is valid for temperatures ranging from 5 to 297 K. The highest scalability is achieved by using a distributed model topology. The model is able to predict the small-signal and noise performance of 2–8 finger transistors with absolute gate widths ranging from 10 to 480 μm. Short gate width transistor fingers (5 μm) and wide transistor fingers (100 μm) are covered by the model. The model is valid over a very broad bandwidth from 0.1 to at least 150 GHz. Furthermore, the model covers all reasonable bias points of the given transistor technology. To the best of the authors' knowledge, this is the first scalable HEMT model that is able to predict the small-signal and noise performance at arbitrary cryogenic temperatures. The scaling ratio related to the absolute gate width of the model is the highest among small-signal models reported in the literature.

Index Terms— Cryogenics, high-electron-mobility transistors (HEMTs), low-noise amplifiers (LNAs), metamorphic HEMTs (mHEMTs), monolithic microwave integrated circuits (MMICs), noise, noise modeling, quantum computing, radio astronomy, small-signal modeling.

I. INTRODUCTION

INTEGRATED circuits in the microwave and millimeter-wave regime benefit from high absolute bandwidth that can be achieved and are used in manifold applications, such as radar, radio astronomy, quantum computing, and wireless communication. Most circuits today are operating at room temperature, but some applications, as, for example, receivers for radio astronomy and read-out circuits of quantum computers, require the cooling of the circuits to cryogenic temperatures. This reduces the noise added by the electronic components, which improves the signal-to-noise ratio of the extremely weak signals that need to be handled in such applications. The lowest noise temperatures at room temperature and under

cryogenic conditions are achieved by InGaAs high-electron-mobility transistors (HEMTs) realized either directly lattice matched on InP substrate or using a metamorphic buffer on a GaAs substrate. Therefore, the latter ones are called metamorphic HEMTs (mHEMTs).

The design of such circuits relies on accurate computer-aided design models, which describe the electrical behavior of the different components used in the circuit. Accurate models reduce costs and time effort since circuits can be designed in fewer design iterations. Furthermore, accurate models allow the fine-tuning of the circuits for optimal performance. Some parameters can only be accessed by the model since they cannot be extracted directly from a single measurement. Especially, the performance of devices at cryogenic temperatures is hard to access because the preparation of samples and cooling is time-consuming and costly.

While passive components can be modeled sufficiently precise with today's 3-D electromagnetic field solvers, the description of the active devices is still a limiting factor for simulation accuracy. Besides the highest model accuracy, high scalability is important so that the circuit designer can choose arbitrary transistor sizes. Furthermore, a wide range of bias points need to be covered by the model to allow the optimization of the circuit according to the bias point of the transistors. All properties mentioned need to be valid over a broad frequency range to allow the design of circuits in arbitrary bands and enable extrapolations beyond the measurable frequency range [1].

Classical cryogenic applications (e.g., radio astronomy) target cryogenic temperatures around 10 K, but applications at other cryogenic temperatures are upcoming. The scaling of future quantum computers to millions of quantum bits to achieve quantum supremacy requires the integration of more functionalities into the cryostat, which will run at different temperature stages. Furthermore, future remote Earth observation missions might utilize cryogenic electronics at temperatures above 77 K since the corresponding cryogenic cooling engines can be handled easier in satellite missions compared to a liquid helium cryostat.

Several small-signal and noise models of HEMTs at room temperature [1]–[15] and at cryogenic temperatures [7], [16]–[18] are reported in the literature. The dependence of the intrinsic small-signal equivalent circuit elements of a 100-nm InP HEMT process is reported in [19]; however,

no scalability data and no noise model are provided. The impact of the channel noise temperature of InP HEMTs at different temperatures is reported in [7] and [18], but no model scalability is provided.

This article deals with scalable small-signal and noise modeling of mHEMTs with 50-nm gate length at arbitrary temperatures between 5 and 297 K. The model uses the scalable small-signal equivalent circuit model topology proposed in [3] for which the temperature-dependent small-signal parameters and a temperature-dependent noise model of Fraunhofer Institute for Applied Solid State Physics's (Fraunhofer IAF) new noise optimized 50-nm mHEMT technology [20] are extracted.

This article is structured as follows. Section II introduces the 50-nm mHEMT technology that is modeled, followed by Section III where the distributed model topology is briefly recapped. Section IV deals with the temperature-dependent small-signal modeling. Section V discusses the noise modeling, and the model is verified in Section VI. Section VII compares the proposed model to the state of the art, and finally, Section VIII concludes this article.

II. 50-nm mHEMT TECHNOLOGY

This section introduces the 50-nm mHEMT technology of Fraunhofer IAF that is modeled. The processing of the technology is described in detail in [21] and the new noise optimized technology variation is reported in [20] (called Technology C).

A linearly graded InAlGaAs metamorphic buffer is grown on a 100-mm semi-insulating GaAs wafer to adapt the lattice constant from GaAs to the InP value by molecular beam epitaxy (MBE). This allows the lattice-matched growth of $In_{0.52}Al_{0.48}As$. The 2-D electron gas (2DEG) is confined in a $In_{0.8}Ga_{0.2}As$ channel between two $In_{0.52}Al_{0.48}As$ barrier layers, where the upper layer is silicon delta doped. A saturation doped cap layer finalizes the epitaxial layer sequence and provides low loss ohmic contacts. Wet chemical recess etching with a succinic acid opens the cap and electron beam lithography with a four-layer resist is used to structure the 50-nm T-gate with a Pt–Ti–Pt–Au layer sequence. The gate-line resistance is improved by placing a 450-nm-wide metal strip of the first metallization layer on top of the 170-nm gate head, which allows for ultrahigh-frequency and ultralow-noise operation. The gate is passivated with low-k benzocyclobuten (BCB), which allows for low parasitic capacitances. A chemical vapor deposited 250-nm SiN layer passivates the wafer and provides metal–insulator–metal (MIM) capacitors with 0.225-fF/μm^2 sheet capacitance.

The process supports full monolithic microwave integrated circuit (MMIC) design with MIM capacitors and temperature-independent NiCr thin-film resistors with 50-Ω/square sheet resistance. The process utilizes two metal layers for interconnections and the upper layer can be used to realize air bridges. The active devices are isolated by wet chemical mesa etching. After front-side processing, the wafer is thinned to 50-μm thickness to suppress parasitic substrate modes at high frequencies. Through-substrate via holes are dry-etched and a 2.7-μm plated gold layer forms the backside connection.
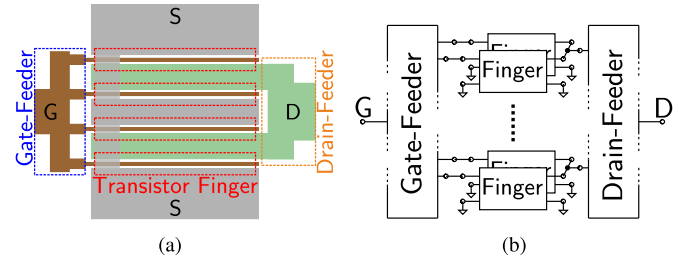

Fig. 1. (a) Simplified layout of a four-finger common-source HEMT. (b) Block diagram showing how the HEMT structure is divided into model subparts (both adapted from [3]).
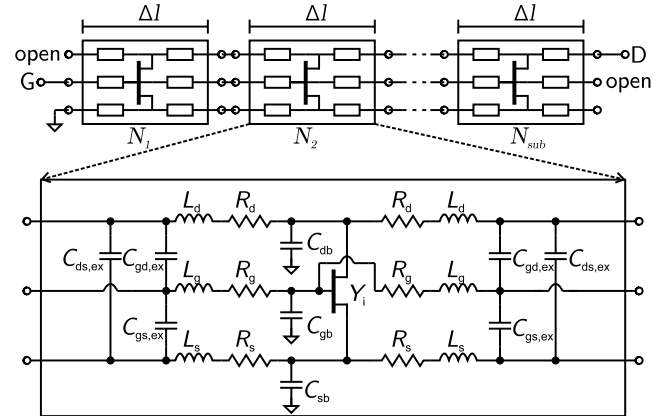

Fig. 2. Distributed transistor finger region model (adapted from [3]) consisting of a line-up of N_{sub} six-ports describing a subpart of the finger of width Δl. The equivalent circuit of a six-port subpart is shown on the bottom. All lumped elements belong to the extrinsic, parasitic transistor structure, and the transistor symbol depicts the intrinsic model (one per subpart).

III. DISTRIBUTED MODEL TOPOLOGY

The distributed model topology presented in detail in [3] is briefly summarized to provide an overview on the model topology that is used. The topology models the transistor structure according to its layout to allow for realistic implementation of parasitic effects. Fig. 1 shows the simplified layout of a four-finger common-source HEMT and the corresponding block diagram of the model depicting how the model divides the transistor into subparts according to its layout.

The model subparts describe the connections to other circuit elements (called "feeder") and the transistor finger region. The feeder networks are modeled as RLC circuits describing the lines to the transistor fingers, which are connected according to the number of fingers and the position of the feeder (gate or drain). The transistor finger area includes the region of the actual gate finger in combination with the neighboring source and drain electrodes to allow for realistic modeling of coupling effects. Each transistor finger is modeled as the line-up of ten six-ports each describing a tenth of the finger. This leads to a high distribution of the finger elements and allows to model distributed effects at longer finger structures and high frequencies with linear scaling rules according to the gate width. The number of the finger unit cells is based on the investigation of [3] and ensures that a large unit gate width can be modeled correctly up to high frequencies. Fig. 2 shows the extrinsic six-port structure containing one intrinsic model core.
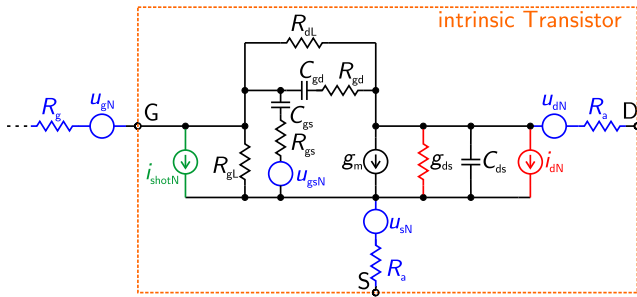
Fig. 3. Intrinsic equivalent circuit with access resistances (encapsulated by dashed line) as inserted at the position of the transistor symbol in Fig. 2. Furthermore, a subpart of the gate-line resistance is shown. Different sources of noise are depicted by noise voltage and current generators. Blue: thermal noise at T_a. Green: shot noise due to gate current. Red: channel noise.
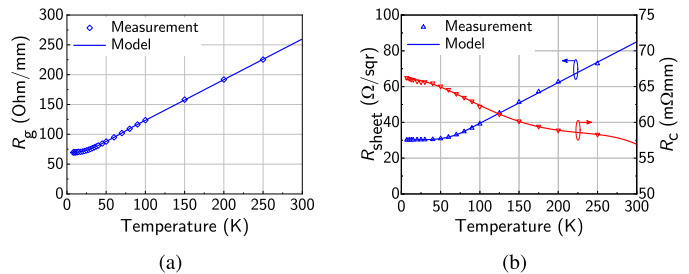

Fig. 4. Measured temperature dependence of the extrinsic resistances (symbols) and corresponding model fit (lines). (a) Gate-line resistance from a four-wire measurement of a gate-line meander structure. (b) Sheet resistance (blue triangles) and contact resistance (red flipped triangles) from four-wire TLM structure measurements are shown.

