Effects of traps in the gate stack on the small-signal RF response of III-V nanowire MOSFETs

Hellenbrand, Markus; Lind, Erik; Kilpi, Olli-Pekka; Wernersson, Lars-Erik

Published in:
Solid-State Electronics

DOI:
10.1016/j.sse.2020.107840

2020

Document Version:
Peer reviewed version (aka post-print)

Link to publication

Citation for published version (APA):
Hellenbrand, M., Lind, E., Kilpi, O-P., & Wernersson, L-E. (2020). Effects of traps in the gate stack on the small-signal RF response of III-V nanowire MOSFETs. Solid-State Electronics, 171, [107840]. https://doi.org/10.1016/j.sse.2020.107840

Total number of authors:
4

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Effects of traps in the gate stack on the small-signal RF response of III-V nanowire MOSFETs

Markus Hellenbrand*, Erik Lind, Olli-Pekka Kilpi, Lars-Erik Wernersson

Department of Electrical and Information Technology, Lund University, Ole Römers Väg 3, 223 63 Lund, Sweden

Abstract

We present a detailed study of the effect of gate-oxide-related defects (traps) on the small-signal radio frequency (RF) response of III-V nanowire MOSFETs and find that the effects are clearly identifiable in the measured admittance parameters and in important design parameters such as $h_{21}$ (forward current gain) and MSG (maximum stable gain). We include the identified effects in a small-signal model alongside results from previous investigations of III-V RF MOSFETs and thus provide a comprehensive physical small-signal RF model for this type of transistor, which accurately describes the measured admittance parameters and gains. We verify the physical basis of the model assumptions by calculating the oxide defect density from the measured admittances.

Keywords: Border Traps, Gate Oxide Defects, Interface Defects, III-V, MOSFET, RF, Small-Signal Model

1. Introduction

Metal-oxide-semiconductor field-effect transistors (MOSFETs) are essential components in high frequency electronics. III-V MOSFETs, due to their narrow band gap and high electron mobility, can achieve higher operation frequencies [1] than corresponding Si MOSFETs, which dominate the digital market. For digital applications, even after several years of investigations into the III-V/high-$\kappa$ material system, III-V MOSFETs are severely impaired by scalability and defects related to the gate oxide, so-called border and interface traps, which degrade reliability and lifetime of the transistors [2]. Two of the most common characterization techniques for traps are bias temperature instability (BTI) measurements in the time domain [3] and capacitance voltage (CV) measurements in the frequency domain [4]. Both techniques have greatly advanced our understanding of defects and led to significant improvement of transistor material qualities, which is reflected in nearly ideal inverse subthreshold slopes and high transconductances [5]. In both techniques, however, the direct observation of traps is typically limited to the megahertz regime and faster traps can only be accessed by decreasing the measurement temperature [6]. The effects of traps in the gate stack on the MOSFET radio frequency (RF) response have thus not yet been investigated in much detail. The few publications which take into account traps in RF investigations to a certain extent, usually focus on a single aspect only and simplify the small-signal model, in which the traps are studied, e.g. by disregarding non-quasistatic (NQS) effects [7–9]. A comprehensive description of the small-signal RF model, which takes into account all different aspects of the contributions of gate-oxide-related traps to the admittance parameters ($y$-parameters), including both real and imaginary parts, and which does not disregard NQS effects, is yet to be developed.

Here, we demonstrate that despite the improvements in recent years, frequency dispersion in III-V MOSFETs at room temperature due to traps can be considerable even at gigahertz frequencies and we investigate the resulting effects on the small-signal $y$-parameters, the RF gains, and the stability factor. Based on our measurements, we provide a comprehensive small-signal model, which takes into account the observed dispersion and which accurately models the measurements. This model also takes into account previous results on small-signal analyses and can be applied at all bias points in the on-state of a transistor and in the off-state before the onset of minority carrier effects. We choose a small-signal approach for this analysis rather than a large-signal model to deconvolute the effect of traps on the RF response from their effect on the quiescent point. This simplifies the analysis and allows us to investigate the effects more clearly. As such, the presented model is not intended for circuit simulations, but it can be used e.g. for small-signal-based amplifier design.

The paper is structured as follows. First, in Section 2, we introduce the devices, which were used to develop and verify the small-signal model and explain the measurement setup and data processing. In Section 3, we briefly discuss the different kinds of defects, which can affect MOSFETs and point out assumptions, which enter into our analysis. In Section 4, we develop expressions for all small-signal model components and demonstrate the effects of traps with measured data in Sections 5 and 6. Lastly, in Section 7, we verify that the model assumptions, which take into account oxide traps, are physically meaningful by calculating the trap densities from the measured admittances.