All finger elements scale linearly with the width of the corresponding finger subpart they describe. In order to provide symmetric six-ports, capacitive coupling between the electrodes, electrode resistances, and electrode inductances are included two times per six-port, which demands for halving the corresponding parameters per subpart length. The intrinsic model, which is included as a three-port into the finger six-port and describes the active transistor region, scales with the length of the corresponding finger subpart as well. The equivalent circuit is shown in Fig. 3.

IV. TEMPERATURE-DEPENDENT SMALL-SIGNAL MODELING

The temperature-dependent model parameter extraction for the given technology and model topology is described in this section. On-wafer S-parameter measurements of 50-nm mHEMTs with various gate widths form the basis for the small-signal model. Room-temperature S-parameters have been measured between 0.01 and 150 GHz with an Anritsu VectorStar vector network analyzer system. Cryogenic S-parameters are measured in a Lakeshore CRX-4K probe station between 0.01 and 50 GHz using a PNA-X vector network analyzer from Agilent Technologies. Dedicated on-wafer calibration structures allow to measure the device under test (DUT) directly at its reference planes making deembedding unnecessary. The calibration has been performed with a line-reflect-reflect-match algorithm using a zero-length THRU standard. This has the advantages that the reference planes are accurately determined by the geometrical structure of the calibration standards and the transmission line characteristics (e.g., impedance and electrical length) are not required for calibration. Furthermore, the reference impedance of the match standard can be (dc-) measured under cryogenic conditions allowing the accurate determination of the system impedance. Cryogenic S-parameters have been measured at 10, 50, 100, 150, 200, and 250 K. At each temperature stage, a new calibration needs to be done after all components have reached a static temperature level to ensure the correct determination of losses and the measurement reference plane.

In a first step, a room-temperature model is extracted, which has the advantage that S-parameter measurements up to 150 GHz can be used for extraction. This eases the accurate determination of several equivalent circuit parameters

as, for example, the inductances of the transistor electrodes since their influence on the transistor S-parameters increases with frequency. There is no evidence that parasitic inductances should change with temperature since they are mainly determined by the geometrical dimensions of the transistor electrodes that do not change significantly with cooling. The invariance of inductances against temperature changes has also been reported in [19]. This allows the reuse of the precisely extracted room-temperature inductance values for the temperature-dependent model. The extrinsic inductances per unit gate width depend on the shape of the electrode and the highest values are found for the very narrow gate electrode. The room-temperature extraction uses dc measurements of transfer lengths measurement (TLM) structures and gate-line meander test structures to determine the extrinsic resistances. Due to the distributed model topology, no fitting factors for the gate-line resistance are needed and the dc measurement values can be used directly as model parameters. The cold-FET method according to [22] is used to extract the parasitic capacitances, which are smaller than the corresponding intrinsic capacitances. Several S-parameter measurements of different transistor geometries and several active bias points are used to separate between intrinsic and extrinsic elements and to extract the parasitic inductances. The extraction procedure uses a dedicated simplex optimization algorithm to determine the equivalent circuit parameters by minimizing the squared distance between model output and measured S-parameters and is described in detail in [3].

At first, the temperature dependence of the extrinsic parameters, which model the purely passive, parasitic electrical behavior of the HEMT structure, is extracted. Since inductances of the transistor electrodes are constant over temperature, only the temperature dependence of the extrinsic resistances and capacitances needs to be extracted. The temperature dependence of the gate-line resistance is obtained from four-wire measurements of a gate-line-meander test structure at several temperatures. Fig. 4(a) shows the measured gate-line resistance values as a function of temperature and the function modeling the temperature dependence of the gate-line resistance. The gate-line resistance decreases linearly with temperature until the decrease settles below 80 K. The linear behavior originates from reduced phonon scattering due to cooling and the settlement at very low temperatures is likely to be caused by

impurity and surface scattering mechanisms, which seems to dominate the resistance at low temperatures. The temperature dependence is modeled by a linear function between room temperature and 80 K. Below 80 K, the saturation behavior is modeled by a polynomial function. Since the gate-line resistance is mainly determined by the 450-nm metal strip of the first metallization layer, the ratio of the gate-line resistance at the temperature of interest to the room-temperature value is used as a factor for the resistance of the other electrodes and feeder elements, which consists of the same metallization layer (but are of much lower resistance).

The temperature dependence of the semiconductor sheet resistance and the contact resistance, which form the access to the active intrinsic transistor, is extracted from dc measurements of a four-wire TLM test structure at several temperatures. Fig. 4(b) shows the measured sheet and contact resistances and the corresponding functions modeling their temperature dependence. The semiconductor sheet resistance decreases linearly with temperature until the decrease settles at low temperatures. This seems to originate from increased electron mobility due to lower phonon scattering and the settlement at low temperatures is likely to be dominated by impurity scattering mechanisms. The contact resistance increases slightly with decreasing temperature, which might be caused by lower electron energy that leads to lower thermionic emission probability at the semiconductor–metal interface.

A temperature-dependent factor for the extrinsic capacitances has been extracted at each temperature following the procedure of [3]. A single factor is feasible for parasitic capacitances since only a change in effective permittivity or a change of metal distances due to contraction should cause a change of parasitic capacitances. An increase of extrinsic capacitances by almost 3% at 10 K has been extracted by fitting a linear function dependent on the temperature to the extracted capacitance factors. The extracted change of the extrinsic capacitances is quite small and might be biased by the uncertainty of the determination of the internal reference planes to the intrinsic model. However, the best model agreement with the extraction data is obtained for this extracted increase.

Once the extrinsic model is fully determined, the intrinsic parameters are extracted from S-parameter measurements. S-parameter measurements at different drain currents and drain voltages at all temperatures are used for the extraction of both the temperature and bias dependence of the intrinsic parameters. The dependence of each intrinsic parameter on the drain voltage, drain current, and the temperature has been modeled by a trivariate Taylor polynomial of first to third order. A large set of S-parameter measurements has been used to extract the coefficients of the polynomial functions describing the intrinsic parameters by minimization of the squared distance of the model prediction to the measured S-parameters via a dedicated simplex optimization algorithm. The simplex algorithm is able to find a globally optimal solution of a minimization problem with an exponential time complexity. However, for a feasible set of starting values, the algorithm often only needs to explore a subset of the parameter space and is able to find a solution in a feasible time.
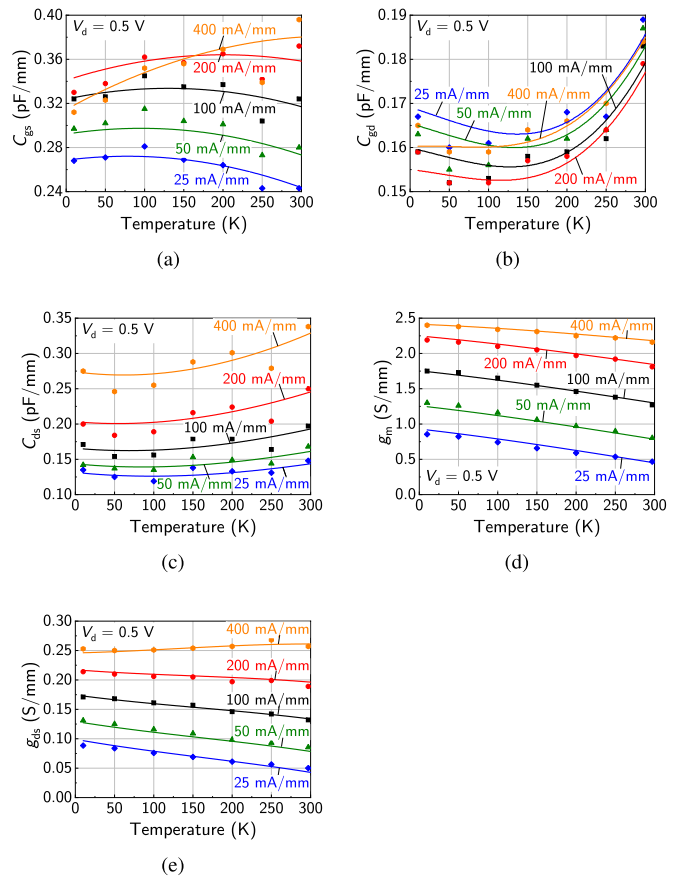

Fig. 5. Modeled intrinsic parameters (lines) and extractions from deembedded S-parameter measurements (symbols) as a function of temperature at $V_d = 0.5$ V. (a) C_{gs}. (b) C_{gd}. (c) C_{ds}. (d) g_m. (e) g_{ds}. All parameters are depicted at $I_d = 25$ mA/mm (blue diamonds), $I_d = 50$ mA/mm (green triangles), $I_d = 100$ mA/mm (black squares), $I_d = 200$ mA/mm (red circles), and $I_d = 400$ mA/mm (orange octagons).

Starting values can be obtained in a short time from a direct extraction in a single bias and temperature point (zero-order series and constant term) and can then successively be refined by simultaneously increasing the order of the series and the extraction data basis. Most intrinsic parameters use a third-order polynomial to achieve the coverage of a wide range of bias points and temperature. However, some intrinsic parameters have a polynomial of reduced order in at least one dimension (drain voltage, drain current, and temperature) to avoid nonplausible dependencies due to the additional extrema of the higher order polynomial appearing in the region of interest (data overfitting).

Fig. 5 shows the most important intrinsic model parameters at different drain currents as a function of ambient temperature T_a. The modeled functions (lines) are compared to extractions from S-parameter measurements, from which the extrinsic circuit has been deembedded (symbols) using the scheme mentioned in [23]. The extractions from the deembedded S-parameter measurements are obtained by transferring the deembedded measurements to Y-parameters and the intrinsic capacitances are obtained from the slope of the imaginary parts of the Y-parameters, whereas the real parts give the conductance values. An excellent fit of the extracted

functions with the direct extractions from the deembedded S-parameter measurements is found indicating the validity of the extracted parameters. A wide coverage of bias points and an excellent modeling of the temperature dependence of the 50-nm mHEMT technology are achieved.

The intrinsic C_{gs} shows a slight increase with temperature (less than 10%) at low drain currents but is generally quite robust against temperature changes. At drain currents above 200 mA/mm, the intrinsic gate–source capacitance decreases with temperature. A qualitatively similar dependence of C_{gs} has been found in [19] for a 100-nm InP HEMT and has been motivated by the different transport mechanisms under the gate, which might cause this change of the gate–source capacitance. Both C_{gd} and C_{ds} show a slight decrease with temperature, which is likely to originate from better carrier confinement at lower temperatures. The intrinsic g_m increases continuously when the temperature is decreased and is approximately 300 mS/mm higher at 10 K compared to the room-temperature value independent of the current bias. This fits the observed dc transfer characteristics well. g_{ds} increases at lower temperatures as well, just for high drain currents above 200 mA/mm, the trend reverses and a slight decrease of g_{ds} is observed.

Besides the continuous description of the temperature dependence of the different intrinsic parameters, their dependence on the drain current and drain voltage bias is important to allow for the highest design flexibility and to allow bias-dependent circuit design optimizations. Typically, most circuits are used most often at room temperature and second most often at deep cryogenic temperatures around 10 K. Therefore, the dependence of the most important intrinsic parameters on the drain current and drain voltage at 297 and 10 K is given in Figs. 6 and 7, respectively. Again, the modeled values (lines) are compared to direct extractions from S-parameter measurements where the extrinsic circuit has been deembedded (symbols). An excellent fit of the model with the extractions from the measurements is observed at both temperatures.

Regarding general trends, the intrinsic parameters show a qualitatively similar dependence on drain current and drain voltage at both temperatures. C_{gs} increases at a high rate with drain current, especially at low drain currents, and the increase settles at drain currents above 200 mA/mm. At both temperatures, a linear increase of C_{gs} with drain voltage is extracted. C_{gd} is quite robust against drain current changes, just at high drain voltages, a more pronounced decrease with drain current is observed. C_{gd} decreases with increasing drain voltage, which is a consequence of the higher voltage applied to the gate–drain terminals of the device. C_{ds} increases almost linearly with drain current. A slight decrease of intrinsic drain–source capacitance with increasing drain voltage is observed at high currents, but the dependence on the drain voltage is quite low, especially at lower currents. g_m increases at a high rate at low drain currents and the increase flattens out above 200 mA/mm. g_m is almost independent on the drain voltage in the bias range that is considered; however, the maximum is reached around 0.5 V and this maximum is more pronounced at cryogenic conditions. g_{ds} increases with drain current at a high rate at first, especially for low drain

Fig. 6. Modeled intrinsic parameters (lines) and extractions from deembedded S-parameter measurements (symbols) as a function of drain current (left column) and as a function of drain voltage (right column) at $T_a = 297$ K. (a) C_{gs} dependent on I_d. (b) C_{gs} dependent on V_d. (c) C_{gd} dependent on I_d. (d) C_{gd} dependent on V_d. (e) C_{ds} dependent on I_d. (f) C_{ds} dependent on V_d. (g) g_m dependent on I_d. (h) g_m dependent on V_d. (i) g_{ds} dependent on I_d. (j) g_{ds} dependent on V_d are shown.

voltages, and settles at higher drain currents. g_{ds} is almost independent of the drain voltage above 0.7 V and a slight increase toward lower drain voltages is observed, which fits the typical transistor output current–voltage characteristics.
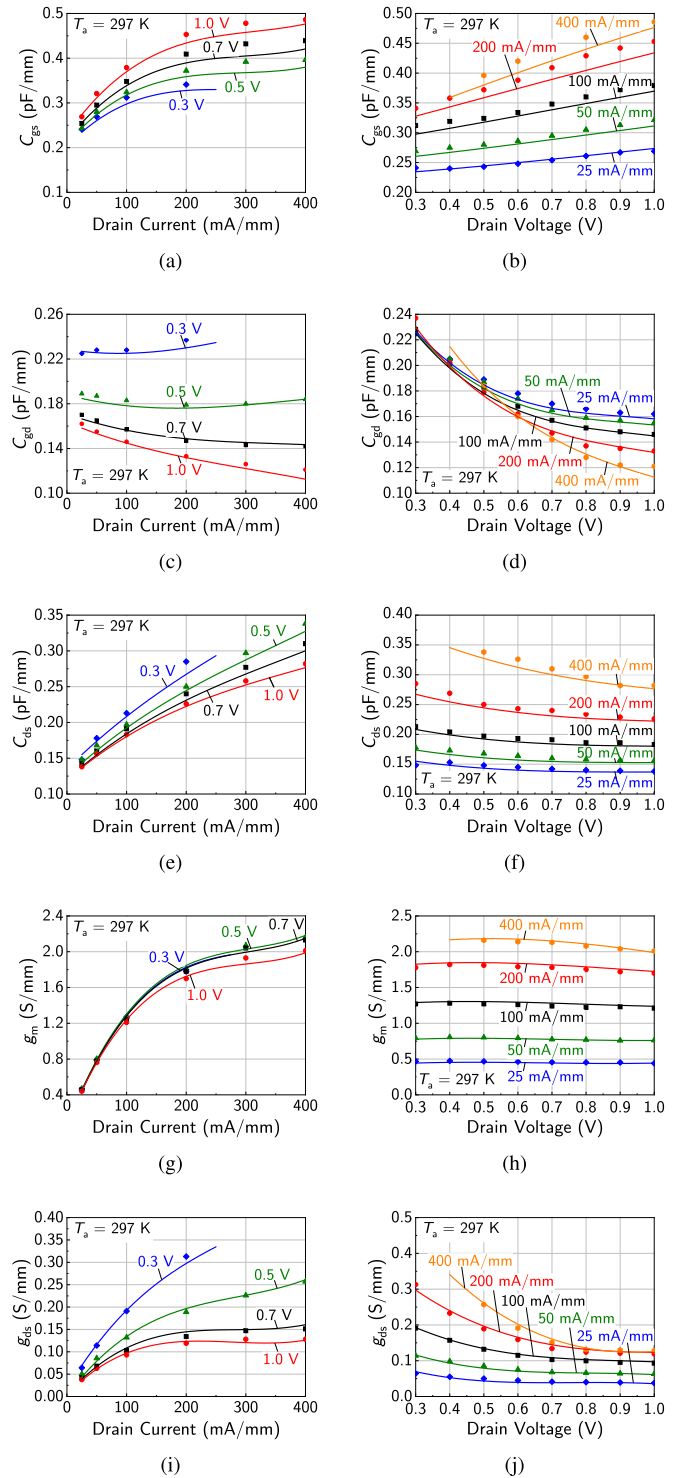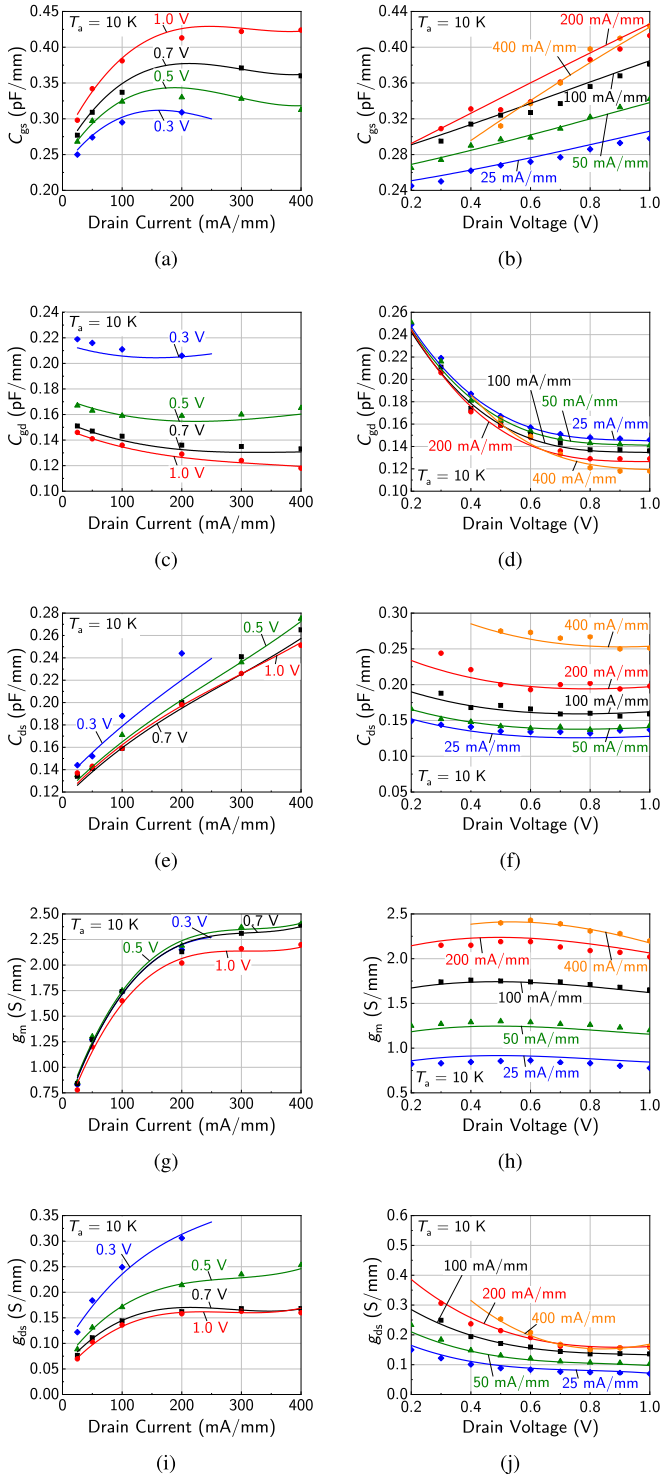
Fig. 7. Modeled intrinsic parameters (lines) and extractions from deembedded S-parameter measurements (symbols) as a function of drain current (left column) and as a function of drain voltage (right column) at $T_a = 10$ K. (a) C_{gs} dependent on I_d. (b) C_{gs} dependent on V_d. (c) C_{gd} dependent on I_d. (d) C_{gd} dependent on V_d. (e) C_{ds} dependent on I_d. (f) C_{ds} dependent on V_d. (g) g_m dependent on I_d. (h) g_m dependent on V_d. (i) g_{ds} dependent on I_d. (j) g_{ds} dependent on V_d.