*Corresponding author

Email address: markus.hellenbrand@eit.lth.se (Markus Hellenbrand)
2. Devices and measurement

The development of the small-signal model presented here was based on measured data from a set of vertical III-V nanowire gate-all-around MOSFETs. The channels consisted of arrays of 180–300 parallel vertical nanowires of InAs graded to InGaAs on the drain side. The nanowire diameters varied between 25 and 30 nm. The high-$\kappa$ gate oxide consisted of a 1 nm/4 nm Al$_2$O$_3$/HfO$_2$ bilayer and gate lengths varied between 100 nm and 220 nm. Details about processing can be found in [10]. For consistency, all measured data presented in this paper are from the same device.

For all investigated devices, scattering parameters ($s$-parameters) were measured with a vector network analyzer between 10 MHz and 67 GHz and for the subsequent analysis the $s$-parameters were converted to $y$-parameters. The measurement setup was calibrated off-chip by a load-reflect-match (LRRM) calibration procedure; on-chip open and short de-embedding structures were used to remove the effect of the pad parasitics. The parameters of the developed small-signal model were then determined by fitting the model to the measured $y$-parameters in Matlab.

The model was developed with a data set measured on the sample described above and verified by subsequent modeling of about thirty additional III-V devices on four different samples. Of the four samples, one was similar to the one described, one contained lateral nanowire MOSFETs, one contained MOSFETs with planar channels, and one contained steep-slope Tunnel FETs. Details about the samples can be found in [11–13]. While the measured parameters differed between samples due to their different geometries, their trends as a function of different bias points were consistent. Furthermore, the effects of traps on the gains, which will be described in Sections 5 and 6, can be discerned in examples from the literature as well [1, 14, 15]. Together with the different modeled samples, this demonstrates the general applicability of the model.

3. Traps in III-V MOSFET gate stacks

The application of high-$\kappa$ oxides on III-V semiconductors can cause many different material defects in the oxide or at the interface between oxide and semiconductor. Conventionally, these defects are often divided into so-called border traps and interface traps, where border traps are commonly associated for instance with oxygen vacancies in the ‘bulk’ of the oxygen layer, i.e. at least a few monolayers away from the oxide/channel interface [16, 17]. Interface traps are believed to be related to disorder, such as dangling bonds, dimers, antisites, or oxygen vacancies at the oxide/channel interface and possibly reaching a few monolayers into the oxide [18, 19]. Density functional theory calculations reveal that all of these defects can create electrically active states in the semiconductor band gap as well as in the conduction band [16–19] in a way that all of these defects can potentially affect MOSFET performance at different biases and time scales. Despite their different chemical natures, the capture and emission of charge carriers by both types of defect can be described by general capture/emission time constants $\tau_{c/e}$, which depend on the tunneling distance and/or the activation energy associated with the charge exchange process [20, 21]:

$$\tau_{c/e} = \tau_0 \exp \left( \frac{E_A}{k_B T} \right).$$

Here, $x$ is the tunneling distance, $\lambda$ the tunneling attenuation length according to the WKB approximation, $E_A$ is the activation energy of a certain defect or defect population, $k_B$ is the Boltzmann constant, and $T$ the temperature. For the effective prefactor $\tau_0$ it is usually assumed that $\tau_0 = 1/(\sigma v_{th})$ with the defect capture cross section $\sigma$, the carrier concentration $n$, and the thermal velocity $v_{th}$ of the charge carriers [20]. In correspondence with a large variation of reported values for $\sigma$, literature values for $\tau_0$ can be found to range from microseconds [22] all the way to the six [23] to below picoseconds [24]. Physically, and in accordance with (1), $\tau_0$ is the time constant below which none of the defects can respond any longer.

In our measurements here, we observe frequency dispersions, which we attribute to traps, even at gigahertz frequencies so that $\tau_0 \approx 0.5$ ps (corresponding to an angular cutoff frequency $\omega_0 = 1/\tau_0 = 2\pi \times 300 \text{ GHz}$) is required to model these dispersions. This is well within the range of values for $\tau_0$ reported in literature and with $v_{th} = 5 \times 10^7 \text{ cm/s}$ and $n = 5 \times 10^{18} \text{ cm}^{-3}$ [25] it results in $\sigma = 8 \times 10^{-15} \text{ cm}^2$, which is a typical value for III-V materials [4, 26, 27]. As a convenient simplification, we typically treated $\tau_0$ as a constant with respect to different bias points, but we point out that this is not a restriction, which is required to achieve agreement between measured values and our developed model.