The excellent agreement between the modeled intrinsic parameters and the direct extractions from S-parameter measurements allows the usage of the model at arbitrary bias points between the drain voltages of 0.2 and 1 V and the

drain currents between 25 and 400 mA/mm. This covers all bias points that are usually used under RF drive in various types of circuits. This includes low-noise-amplifier (LNA) bias points that are typically operated at low drain currents and voltages (especially under cryogenic conditions), as well as bias points used for power amplifiers that typically use higher drain voltages and currents.

V. NOISE MODELING

The proposed small-signal model is used as the basis for the inclusion of an RF-noise model. The noise model is extracted based on noise temperature measurements of different LNAs at room and cryogenic temperatures. On-wafer noise temperature and S-parameters at room temperature have been measured in one-probe contact with a Keysight PNA-X vector network analyzer system. The system includes a sensitive receiver dedicated to noise temperature measurements and an input tuner that allows to measure the DUT's noise parameters and provides vector-corrected noise-figure measurements. Noise temperature measurements have been done using this vector corrected cold-source method at room temperature [24]. For the cryogenic noise temperature measurements, the Y-factor method with a cold attenuator was used [25]. A 20-dB attenuator MMIC with integrated temperature sensor is glued in front of the DUT and the connection to the DUT is formed with short wire bonds between the attenuator output and the DUT input. The input probe arm connects the attenuator chip input and the other probe arm the DUT output. This allows for cryogenic noise temperature measurements directly at the chip reference plane, which enables precise noise model extractions. The noise powers have been measured with an Agilent N8975A noise-figure analyzer (NFA) and a custom preamplifier. The cold and hot states for the Y-factor method are provided by a Keysight 346C noise diode. The overall worst case measurement uncertainty in the Ku-band has been estimated to be ± 1.4 K accounting for the diode excess noise ratio (ENR) uncertainty, the uncertainty of the attenuator temperature, the uncertainty of the probe arm noise temperature, the NFA measurement uncertainty, and the mismatch of the components in the measurement chain.

The extrinsic circuit is purely passive, and therefore, only white, thermal noise at the ambient temperature at which the model shall be used for simulations is generated in the resistive elements of the extrinsic equivalent circuit. The main contributors of the extrinsic equivalent circuit to the overall effective noise temperature are the gate-line resistance and the access resistances. The latter ones are mainly determined by the semiconductor sheet resistance, the ungated recess area, and the contact resistance. The noise temperature of the 50-nm mHEMT technology reduces significantly upon cooling to 10 K. A huge impact on this reduction of noise temperature originates from the reduced thermal noise of the parasitic resistances. Especially, the gate-line resistance is a significant contributor to the overall noise temperature due to its comparably large resistance, which is a consequence of the small cross section of the short gate. The noise of gate-line resistance and sheet resistance drops significantly when cooled

to 10 K since the temperature is lowered by almost factor 30 compared to room temperature and the resistance value drops as well according to Fig. 4. The contact resistance rises slightly for the given technology, but the noise power added by the contact resistance is still reduced significantly due to the enormous change in temperature. Fig. 3 shows the equivalent circuit of the intrinsic model with all noise generators connected (thermal noise depicted as voltage sources in blue).

The active, intrinsic transistor noise can be modeled by connecting noise generators to the corresponding equivalent circuit elements [26]–[28]. A widely used noise modeling approach for HEMTs both at room temperature and under cryogenic conditions has been presented by Pospieszalski [7] and is adapted in this work for the intrinsic model of the proposed topology. Pospieszalski's model introduces two sources of noise to describe the noise of the transistor. A noise source is related to the intrinsic gate–source resistance, which is treated to cause thermal noise at ambient temperature. The second source of noise is related to g_{ds} and is again treated to be thermal, but at much higher temperatures than the lattice temperature. This high temperature is caused by the high electrical field that is placed across the short gate, which accelerates the channel electrons and puts them to a higher energy level. This is interpreted as a high temperature of the 2DEG. Thus, the noise power is not a strong function of the lattice temperature in Pospieszalski's approach.

Pospieszalski's model is extended to have another source of noise related to the gate, which models pure shot noise caused by gate-leakage current flowing through the Schottky barrier forming the gate (see Fig. 3). Equation (1) states the effective noise current of the shot noise source with q being the elementary electron charge, $|I_g|$ being the magnitude of the gate-leakage current, and Δf being the absolute bandwidth that is considered

$$\left| \bar{i}_{shotN} \right| = \sqrt{2 \cdot q \cdot |I_g| \cdot \Delta f}. \tag{1}$$

From (1), it can be seen that the shot noise power is fully determined by the magnitude of the gate-leakage current. The gate-leakage current is a function of the bias point and the temperature. The dependence of the gate-leakage current on the drain voltage, the drain current, and the temperature is extracted from dc measurements at various bias points at temperatures of 10, 50, 100, 150, 200, 250, and 297 K and is modeled by a trivariate Taylor series. At typical low-noise bias points, the influence of the gate leakage on the noise temperature is small, but at some gate-diode forward bias points, it increases so that it has a higher impact on the overall noise temperature.

The intrinsic cores of the model also utilize the two sources of the Pospieszalski model. R_{gs} is treated to generate thermal noise at ambient temperature directly without a fitting factor. However, for the given 50-nm technology, R_{gs} is already at room temperature very low and has only a small influence on the overall noise temperature. At cryogenic temperatures, noise caused by R_{gs}, and consequently, the influence on the overall noise temperature of the device decreases even further.

Once all other sources of noise are determined by physical means, the channel noise temperature (T_d) modeling the
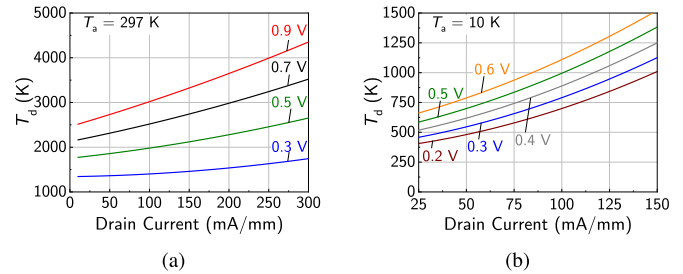

Fig. 8. Modeled channel noise temperature as a function of drain current at (a) 297 and (b) 10 K.

g_{ds}-related noise is extracted (see Fig. 3). The extraction is based on the room temperature and cryogenic noise temperature measurements of different LNAs from lower gigahertz regime up to W-band (75–110 GHz) at various bias points. The corresponding circuit simulations are set up with the proposed HEMT model in combination with a well-established passive grounded coplanar waveguide (GCPW) library [29] in which the circuits are built. The passive library's attenuation is adapted for the cryogenic circuit simulation to meet the reduced losses of the matching networks, which is especially important for the input matching network. The cryogenic loss of the GCPW is extracted from S-parameter measurements of a resonator structure built of the GCPW. The channel noise temperature T_d is adapted to minimize the root-mean-square distance between noise temperature measurement and circuit simulation at various bias points. The resulting dependence of T_d at 10 and 297 K on the drain current and the drain voltage is modeled by two bivariate Taylor series. Fig. 8 shows the modeled T_d at 297 K (left) and at 10 K (right) as a function of drain current at various drain voltages (note the different scales and voltages).

The extracted T_d at 10 K is approximately half the room-temperature value for a comparable drain voltage and current. The temperature dependence of T_d is extracted from noise temperature measurements of the single-ended W-band LNA presented in [30] at different cryogenic temperatures. Fig. 9 shows T_d (discrete extractions and model), the intrinsic g_{ds} to which T_d corresponds, and the resulting effective channel noise current density as a function of ambient temperature.

A linear temperature dependence of T_d has been extracted for temperatures above 50 K. Below 50 K, the model fit is refined by modeling the observed settling behavior by a quadratic function. A simple linear model of the temperature dependence of T_d would also provide feasible results, but with slightly higher deviation between model and measurements. The modeled temperature dependence is used for the calculation of T_d at arbitrary temperatures and bias points by using the function to interpolate between cryogenic and room-temperature values.

An interesting finding is that the actually modeled magnitude of the effective channel noise current density (i_{dN}) reduces only by approximately 25% although T_d (and therefore, the corresponding noise power) is approximately halved when the device is cooled to cryogenic temperatures. This is due to the increase of g_{ds} with decreasing temperature
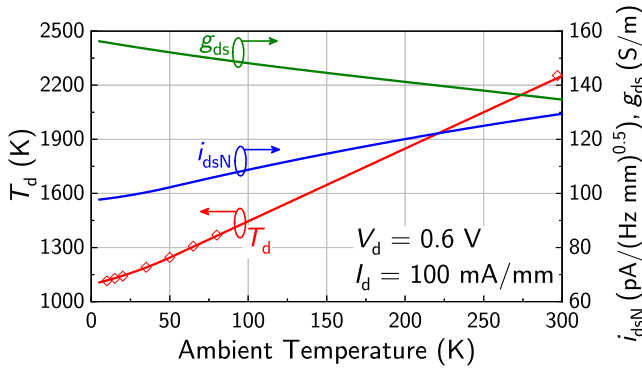
Fig. 9.   Modeled channel noise temperature (red line), modeled g_{ds} (green line), and resulting magnitude of the effective channel noise current density (blue line) as a function of ambient temperature when the device is biased at $V_d = 0.6$ V and $I_d = 100$ mA/mm. Symbols depict T_d extractions from measurements.

Fig. 10.   Simulated T_{min} at 6 GHz of a two-finger common-source 50-nm mHEMT operated at $V_d = 0.5$ V and $I_d = 50$ mA/mm as a function of the finger width (5–100 μm) and the temperature (5–300 K).

(see Figs. 5 and 9), which is the conductance in which the noise is generated in the model. As a consequence, the correctness of the extraction of the g_{ds} temperature dependence affects the extraction of the T_d temperature dependence (and therefore, the output-related noise power), which is frequently used for device comparisons. The extraction of T_d might be easily inaccurate if the dependence of g_{ds} on temperature or bias is not considered. In this work, an extensive study of the dependence of the various small-signal equivalent circuit parameters has been given. Furthermore, the noise model extraction focuses on the determination of various parameters that can be accessed by physical means before extracting T_d. This combination allows for precise extraction of T_d.

The proposed model is in good agreement with extracted temperature dependencies of T_d reported in the literature. A significant reduction of T_d with T_a in HEMTs has been found by several other groups [7], [16], [18] as well. A qualitatively comparable settling behavior of T_d at low temperatures has also been reported in [18] and [31]. The findings of this article indicate that the actual fluctuations causing the noise current are only weakly influenced by a change of ambient temperature. This is based on Pospieszalski's assumption that the small-signal g_{ds} directly represents the inverse of the physical source–drain resistance of the intrinsic device [7]. If this assumption is correct, only a reduction in noise current of approximately 25% is caused by the change in ambient temperature. These findings let the questions arise, in which physical processes cause the observed temperature dependence of the effective noise current, which has not been answered conclusively to date. However, the weak temperature dependence of i_{dN} supports the hypothesis [7] that T_d models channel noise caused by a high electrical field under the gate.

Recently, a new theory on the channel noise has been proposed, which states that the channel noise is suppressed shot noise [32]. The suppression factor (Fano factor) should be only weakly dependent on drain current and ambient temperature. However, this contradicts the channel noise extractions presented in the original work of Pospieszalski [7], the ones of other cryogenic noise models [16], [18], and the extractions of this work. A temperature-independent suppression factor
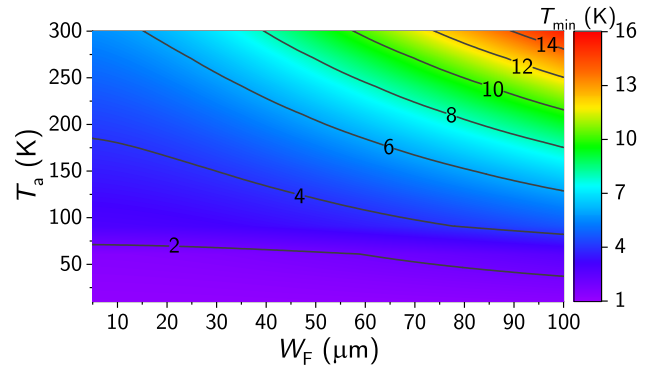
cannot qualitatively reproduce the reported change in channel noise power upon cooling. A Fano factor ranging from 0.52 to 0.29 would be able to reproduce the data (i_{dN}^2) provided in Fig. 9. Given that the theory of temperature-independent suppression of shot noise holds true, other mechanisms would have to cause the temperature dependence of the channel noise.

The noise temperature that can be obtained for a specific device at a certain temperature is highly relevant for circuit design. Fig. 10 shows T_{min} of a two-finger common-source device biased at $V_d = 0.5$ V and $I_d = 50$ mA/mm at 6 GHz as a function of temperature and finger width (W_F). The lowest noise temperatures are obtained for short unit gate width devices, but a larger impedance transformation is typically needed at 6 GHz to achieve proper noise matching. This is due to Γ_{opt} that is turning close to the open-circuit impedances for short gate width devices. The losses of a large matching network will usually degrade the achievable noise performance for such sizes at 6 GHz. Furthermore, HEMTs with a higher W_F exhibit a lower effective noise resistance R_n, which eases matching further. Therefore, considerably large gate width transistors are typically used at 6 GHz although a lower T_{min} is found for smaller devices. At higher frequencies, lower absolute gate width devices become more feasible since a lower impedance transformation is needed for noise matching at the same device size. The cryogenic bias point shown in Fig. 10 might not be optimal for all temperature levels.

Fig. 11 shows the modeled noise parameters at the example of a $W_g = 2 \times 60$-μm HEMT biased at $V_d = 0.5$ V and $I_d = 50$ mA/mm operated at 10 K.

VI. MODEL VERIFICATION

The small-signal prediction of the proposed model is verified upon on-wafer S-parameter measurements of 50-nm mHEMTs of various absolute gate widths (W_g), unit gate widths, and finger numbers (n_F). The model prediction of the S-parameters of the measured devices is evaluated and compared to the measurements. Fig. 12 shows the on-wafer S-parameter measurements compared to the corresponding model prediction at 10 and 297 K. Room-temperature S-parameters are shown between 0.1 and 150 GHz, while the measurements at 10 K cover the frequency regime between
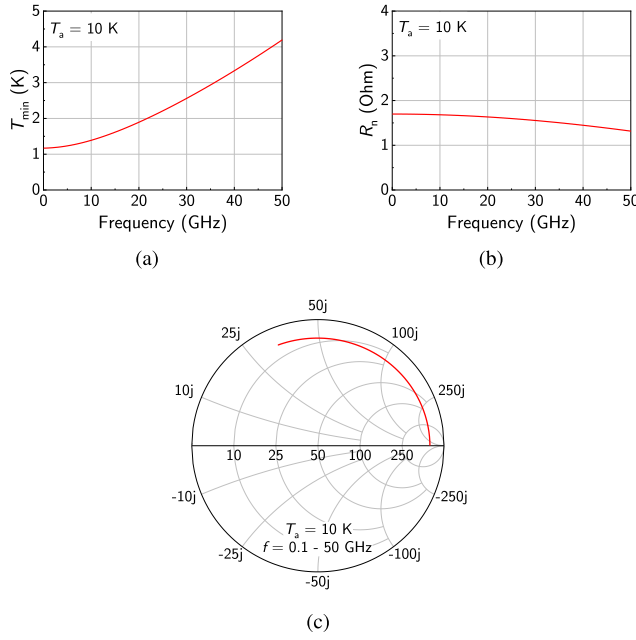
Fig. 11. Modeled noise parameters at 10 K of $W_g = 2 \times 60$ μm 50-nm mHEMT biased at $V_d = 0.5$ V and $I_d = 50$ mA/mm. (a) T_{min}. (b) R_n. (c) Γ_{opt} between 0.1 and 50 GHz.

0.1 and 50 GHz. Devices with extremely short absolute gate width ($W_g = 10$ μm) and extremely high absolute gate width ($W_g = 480$ μm) are presented. Very short single finger width ($W_F = 5$ μm) devices as well as very long finger width ($W_F = 100$ μm) devices are correctly described. Furthermore, all finger numbers supported by the HEMT process ranging from two-finger devices to eight-finger devices are shown.

An excellent agreement of the model prediction with the measurements is found for all device geometries at both temperatures presented. Since the small-signal performance of all the extreme device geometries is correctly reproduced, an extremely high scalability is achieved by the model. The verification at two temperatures at the borders of the covered temperature range of the model, which, furthermore, are most often targeted by designs, ensures that the scalability is achieved at arbitrary temperatures between 5 and 297 K. At room temperature, the measurements in the frequency regime between 0.1 and 150 GHz are correctly modeled, which demonstrates the ability of the model to cover high frequencies. Cryogenic S-parameters have only been measured up to 50 GHz, but since the model has demonstrated the ability to correctly describe the small-signal performance up to at least 150 GHz at room temperature, a similar high coverage of frequencies is to be expected at all temperatures in which the model is valid due to the same model topology.

The ability to describe the small-signal behavior of the 50-nm mHEMTs at arbitrary temperatures between 5 and 297 K is demonstrated at S-parameter measurements of a single mHEMT at temperatures between 10 and 297 K. Fig. 13 shows the S-parameter measurements of a 2×60-μm mHEMT biased at $V_d = 0.5$ V and $I_d = 50$ mA/mm at various temperatures compared with the corresponding model.
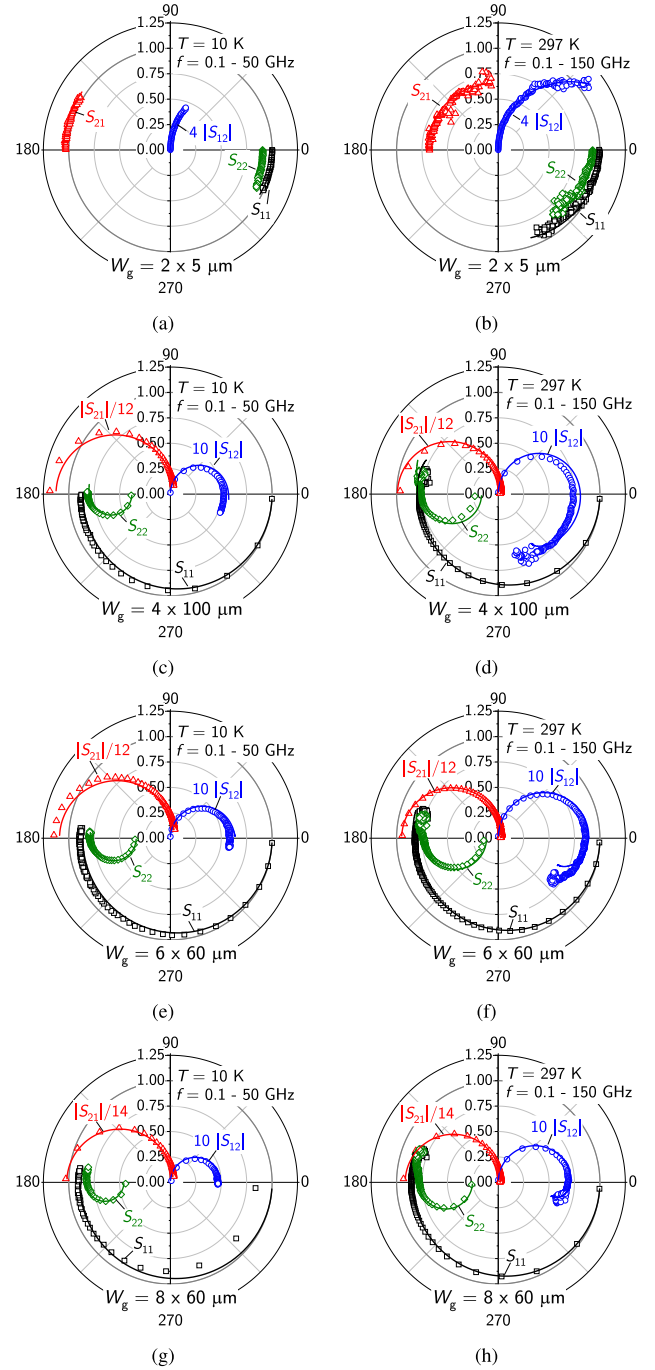

Fig. 12. On-chip S-parameter measurements (symbols) and corresponding model predictions (lines) of 50-nm mHEMTs with various absolute gate widths at an ambient temperature of (left) 10 and (right) 297 K. HEMTs with (a) $W_g = 2 \times 5$ μm at 10 K, (b) $W_g = 2 \times 5$ μm at 297 K, (c) $W_g = 4 \times 100$ μm at 10 K, (d) $W_g = 4 \times 100$ μm at 297 K, (e) $W_g = 6 \times 60$ μm at 10 K, (f) $W_g = 6 \times 60$ μm at 297 K, (g) $W_g = 8 \times 60$ μm at 10 K, and (h) $W_g = 8 \times 60$ μm at 10 K. S-parameters are depicted as follows: S_{11} (black squares), S_{12} (blue circles), S_{21} (red triangles), and S_{22} (green diamonds). The model deviation of S_{11} in (g) is due to increased leakage current. Although representative DUTs are selected by room-temperature wafer mappings, it can happen that a device is damaged during cooling.