Results from BTI measurements for example, have demonstrated that at relatively long time scales, the second exponential function in (1) usually dominates the overall time constant [28]. Eq. (1) then becomes $\tau_{c/e} = \tau'_0 \exp(E_A/(k_B T))$ instead and the information about the location of the probed defects is hidden in the effective prefactor $\tau'_0$ with typical values for $\tau'_0$ in the nanosecond range [6, 29]. This, together with $\lambda = 0.13$ nm (for an InAs/Al$_2$O$_3$ interface), our assumption of $\tau_0 \approx 0.5$ ps, and $\tau'_0 = \tau_0 \exp(\lambda/\lambda)$ indicates that the RF measurements here are probing defects at a depth of about 1 nm from the channel interface into the oxide. At the same time, with the measurement frequencies from 10 MHz to 67 GHz, the corresponding time constants $\tau$ are already in the nanosecond range or below even without the exponential energy term. Thus, in the wide distribution of activation energies $E_A$ that is assumed in BTI analyses [2, 6, 30], our measurements here seem to be probing the very smallest $E_A$ so that the inelastic BTI models almost behave like an elastic model. In a purely elastic formulation (i.e. $E_A = 0$), the given measurement frequencies would probe depths in the oxide between 1.3 nm and 0.2 nm, see (1). Based on these considerations, we expect to probe a blend of the aforementioned interface and border traps as well as a combination of elastically and inelastically responding traps.

In the following, we will demonstrate that at radio frequencies, this blend of electrically active traps can be successfully modeled by a simple distributed RC circuit. In CV analyses, sometimes, additional elements are added to this network to
separate interface defects as a distinct factor [22, 31, 32]. Since a distributed RC network simply models time constants without assumptions about the chemical nature or location of the defects in question, this addition proved to be unnecessary in our analysis. We do not consider this as a claim for or against either of the two kinds of defects; however, and just note that the transition between the two is probably fluent. This makes nomenclature ambiguous for the defects probed at radio frequencies so that for the remainder of the paper we will refer to the blend of probed defects simply as ‘traps’. The purpose of this paper is the inclusion of the blend of probed defects simply as ‘traps’. The purpose of our analysis. We do not consider this as a claim for or against all kinds of traps in a MOSFET, can be modeled by a distributed network, which models the exact location, the chemical nature, or the exact capacitance of RC elements as illustrated in Fig. 1(a), where the blend of probed defects simply as ‘traps’. The purpose of this paper is the inclusion of the effect of all kinds of traps in a small-signal model for the high frequency response of III-V MOSFETs.

4. Small-signal model components

The wide distribution of traps with different time constants, which affect a MOSFET, can be modeled by a distributed network of RC elements as illustrated in Fig. 1(a), where the $\Delta C_t$ model the capacitances associated with a certain trap population and the $\Delta G_t$ in series determine the corresponding time constants. The different contributions are separated by incremental parts $\Delta C_{ox}$ of the gate oxide and the RC branch next to the semiconductor capacitance $C_s$ takes into account traps right at the interface. Carrier recombination far in the off-state is not taken into account in this model. Since the RC elements only model time constants, no assumptions are required at this point about the exact location, the chemical nature, or the exact capture mechanism of the corresponding defects.

With a numerical solution of the voltage over $C_s$ in the distributed network in Fig. 1(a) it can be shown that in the frequency range measured here, the distributed network can be approximated by a lumped admittance as depicted in Fig. 1(b), where the real part is linearly dependent on the angular frequency $\omega$, and the imaginary part is linearly dependent on $-\Im(\omega)$ with $\omega_0$ as discussed in Section 3. In a physical interpretation, the real part models losses due to a phase shift between the applied signal and the trap response, and the imaginary part models the capacitance due to traps capturing and emitting charges. The validity of this approximation is demonstrated by the measured data in Sections 5 and 6 and is illustrated in Fig. 4 and Fig. 6. Furthermore, similar RC networks have been successfully employed in BTI [33] and CV investigations [23] and the latter approach led to the same conclusion about the lumped frequency dependencies that we use here.