Furthermore, the maximum stable gain (MSG) calculated from S-parameter measurements and the model prediction at 10, 100, 200, and 297 K is shown. At each temperature,
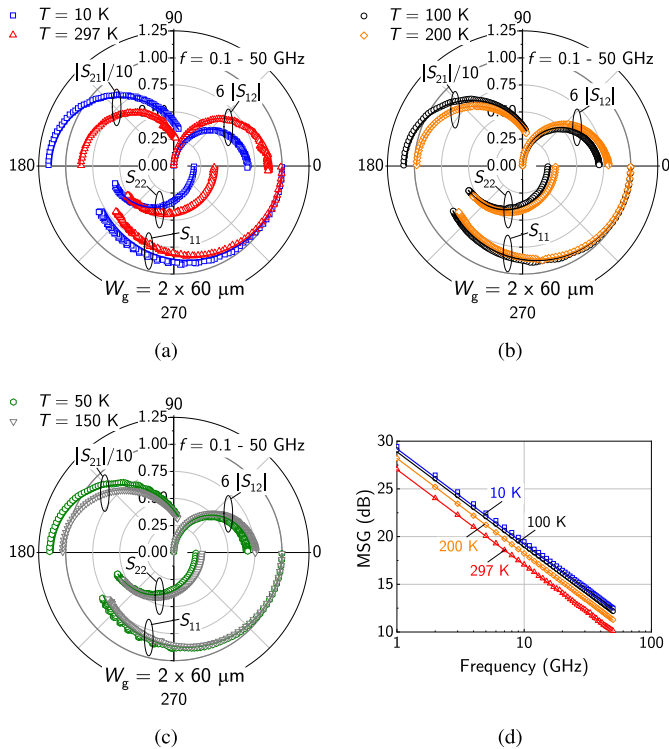
Fig. 13. On-chip S-parameter measurements (symbols) and corresponding model predictions (lines) of 50-nm mHEMTs with $W_g = 2 \times 60$ μm at different temperatures biased at $V_d = 0.5$ V and $I_d = 50$ mA/mm. Measurements at (a) 10 K (blue squares) and 297 K (red triangles), (b) 100 K (black circles) and 200 K (orange diamonds), and (c) 50 K (green octagons) and 150 K (gray flipped triangles). (d) Maximum gain at 10, 100, 200, and 297 K.

Fig. 14. On-wafer S-parameter and noise temperature measurements (symbols) from [20] of a three-stage Ku-band LNA MMIC compared to the circuit simulation utilizing the proposed mHEMT model (solid lines) at room temperature.

Fig. 15. Cryogenic gain and noise temperature measurement (symbols) from [20] of a three-stage Ku-band LNA MMIC compared to the circuit simulation utilizing the proposed mHEMT model (solid lines).

an excellent agreement between model and measurement is observed indicating the validity of the extracted temperature dependencies.

Besides a verification on device level, a verification at circuit level ensures that the model is suitable to be used in computer-aided circuit simulations. Furthermore, a circuit level verification allows to confirm the noise prediction of the model. Fig. 14 compares the room-temperature on-wafer measurement of a three-stage Ku-band LNA published in [20] with the circuit simulation utilizing the proposed model. The MMIC consists of three 4×40-μm common-source stages with symmetrical inductive source degeneration. Both S-parameters and noise temperature measurements are described correctly by the model. Resonances of input and output match are predicted precisely, as well as the magnitude and bandwidth of the small-signal gain. The modeled noise temperature is in very good agreement with the measurement.

The 10 K gain and noise temperature measurements of the same three-stage Ku-band LNA published in [20] are used to verify the model response at 10 K. Fig. 15 compares the on-chip measurements published in [20] with the circuit simulation using the proposed transistor model. An excellent agreement between measurements and simulation is achieved for both gain and noise temperature, demonstrating the capability of the proposed model for usage in the state-of-the-art cryogenic circuit design.
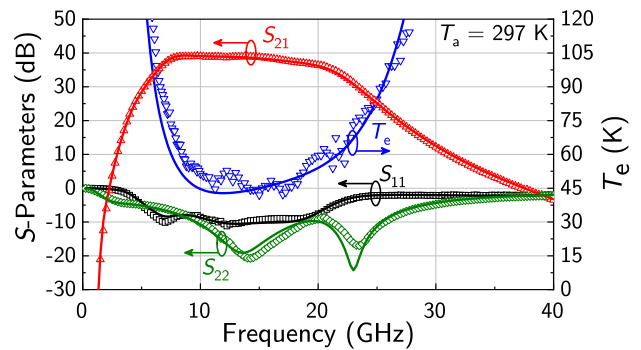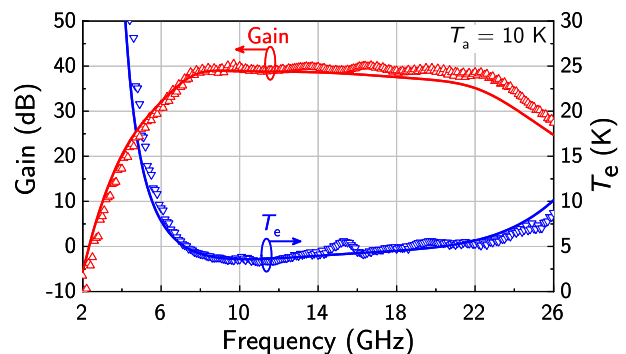
VII. STATE-OF-THE-ART COMPARISON

The proposed temperature-dependent small-signal and noise model is compared to the state of the art. A quantitative measure for the scalability of the model is needed to benchmark the scaling performance of the models compared. A simple, but quite effective measure for scalability is the ratio of the largest absolute gate width to the lowest absolute gate width transistor that is correctly described by the model. Another useful measure is the ratio of the highest gate width finger to the shortest gate width finger that can be predicted. The finger numbers that can be predicted are of course another useful measure; however, one has to consider that not all technologies allow for arbitrary finger numbers. The combination of the three mentioned parameters is useful for a first quantitative insight into the model properties. However, no statement about the frequency regime that is correctly predicted by the model can be made from this alone. This limits the informative value of the three proposed parameters on their own since it is easier to achieve a wide scaling ratio when only low frequencies are considered. Especially, when only the lower gigahertz regime is covered by a model, it is significantly easier to obtain high scalability with a comparably simple model since many parasitic effects are not as pronounced in the low-frequency S-parameters. Therefore, it makes sense to additionally account for the absolute bandwidth in which the

TABLE I
STATE-OF-THE-ART SCALABLE ROOM-TEMPERATURE SMALL-SIGNAL MODELS

Ref.	Freq. (GHz)	W_g (min. − max.) (µm)	W_F (min. − max.) (µm)	n_F (min. − max.) (a.u.)	$\frac{W_{g,max}}{W_{g,min}}$ (a.u.)	$\frac{W_{F,max}}{W_{F,min}}$ (a.u.)	FOM (GHz)
[8]	0.5 − 50	48 − 240	24 − 48	2 − 7	5	2	248
[8]	0.5 − 75	40 − 120	20 − 60	2	3	3	224
[9]	0.5 − 20	1000 − 4000	125 − 250	8, 16	4	2	78
[11]	0.05 − 40	64 − 96	2 − 18	4, 8, 32	1.5	9	60
[2]	0.05 − 450	10 − 120	5 − 30	2 − 4	12	6	5399
[4]	0.05 − 330	20 − 120	5 − 30	2 − 8	6	6	1980
[6]	0.25 − 110	60 − 200	15 − 100	2 − 8	3.33	6.66	366
[12]	1 − 15	120, 3600	60, 300	2, 12	30	5	420
[1]	10 − 100	180 − 200	50 − 60	4 − 6	1.11	1.2	100
[1]	2 − 48	100 − 1920	25 − 240	4 − 8	19.2	9.6	883
[1]	10 − 100	300 − 600	30 − 75	6 − 10	2	2.5	180
This Work	0.1 − 150	10 − 480	5 − 100	2 − 8	48	20	7200

TABLE II
STATE-OF-THE-ART TEMPERATURE-DEPENDENT SMALL-SIGNAL MODELS

Ref.	Freq. (GHz)	W_g (min. − max.) (µm)	W_F (min. − max.) (µm)	n_F (min. − max.) (a.u.)	$\frac{W_{g,max}}{W_{g,min}}$ (a.u.)	FOM (GHz)	T_a (K)	Noise
[19]	0.2 − 50	100	50	2	1	50	5 − 350	no
[18]	0.2 − 50	80	-	-	1	50	18 − 300	yes
[16]	0.2 − 67	200	100	2	1	67	10, 300	yes
[17]	0.2 − 50	30 − 240	15 − 60	2, 4	8	200	15, 297	yes
This Work	0.1 − 50	10 − 480	5 − 100	2 − 8	48	2400	5 − 297	yes

model is valid. A figure-of-merit (FOM) that accounts for scalability and frequency range coverage in a single number is proposed in (2), with f_{max} and f_{min} being the maximal and minimal frequency that is accurately predicted by the model, respectively, and $W_{g,max}$ and $W_{g,min}$ being the maximal and minimal absolute gate width that is described by the model, respectively

$$\text{FOM} = (f_{max} - f_{min}) \cdot \frac{W_{g,max}}{W_{g,min}}. \tag{2}$$

Table I compares the proposed model to state-of-the-art scalable small-signal models at room temperature based on the metric discussed. To the best of the authors' knowledge, the highest scaling ratio among room-temperature small-signal models in the literature is achieved. Transistors with very short finger width (5 µm) and very long finger width (100 µm) are accurately described by the model. Furthermore, a very wide frequency range is covered by the model, which leads to a very good FOM. Some model topologies [2], [4] have been verified to work at even higher frequencies but do not achieve the scaling ratio of the model presented in this work.

State-of-the-art cryogenic and temperature-dependent small-signal models are compared in Table II. Only a few models in the literature demonstrated scalability at cryogenic temperatures.