In order to include the distributed RC network in the MOSFET small-signal model, equivalent lumped admittances like the one in Fig. 1(b) replace the otherwise constant intrinsic gate-to-source and gate-to-drain capacitances $C_{gs}$ and $C_{gd}$, respectively. In Fig. 2, the equivalent admittances are denoted $y_{gs}$ and $y_{gd}$ and with the explanation above, for $\omega < \omega_0$, they can be expressed as

$$y_{gs} = \frac{1}{C_{gs}} + j\omega C_{gs} \ln(\omega/\omega_0),$$

where ‘x’ stands for either ‘s’ or ‘d’ (source or drain), $C_{gs}$ and $C_{gd}$ are the frequency-dependent components of the conductance and the capacitance, respectively, and the $C_{gs,i}$ are the respective intrinsic capacitances without the effect of traps. While linearly frequency-dependent conductances were employed in small-signal models before [8, 9], none of those models included the trap contribution to the capacitances.

Furthermore, those previous models disregarded the NQS channel resistances $R_g$ and $R_d$ in Fig. 2 in series with $y_{gs}$ and $y_{gd}$, respectively, which take into account the delay of charge carriers moving in the channel. This means that no distinction was made between intrinsic ($y_{gs}$ and $y_{gd}$) and parasitic

![Figure 2: Complete small-signal model. $y_{gs}$, $y_{gd}$, $y_{gs}$, and $y_{gd}$ each represent a lumped circuit as in Fig. 1(b). Detailed expressions for the different components are given by (1)–(6) and example values are listed in Table 1.](image-url)

Table 1: Example parameters fitted to the small-signal model in Fig. 2 and used to calculate the modeled $Y$-parameters and gains in Fig. 5.

| Parameter | Value |
|-----------|-------|
| $V_{GS}$, $V_{DS}$ | 0.4 V, 0.5 V |
| $V_T$ | 0.1 V |
| $R_g$ | 5 $\Omega$ |
| $R_d$ | 4 $\Omega$ |
| $g_{ds}$ | 16 $\Omega$ |
| $g_{gd}$ | 10 $\mu S$ |
| $g_{gso}$ | 2.5 $\text{fS/}(\text{rad/s})$ |
| $C_{gs}$ | 20 $\text{fF}$ |
| $C_{gd}$ | 6 $\text{fF}$ |
| $C_{gs,i}$ | 6.0 $\text{fF}$ |
| $C_{gd,i}$ | 1.0 $\text{fF}$ |
| $C_{gs,o}$ | 0.35 $\text{fF}$ |
| $C_{gd,o}$ | 0.2 $\text{fF}$ |
| $g_{ms}$ | 18.7 mS |
| $g_{md}$ | 0.03 |
| $\gamma_1$ | $6 \times 10^{-3}$ |
| $\gamma_2$ | $80 \times 10^{-3}$ |
| $g_{20}$ | 1.55 mS |

$g_{20}$ is zero at this bias point, but differs from zero at lower $V_{GS}$. 
For most MOSFETs even the intrinsic device, i.e. after de-embedding pad parasitics, is subject to some parasitic elements, which contain oxides, they are expected to exhibit the same frequency dependences. This approach works well enough for modelling the \( y \)-parameters and gains and it only adds to the limitation of the quantitative analysis of the trap densities in Section 7.

Besides the admittances from (2) and (4), the transconductance \( g_m \) is also affected by traps. The numerical solution for the voltage over \( C_i \) in the distributed network in Fig. 1(a) reveals that the resulting dispersion affects both the real and the imaginary part. The complete expression for \( g_m \) below \( \omega_0 \) is then

\[
g_m = g_{m,i} \left[ 1 + \gamma_1 \ln (\omega/\omega_0) + j \omega (1 + \gamma_2 \ln (\omega/\omega_0)) \right] - \omega \left( C_m - C_{m,\omega} \ln (\omega/\omega_0) \right),
\]

where \( g_{m,i} \) is the intrinsic transconductance, \( \gamma_1 \) and \( \gamma_2 \) determine the strength of the frequency dispersions of the real and the imaginary part, respectively, and \( \alpha \) scales the imaginary part of this dispersion, since it is significantly smaller than the real part. A derivation of this transconductance-frequency \( (g_m-f) \) dispersion can be found in [7]. \( C_m \) in the purely imaginary part of (5) is the mutual differential capacitance \( C_m = C_{dp} - C_{gd} \), which balances the charge in the channel. Since in the presence of traps, \( C_{dp} \) and \( C_{gd} \) are both frequency-dependent according to (2), \( C_m \) in (5) exhibits the respective frequency dependence as well. In all our measurements, however, this frequency dependence of \( C_m \) was masked by the \( \gamma_2 \) term in (5) so that we set \( C_{m,\omega} = C_{gd,\omega} \).