VIII. CONCLUSION

In this article, a scalable small-signal and noise model of a 50-nm mHEMT technology optimized for cryogenic

ultralow-noise amplification has been presented. The model is able to accurately describe the small-signal and noise performance at arbitrary temperatures between 5 and 297 K. The scalability of the model is preserved at all temperatures, which allows to design circuits optimized for the usage at dedicated temperature levels.

The outstanding scalability of the model is achieved by a distributed model topology, which allows to model parasitic effects closely related to the actual transistor layout. This allows for linear scaling rules of extrinsic and intrinsic parameters, which are valid from very short up to very long finger width. To the best of the authors' knowledge, the proposed model demonstrates the highest scalability among models presented in the literature. The ability to describe HEMTs at arbitrary temperatures between 5 and 297 K is achieved by temperature-dependent modeling of the different equivalent circuit elements.

A wide range of bias points is accurately described by the model at all temperatures considered. The dependence of the intrinsic parameters on drain current, drain voltage, and ambient temperature has been investigated on the basis of S-parameter measurements, from which the extrinsic circuit has been deembedded. The dependence of each intrinsic parameter has been modeled by a Taylor series of different order leading to the wide bias range of the model from drain currents $I_d = 25$ mA/mm to $I_d = 400$ mA/mm and drain voltages from $V_d = 0.2$ V to $V_d = 1$ V. The given bias range is validly modeled at arbitrary temperatures between 5 and 297 K.

At room temperature, a broad frequency range up to at least 150 GHz has been verified to be covered by the model. At cryogenic temperatures, S-parameters could only be measured up to 50 GHz. However, it is likely that the model achieves the same bandwidth at cryogenic temperatures as at room temperature. This is due to the fact that the model topology demonstrated its ability to describe such a bandwidth at room temperature and just the model parameters have been adapted according to the temperature.

The temperature-dependent noise model, which adapts Pospieszalski's approach, allows to predict the noise performance of transistors and amplifiers at arbitrary temperature. The extraction relies on physical parameters to model different sources of noise, as, for example, thermal noise of parasitics and shot noise caused by leakage. Once these sources have been extracted, the effective drain temperature T_d has been determined. An excellent agreement between the circuit simulation utilizing the proposed transistor model and measurements of a Ku-band amplifier indicates the validity of the noise modeling approach.

To the best of the authors' knowledge, this is the first time that a scalable small-signal and noise model is able to describe HEMTs at arbitrary temperatures between 5 and 297 K. This allows for high-performance circuit design dedicated to applications at different temperature levels, for example, satellite remote sensing missions, quantum computing, and radio astronomy.

APPENDIX

A brief overview on the arithmetics of the proposed model is given in the following. The well-known cascadeable representation of a two-port (A-matrix) can be extended to higher port (even) numbers when half of the ports is assigned to be the input and the other half to be the output. Equation (3) gives the definition of an $n \times n$ A-matrix with $n = 2\,m$. The currents at the output ports ($m + 1$ to n) are defined into the opposite direction compared to Y- or Z-matrices to allow for the cascade connection by matrix multiplication

$$\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \\ i_1 \\ i_2 \\ \vdots \\ i_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \\ a_{(m+1)1} & a_{(m+1)2} & \cdots & a_{(m+1)n} \\ a_{(m+2)1} & a_{(m+2)2} & \cdots & a_{(m+2)n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} u_{(m+1)} \\ u_{(m+2)} \\ \vdots \\ u_n \\ i_{(m+1)} \\ i_{(m+2)} \\ \vdots \\ i_n \end{pmatrix}. \quad (3)$$

With the given A-matrix definition, multiports describing subparts of coupled lines or in this case transistor electrodes can be defined. An n-port admittance matrix describing a mesh network can be extended to a $2n$-port in the A-matrix representation, where a second port is introduced at each original port. Every new port is shorted to the node of the original port. The A-matrix describing such a $2n$-port is given in (4), with the $n \times n$ submatrices $[I]$ (identity matrix),

[0] (zero matrix), and $[Y]$ being the original n-port Y-matrix

$$A_{2n \times 2n, Y} = \begin{pmatrix} [I] & [0] \\ [Y] & [I] \end{pmatrix}. \quad (4)$$

In the special case of $n = 1$, the matrix becomes the well-known two-port A-matrix of a shunted admittance [33]. This representation can be used to describe the electrode coupling by the extrinsic capacitances $C_{gs,ex}$, $C_{gd,ex}$, and $C_{ds,ex}$, as well as the capacitances C_{gb}, C_{db}, and C_{sb} and the intrinsic model by extending the corresponding three-port Y-parameter matrices to six-port A-matrices.

A similar scheme to extend a floating star n-port network in the Z-matrix representation to a $2n$-port network exists. The star point is disconnected and a new second port for each original port is connected in place of the star point to form the resulting $2n$-port network. The A-matrix describing such a $2n$-port is given in (5), with $[Z_T]$ being the original floating star n-port network. Note that this scheme is only applicable to floating star networks in the presented form since the star point is cut open

$$A_{2n \times 2n, Z} = \begin{pmatrix} [I] & [Z_T] \\ [0] & [I] \end{pmatrix}. \quad (5)$$

In the special case of $n = 1$, the matrix becomes the well-known two-port A-matrix of a series impedance [33]. This representation can be used to describe the electrode resistances R_g, R_d, and R_s and inductances L_g, L_d, and L_s.

The overall A-matrix of a transistor finger (A_F) is obtained by multiplication of the finger unit cell A-matrices and these are obtained by multiplication of the coupled electrode six-port A-matrices according to the following equation:

$$A_F = \left(A_{C,ex} \cdot A_{RL} \cdot A_{Cb} \cdot A_{intr} \cdot A_{RL} \cdot A_{C,ex} \right)^{N_{sub}}. \quad (6)$$

$A_{C,ex}$ describes the coupling of $C_{gs,ex}$, $C_{gd,ex}$, and $C_{ds,ex}$. A_{RL} models R_g, R_d, and R_s and L_g, L_d, and L_s. A_{Cb} describes C_{gb}, C_{db}, and C_{sb}. A_{intr} is the intrinsic model and N_{sub} is the number of the finger unit cells.

The Y-parameter matrix is the candidate of choice for the parallelization of multiple fingers. The two-port conversions published in [34] and corrected in [35] can be extended to the presented multiport case. Equations (7)–(10) give the conversion of the $2n \times 2n$ A-matrix to a $2n \times 2n$ Y-matrix with the $n \times n$ submatrices $[A_{11}]$, $[A_{12}]$, $[A_{21}]$, and $[A_{22}]$ and $[Y_{11}]$, $[Y_{12}]$, $[Y_{21}]$, and $[Y_{22}]$, respectively

$$[Y_{11}] = [A_{22}] \cdot [A_{12}]^{-1} \quad (7)$$

$$[Y_{12}] = [A_{21}] - [A_{22}] \cdot [A_{12}]^{-1} \cdot [A_{11}] \quad (8)$$

$$[Y_{12}] = -[A_{12}]^{-1} \quad (9)$$

$$[Y_{22}] = [A_{12}]^{-1} \cdot [A_{11}]. \quad (10)$$

The resulting six-port Y-parameter matrix is reduced to a two-port by terminating unconnected ports with an open and by shorting ground ports. A general scheme to reduce an n-port network in the Y representation to an $n-1$-port network by terminating port number t with an admittance y_L is given in the following.

Let the original Y-matrix be a quadratic $(m+1) \times (n+1)$ matrix, where the port number t will be terminated with an admittance y_L. The row and column at position t are both discarded leading to a preliminary $m \times n$ Y-matrix. The entries $y_{f,mn}$ of the final reduced $m \times n$ matrix are calculated from the entries of the preliminary matrix $y_{p,mn}$ and the original matrix $y_{o,mn}$ at row m and column n according to the following equation:

$$y_{f,mn} = y_{p,mn} - \frac{y_{o,mt} \cdot y_{o,tn}}{y_{o,tt} + y_L}. \tag{11}$$

In the case of shorting port t (y_L becomes infinite), the preliminary matrix is the final matrix directly. This scheme is used to terminate electrode ports with dedicated admittances modeling an RF open or RF short. Iteratively applying this scheme to the six-port Y-parameter matrix allows to reduce it to a two-port Y-matrix with one port being a gate connection and the other port being a drain connection. From this point on, the setup of the model is straightforward and can be done by well-known two-port arithmetics [33]–[35] according to Fig. 1(b) and [3].

ACKNOWLEDGMENT

REFERENCES

[1] J. Wood and D. E. Root, "Bias-dependent linear scalable millimeter-wave FET model," *IEEE Trans. Microw. Theory Techn.*, vol. 48, no. 12, pp. 2352–2360, Dec. 2000.

[2] M. Ohlrogge *et al.*, "A scalable compact small-signal mHEMT model accounting for distributed effects in sub-millimeter wave and terahertz applications," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2014, pp. 1–4.

[3] F. Heinz, D. Schwantuschke, A. Leuther, and O. Ambacher, "Highly scalable distributed high electron mobility transistor model," in *Proc. IEEE Asia–Pacific Microw. Conf.*, Dec. 2019, pp. 1286–1288.

[4] M. Ohlrogge *et al.*, "Small signal modelling approach for submillimeter wave III–V HEMTs with analysation and optimization possibilities," in *IEEE MTT-S Int. Microw. Symp. Dig.*, May 2016, pp. 1–4.

[5] F. Heinz, F. Thome, A. Leuther, and O. Ambacher, "Noise performance of sub-100-nm metamorphic HEMT technologies," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Aug. 2020, pp. 293–296.

[6] D. Schwantuschke *et al.*, "A fully scalable compact small-signal modeling approach for 100 nm AlGaN/GaN HEMTs," in *Proc. Eur. Microw. Integr. Circuit Conf. (EuMIC)*, Oct. 2013, pp. 284–287.

[7] M. W. Pospieszalski, "Modeling of noise parameters of MESFETs and MODFETs and their frequency and temperature dependence," *IEEE Trans. Microw. Theory Techn.*, vol. 37, no. 9, pp. 1340–1350, Sep. 1989.

[8] W. Choi, G. Jung, J. Kim, and Y. Kwon, "Scalable small-signal modeling of RF CMOS FET based on 3-D EM-based extraction of parasitic effects and its application to millimeter-wave amplifier design," *IEEE Trans. Microw. Theory Techn.*, vol. 57, no. 12, pp. 3345–3353, Dec. 2009.