Physically, equations like (2)–(5) should also apply to the source-drain elements \( C_{sd} \) and \( g_{ds} \), since \( y_{22} \) is subject to the effect of traps as well. A corresponding resistance in series with \( C_{sd} \), however, would always be masked by the large \( g_{ds} \) in par-

\( (y_{gs,p} \text{ and } y_{gd,p} \text{ in Fig. 2 – see next paragraph}) \) components of those models. In order to keep the complexity of the equations for our model to a minimum, we derived analytical expressions for \( R_i \) and \( R_j \) based on an RC relaxation time model [34, 35]. The expressions

\[
R_i = \frac{1}{1.4 \ g_{m,i}} \quad \text{and} \quad R_j = \frac{1}{1.4 \ g_{m,i} \times C_{gd,i}/C_{gs,i}}, \tag{3}
\]

can be obtained by equating the time it takes charge carriers to move through half of the channel with the RC constant consisting of \( R_j(R_i) \) and \( C_{gs,i}(C_{gd,i}) \). \( g_{m,i} \) is the intrinsic transconductance, i.e. without the effect of traps.

For most MOSFETs even the intrinsic device, i.e. after de-embedding pad parasitics, is subject to some parasitic elements, which contain oxides, they are expected to exhibit the same frequency dependences as (2). Furthermore, DC gate leakage can be considered as a parasitic effect and is included in \( y_{gs,p} \) and \( y_{gd,p} \) as well, so that for \( \omega < \omega_0 \),

\[
y_{gs,p} = \omega \times \exp_{,o} + j \omega \left( C_{exp,0} - C_{exp,o} \ln (\omega/\omega_0) \right) + g_{gs,i}.
\tag{4}
\]

Here again, ‘x’ stands for ‘s’ or ‘d’ (source or drain), \( g_{gs,i} \) takes into account the DC gate leakage, \( C_{exp,i} \) is the parasitic capacitance without the effect of traps, and all other parameters are analogous to (2). While the constant parasitic \( C_{exp,0} \) can be determined from the off-state of the transistor, the \( \exp_{,o} \) and \( C_{exp,o} \) are energy- and thus voltage-dependent due to their dependence on the trap energy distribution so that they cannot be subtracted as constant values. As a feasible approximation, we have assumed the same values for the intrinsic and the parasitic frequency dependences. This approach works well enough for modelling the \( y \)-parameters and gains and it only adds to the limitation of the quantitative analysis of the trap densities in Section 7.

\[\text{Figure 3: All } y \text{-parameters, measured (solid, blue lines) and modeled according to Fig. 2 (dashed, red lines), for three representative gate voltages } V_{GS} = -0.2 \text{ V, } 0.1 \text{ V, and } 0.4 \text{ V. } V_{GS} = 0.1 \text{ V. } \text{Re}(y_{11}) \text{ and } \text{Re}(y_{22}) \text{ are plotted with logarithmic frequency axes to illustrate the frequency dispersion and the effects of impact ionization and band-to-band tunneling, respectively. In the subfigures for } \text{Re}(y_{11}) \text{ and } \text{Re}(y_{22}), \text{ example curves are provided for the model without traps at } V_{GS} = -0.2 \text{ V and the deviation is evident. All other modeled curves fit well with the measured data. For these figures, } R_i \text{ and } R_j \text{ have been subtracted, but not } R_g. \text{ Complete expressions for the modeled } y \text{-parameters are provided in the Appendix.}\]
allel with $C_{gd}$ and a $g_{ds} - f$ dispersion corresponding to (5) was not observed. Instead, in virtually all our measurements we observed the onset of band-to-band tunneling (BTBT) or impact ionization (II) already at moderate electric fields between gate and drain. BTBT and II are well known effects in narrow band-gap MOSFETs even in DC measurements. They are not related to traps, but should be included in a comprehensive III-V MOSFET small-signal model. In Fig. 2, BTBT and II are modeled together by the current sources $g_{11}$ and $g_{12}$ with

$$g_a = \frac{g_{i0}}{1 + j \omega \tau_i}.$$  \hspace{1cm} (6)

Here, $g_{i0}$ (with $k = 1, 2$) determines the magnitude and $\tau_i$ is the characteristic time constant. As an approximation, a common $\tau_i$ was used for both effects. In fact, BTBT and II also affect Re($\gamma_{11}$) (cf. $\gamma_{21}$ in the Appendix), but it requires higher gate-to-drain electric fields than in Re($\gamma_{22}$) for these effects to emerge from the $g_m - f$ dispersion in Re($\gamma_{21}$). Since the time constant for BTBT and II is different from the dispersion time constant of the transconductance, the effects can be separated despite appearing in parallel in the model in Fig. 2. In Fig. 3, BTBT and II can be identified by the changing slope of Re($\gamma_{21}$) for $V_{GS} = -0.2$ V and $V_{GS} = 0.1$ V at about 0.1 GHz and by the increase of the two lower curves of Re($\gamma_{22}$) below about 1 GHz. Here it should be noted that a change in the RF output conductance Re($\gamma_{22}$) can also be caused by self-heating. This would result in a distinct stepwise increase in Re($\gamma_{22}$) at about a few megahertz and the effect should increase with increasing $V_{GS}$ [36]. We did not observe this in any of our measurements so that a convolution of our analysis with effects of self-heating is unlikely. Further details about BTBT and II in the small-signal model can be found in [37, 38].

The frequency-dependent expressions (2), (4), and (5) were provided with the constraint that $\omega < \omega_0$. Above $\omega_0$, none of the traps can respond anymore and the expressions become drastically simpler. The transconductance (5) becomes $g_{m,i} = j \omega C_m$ and the intrinsic and parasitic conductances (2) and (4), respectively, become $g_{px,i} = j \omega C_{px,i}$ and $g_{px,p} = j \omega C_{px,p} + g_{px,1}$, respectively. For a physical model, this transition from the expressions below $\omega_0$ to the ones above needs to be continuous. For the logarithmic parts, acceptable continuity is provided by the shape of the logarithm itself. For the real parts of the $g_{px,i}$ and $g_{px,p}$, a transition function would be required, which describes the decrease of the modeled losses back to zero. Since the physical $\omega_0$ seems to lie at higher frequencies than what we were able to measure, we could not investigate this transition of the real part. This does not affect the modeled $\gamma$-parameters, however, since such a transition function would only affect the model close to $\omega_0$ and thus outside of the measurement range.

The complete expressions for the $\gamma$-parameters (below $\omega_0$) are provided in the Appendix and an example of a complete set of small-signal parameters is provided in Table 1.

5. Linear components and their effect on the unilateral power gain and the stability factor

The importance of the linear contributions in (2) and (4) at high frequencies is demonstrated in Fig. 3 by the curves of Re($\gamma_{11}$) and Re($\gamma_{12}$) at $V_{GS} = -0.2$ V. Without it, the modeled Re($\gamma_{11}$) is clearly lower and the modeled Re($\gamma_{12}$) is clearly higher than the respective measured values. At first glance, on the linear scale of Fig. 3, better fits for Re($\gamma_{11}$) and Re($\gamma_{12}$) could be achieved by increasing their quadratic contributions, for example by increasing the NQS resistances or the gate resistance $R_g$. On the logarithmic scale of Fig. 4(a), however, it becomes evident that such a purely quadratic function cannot describe the frequency dependence of Re($\gamma_{11}$) or Re($\gamma_{12}$) correctly. The inclusion of the linear trap contribution greatly alleviates the deviation between the measured values and the purely quadratic model, as demonstrated in Fig. 4(b), where the relative differences for the models with and without traps are compared. Furthermore, the linear contributions to Re($\gamma_{11}$) and Re($\gamma_{12}$) are typically of different magnitudes, so that they cannot be corrected by a single common model parameter such as the gate resistance.

The linearly frequency-dependent contributions in Re($\gamma_{11}$) and Re($\gamma_{12}$) also readily explain the change of the slope of the unilateral power gain $U$, which is evident in Fig. 5. This can be understood from the equation describing $U$, when traps are taken into account according to (2) and (4). Starting from the general definition of $U$ [39], at low and intermediate frequencies ($\leq$ 10 GHz in Fig. 5), where second-order frequency terms