[9] J. Gao and A. Werthof, "Scalable small-signal and noise modeling for deep-submicrometer MOSFETs," *IEEE Trans. Microw. Theory Techn.*, vol. 57, no. 4, pp. 737–744, Apr. 2009.

[10] A. R. Alt, D. Marti, and C. R. Bolognesi, "Transistor modeling: Robust small-signal equivalent circuit extraction in various HEMT technologies," *IEEE Microw. Mag.*, vol. 14, no. 4, pp. 83–101, Jun. 2013.

[11] A. Jarndal and G. Kompa, "An accurate small-signal model for AlGaN-GaNHEMT suitable for scalable large-signal model construction," *IEEE Microw. Wireless Compon. Lett.*, vol. 16, no. 6, pp. 333–335, Jun. 2006.

[12] G. Van Der Bent, A. P. De Hek, and F. E. Van Vliet, "EM–based GaN transistor small-signal model scaling," in *Proc. 13th Eur. Microw. Integr. Circuits Conf. (EuMIC)*, Sep. 2018, pp. 134–137.

[13] M. Rudolph and R. Doerner, "Bias-dependent Pospieszalski noise model for GaN HEMT devices," in *Proc. German Microw. Conf.*, Mar. 2014, pp. 1–4.

[14] M. Rudolph, R. Doerner, P. Heymann, L. Klapproth, and G. Bock, "Direct extraction of FET noise models from noise figure measurements," *IEEE Trans. Microw. Theory Techn.*, vol. 50, no. 2, pp. 461–464, Feb. 2002.

[15] R. A. Pucel, H. A. Haus, and H. Statz, "Signal and noise properties of gallium arsenide field effect transistors," *Adv. Electron. Electron Phys.*, vol. 38, pp. 195–265, Jan. 1975.

[16] J. Schleeh, H. Rodilla, N. Wadefalk, P. Nilsson, and J. Grahn, "Characterization and modeling of cryogenic ultralow-noise InP HEMTs," *IEEE Trans. Electron Devices*, vol. 60, no. 1, pp. 206–212, Jan. 2013.

[17] M. Seelmann-Eggebert, F. Schäfer, A. Leuther, and H. Massler, "A versatile and cryogenic mHEMT-model including noise," in *IEEE MTT-S Int. Microw. Symp. Dig.*, May 2010, pp. 501–504.

[18] M. R. Murti *et al.*, "Temperature-dependent small-signal and noise parameter measurements and modeling on InP HEMTs," *IEEE Trans. Microw. Theory Techn.*, vol. 48, no. 12, pp. 2579–2587, Dec. 2000.

[19] A. R. Alt and C. R. Bolognesi, "(InP) HEMT small-signal equivalent-circuit extraction as a function of temperature," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 9, pp. 2751–2755, Sep. 2015.

[20] F. Heinz, F. Thome, A. Leuther, and O. Ambacher, "A 50-nm gate-length metamorphic HEMT technology optimized for cryogenic ultra-low-noise operation," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 8, pp. 3896–3907, Aug. 2021.

[21] A. Leuther, A. Tessmann, H. Massler, R. Aidam, M. Schlechtweg, and O. Ambacher, "450 GHz amplifier MMIC in 50 nm metamorphic HEMT technology," in *Proc. Int. Conf. Indium Phosphide Rel. Mater.*, Aug. 2012, pp. 229–232.

[22] G. Dambrine, A. Cappy, F. Heliodore, and E. Playez, "A new method for determining the FET small-signal equivalent circuit," *IEEE Trans. Microw. Theory Techn.*, vol. MTT-36, no. 7, pp. 1151–1159, Jul. 1988.

[23] F. Heinz, D. Schwantuschke, M. Ohlrogge, A. Leuther, and O. Ambacher, "RF-noise model extraction procedure for distributed multiport models," in *Proc. 12th German Microw. Conf.*, Mar. 2019, pp. 260–263.

[24] "High-accuracy noise figure measurements using the PNA-X series network analyzer," Agilent Technol., Santa Clara, CA, USA, Appl. Note 1408-20, Jan. 2013.

[25] J. E. Fernández, "A noise-temperature measurement system using a cryogenic attenuator," Telecommun. Mission Oper. Prog. Rep., Jet Propuls. Lab., California Inst. Technol., Pasadena, CA, USA, TMO Prog. Rep. 42-135, Nov. 1998, pp. 1–9.

[26] R. Q. Twiss, "Nyquist's and Thevenin's theorems generalized for nonreciprocal linear networks," *J. Appl. Phys.*, vol. 26, pp. 599–602, May 1955.

[27] H. Rothe and W. Dahlke, "Theory of noisy fourpoles," *Proc. IRE*, vol. 44, no. 6, pp. 811–818, Jun. 1956.

[28] H. Hillbrand and P. Russer, "An efficient method for computer aided noise analysis of linear amplifier networks," *IEEE Trans. Circuits Syst.*, vol. CAS-23, no. 4, pp. 235–238, Apr. 1976.

[29] S. Diebold *et al.*, "A fully-scalable coplanar waveguide passive library for millimeter-wave monolithic integrated circuit design," in *Proc. Eur. Microw. Conf.*, Oct. 2011, pp. 293–296.

[30] F. Thome, A. Leuther, F. Heinz, and O. Ambacher, "W-band LNA MMICs based on a noise-optimized 50-nm gate-length metamorphic HEMT technology," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2019, pp. 168–171.

[31] M. A. McCulloch *et al.*, "Dependence of noise temperature on physical temperature for cryogenic low-noise amplifiers," *J. Astron. Telescopes, Instrum., Syst.*, vol. 3, no. 1, Mar. 2017, Art. no. 014003.

[32] M. W. Pospieszalski, "On the limits of noise performance of field effect transistors," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2017, pp. 1953–1956.

[33] D. M. Pozar, *Microwave Engineering*, 4th ed. Hoboken, NJ, USA: Wiley, 2011.

[34] D. A. Frickey, "Conversions between S, Z, Y, H, ABCD, and T parameters which are valid for complex source and load impedances," *IEEE Trans. Microw. Theory Techn.*, vol. 42, no. 2, pp. 205–211, Feb. 1994.

[35] R. B. Marks, D. F. Williams, and D. A. Frickey, "Comments on, 'Conversions between S, Z, Y, h, ABCD, and T parameters which are valid for complex source and load impedances' [with reply]," *IEEE Trans. Microw. Theory Techn.*, vol. 43, no. 4, pp. 914–915, Apr. 1995.

Felix Heinz received the M.Sc. degree in embedded systems engineering from Albert-Ludwigs-University Freiburg, Freiburg im Breisgau, Germany, in 2017. He is currently pursuing the Ph.D. degree at the Device Modeling Group, Microelectronics Department, Fraunhofer Institute for Applied Solid State Physics (Fraunhofer IAF), Freiburg im Breisgau.

He joined the Microelectronics Department, Fraunhofer Institute for Applied Solid State Physics (Fraunhofer IAF), in 2017. His current research interests include monolithic microwave integrated circuit design, as well as the characterization and modeling of III–V RF technologies both at room and at cryogenic temperature for applications, such as wireless communication, quantum computing, and radio astronomy.

Fabian Thome received the Dipl.-Ing. degree in electrical engineering and information technologies from the Karlsruhe Institute of Technology, Karlsruhe, Germany, in 2011, and the Ph.D. degree in semiconductor devices and wireless communication from the University of Freiburg, Freiburg im Breisgau, Germany, in 2020.

In 2010, he joined the Fraunhofer Institute for Applied Solid State Physics (Fraunhofer IAF), Freiburg im Breisgau, where he has been a Researcher and a Project Manager since 2012. His current research interests include the development of InGaAs and GaN high-electron-mobility transistor (HEMT) technologies as well as the design and characterization of linear and nonlinear semiconductor devices and monolithic microwave integrated circuits from the microwave-to-submillimeter-wave frequency range for applications in radio astronomy, wireless communications, quantum computing, and measurement equipment.

Dirk Schwantuschke received the Dipl.-Ing. degree in electrical engineering and information technology from the Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, in 2010, and the Dr.-Ing. degree from the University of Stuttgart, Stuttgart, Germany, in 2015.

Since 2011, he has been with the Fraunhofer Institute for Applied Solid State Physics (IAF), Freiburg, Germany, where his research interests include the modeling of AlGaN/GaN high-electron-mobility transistors (HEMTs) and the design of power amplifiers for millimeter-wave (mm-wave) applications. He is currently the Head of the Device Modeling Group, Microelectronics Department.

Arnulf Leuther received the Dipl.-Phys. and Ph.D. degrees in physics from the Technical University of Aachen, Aachen, Germany, 1992 and 1996, respectively.

From 1992 to 1996, he was with Forschungszentrum Jülich, Jülich, Germany. In 1996, he joined the Fraunhofer Institute for Applied Solid State Physics (Fraunhofer IAF), Freiburg, Germany, where he is currently the Head of the Lithography Group. His current research interests include the development of advanced III–V process technologies for metamorphic high-electron-mobility transistors (HEMTs) and the fabrication of millimeter- and submillimeter-wave monolithic microwave integrated circuits (MMICs).

Oliver Ambacher received the Diploma and Ph.D. degrees (Hons.) in natural sciences from the Ludwig Maximilian University of Munich, Munich, Germany, in 1989 and 1993, respectively.

In 1993, he was a Research Assistant at the Walter Schottky Institute, Technical University of Munich, where he dealt with the growth of gallium nitride and its alloys with the help of molecular beam epitaxy and chemical vapor deposition. In 1995, he focused the research work of his group on the processing of GaN-based electronic and optical components. He was significantly involved in the implementation of the first UV detectors, surface acoustic wave components, microwave amplifiers and sensors, as well as in the research of polarization-induced effects in GaN-based heterostructures and quantum structures. From 1998 to 1999, he was offered the Feodor Lynen Grant from the Alexander von Humboldt Foundation at Cornell University (USA) to deepen his work in the field of AlGaN/GaN transistors for high-frequency power amplifiers. Following his habilitation in experimental physics in 2000 and his promotion to Senior Assistant in 2001, he was appointed as a Professor of nanotechnology at the Technical University of Ilmenau, Ilmenau, Germany, a year later. In 2002, he was elected as the Director of the Institute for Solid State Electronics, and two years later, he was appointed as the Director of the Center for Micro and Nanotechnologies. Since October 2007, he has been a Professor at Albert-Ludwigs-University Freiburg, Freiburg im Breisgau, Germany, and the Head of the Fraunhofer Institute for Applied Solid State Physics, Freiburg im Breisgau, where he is currently involved in the development of quantum sensors and electronic components for quantum computers.