![Diagram](image-url)
are small, \( U \) can be written as

\[
U = \frac{[y_{21} - y_{12}]^2}{4[\text{Re}(y_{11}) \text{Re}(y_{22}) - \text{Re}(y_{12}) \text{Re}(y_{21})]} \approx \frac{|g_{m1}|^2}{4 \omega (g_{gs,0} + g_{gd,0}) g_{ds} + \omega g_{gd,0} \text{Re}(g_{m})},
\]

where the \( g_{gs,0} \) contain the frequency dependences of both the intrinsic and the parasitic elements, i.e. from (2) and (4). Since (7) aims to highlight the importance of the linear frequency dependence (and for better readability), \( g_{ps,1} \) and \( g_{gd,1} \) from (4) are omitted in this approximation. In (7), the linearly frequency-dependent contributions in \( \text{Re}(y_{11}) \) and \( \text{Re}(y_{12}) \) dominate over the usually squared frequency dependences, which would have resulted in the well-known expression \( U \propto g_m^2/(4\omega^2 c_{gg}^2) \) with the total gate capacitance \( C_{gg} \). The now predominantly linear frequency dependence in the denominator of (7) changes the roll-off of \( U \) from the typical -20 dB/decade to almost -10 dB/decade. This can be observed in examples from literature as well [14, 15]. When compared with the model without traps, in a certain frequency range this leads to a decrease of the unilateral power gain by almost 10 dB. For the design of unilateral amplifiers, which are limited by \( U \), this is a severe penalty. The constant level of \( U \) at low frequencies is caused by the constant leakage contributions \( g_{ps,1} \) and \( g_{gd,1} \) in (4). Furthermore, it can be noted that \( U \) is affected by the logarithmic frequency dependence of \( g_m \). However, the linear components in (7) largely dominate over this logarithmic contribution.

Besides the roll-off of \( U \), the linear components in \( \text{Re}(y_{11}) \) and \( \text{Re}(y_{12}) \) affect the stability factor \( k \), as illustrated in Fig. 5 as well. The modeled \( k \) without the inclusion of traps is always lower than the measured values. In the equation for \( k \), the effect of the linear contributions can be seen clearest at frequencies, where \( k \) is almost constant, i.e. between 1 GHz and 10 GHz in Fig. 5. At these frequencies, starting from its general definition [39], \( k \) can be written as

\[
k = \frac{2 \text{Re}(y_{11}) \text{Re}(y_{22}) - \text{Re}(y_{12}) y_{21}}{[y_{21} y_{12}]} \approx \frac{2\omega (g_{gs,0} + g_{gd,0}) g_{ds} + \omega g_{gd,0} \text{Re}(g_{m})}{\left(-\omega g_{gd,0} - j\omega (C_{gd,1} + C_{gd,p}) g_{m}\right)},
\]

where again, the \( g_{gs,0} \) contain the frequency dependences of both the intrinsic and the parasitic elements and \( g_{ps,1} \) and \( g_{gd,1} \) from (4) are omitted for increased readability. In this approximation it is obvious that without the effect of traps \( g_{gs,0} \), \( k \) would become zero. Indeed, in Fig. 5, for the model without traps, \( k \) almost vanishes in the range between 1 GHz and 10 GHz, so that by the presence of traps, \( k \) is increased by a factor of up to five. Since the shift of \( k \) appears to be mostly constant in Fig. 5, the relative difference decreases towards higher frequencies, but can still amount to up to 30 % at 60 GHz. In Fig. 5, together with the effect of the traps on the gain, this lowers the \( k \)-point, i.e. the point, where MSG changes to the maximum available gain MAG, by 15 to 20 GHz. The remaining small difference of \( k \) from zero in the measured values stems from the constant leakage contributions \( g_{ps,1} \) and \( g_{gd,1} \) in \( y_{gs,p} \) and \( y_{gd,p} \), which were omitted from (8). The lower \( k \)-point decreases the frequency range available for circuit design.

6. Logarithmic components and their effect on the forward current gain and the maximum stable gain

Logarithmic frequency dispersions due to traps are present in the transconductance, (5), and in the intrinsic and parasitic capacitances in (2) and (4), respectively. The frequency dispersion in the real part of the transconductance is visible in \( \text{Re}(y_{21}) \) in Fig. 3. The further increase of the measured \( \text{Re}(y_{21}) \) at the higher \( V_{GS} \) beyond approx. 20 GHz, as well as the further increase of \( \text{Re}(y_{11}) \) above approx. 35 GHz, is most likely due to the shortcomings of the open-short de-embedding at high frequencies.

Frequency dependences in capacitances due to traps have been resolved in CV measurements before [23], but not in transistor measurements. In Fig. 6(a), the measured \( \text{Im}(y_{11})/\omega \) and \( \text{Im}(y_{12})/\omega \), which approximately correspond to the capacitances in (2) and (4), are compared with the models with and without the effect of traps. The model without traps results in horizontal lines, which corresponds to ordinary capacitances, whereas the measured values clearly deviate from this behavior. Upon inclusion of the logarithmic frequency dependences in (2) and (4), the deviation readily disappears. To quantify this, the relative differences of the measured and modelled \( \text{Im}(y_{11}) \) and \( \text{Im}(y_{12}) \) are plotted in Fig. 6(b). At high frequencies, less traps can respond, so that the capacitances of the model, which includes traps, are reduced to the intrinsic capacitances and the difference between the models with and without traps vanishes. Towards lower frequencies, the agreement between measured and modeled values is improved by up to 20 % by the inclusion of traps in the model.

The same logarithmic deviation between measured and modeled values, which is observed in \( \text{Im}(y_{11}) \) and \( \text{Im}(y_{12}) \), can be identified in the forward current gain \( h_{21} \) and the maximum stable gain MSG in Fig. 5. Just as for the \( y \)-parameters, the deviation disappears upon inclusion of the frequency-dependent capacitances. Again, the difference between the two models can
be explained by the equations of the two quantities. Starting with its general definition [39], $h_{21}$ can be written as

$$h_{21} = \frac{|y_{21}|}{|y_{11}|} \approx \left| \frac{G_m}{j\omega (C_{gs} + C_{gs,p} + C_{gd} + C_{gd,p})} \right|,$$  \hspace{1cm} (9)

where all four capacitances denote the total frequency-dependent capacitances from (2) and (4), e.g. $C_{gd} = C_{gd,i} - C_{gd,ox}\times \ln(\omega/\omega_0)$. In the approximation in (9), typically, the real parts of $y_{11}$, which are important for e.g. $U$ and $k$, are significantly smaller than the imaginary parts. Without the trap contribution, in a Bode plot, (9) would yield the typical -20 dB/decade roll-off. When the effects of traps are included, logarithmic frequency dependences are added to both the numerator (decrease towards lower frequencies) and to the denominator (increase towards lower frequencies), which causes the lower slope for $h_{21}$ in Fig. 5. The same applies for MSG, which, again, starting from its general definition [39], can be written as

$$MSG = \frac{|y_{21}|}{|y_{11}|} \approx \left| \frac{G_m}{j\omega (C_{gs} + C_{gs,p})} \right|,$$ \hspace{1cm} (10)

As in (9), both capacitances include the frequency dependences from (2) and (4) and as for $U$, the deviations of the slopes of $h_{21}$ and MSG can be observed in examples in literature as well [1, 14]. MSG is an important metric for the design of small-signal amplifiers and the dispersion in the transistor MSG is directly translated to the amplifier. Since the dispersion in MSG is much smaller than in $U$, the resulting penalty in the amplifier should be acceptable as long as the design is not dependent on single decibels of gain. In any case, it is important to be aware of the dispersion, so that the amplifier design will not be based on an idealized and thus incorrect transistor gain.

7. Physical origin of trap model parameters

In the following, the trap density $N_t$ will be calculated from the measured admittances to verify that the small-signal parameters, which model the oxide traps, are physically meaningful. The approach to deriving an expression for $N_t$ is indicated in Fig. 1. At a frequency $\omega$, the part of the distributed network, which responds at frequencies up to this $\omega$, can be seen as an admittance $Y(\omega)$. The change $\Delta Y(\omega)$ for a change $\Delta \omega$ can then be expressed as $\Delta Y(\omega) = Y(\omega) - Y(\omega - \Delta \omega)$, where $Y(\omega - \Delta \omega)$ is the admittance corresponding to an incrementally lower $\omega$. The trap density $N_t$, as detailed in [23], in included in this model as $\Delta G = q^2 N_t \lambda x$, where $q$ is the elemental charge, and the $\Delta G$ are the incremental parts of the trap population within a thickness $\Delta x$ in the gate oxide. The time constants $\tau = \Delta C/G$, thus the different $\tau$ can still be modeled independently of the position $\Delta x$ in the oxide. To be able to calculate $N_t$ from the measured admittances, the angular frequency $\omega$, and thus the time constants $\tau$, are related to $\Delta x$ via elastic tunneling as $\Delta x = \lambda \Delta \ln(\omega)$, where $\lambda = 0.13$ nm is the tunneling attenuation length according to the WKB approximation as in Section 3. As discussed in Section 3 as well, elastic tunneling should be a viable approximation at gigahertz frequencies. With this, a change in the admittance $Y(\omega)$ of the distributed RC network in Fig. 1 ($y_{11}$ or $y_{12}$ in the small-signal model) for a change in the angular frequency $\omega$ can be expressed as

$$\frac{\Delta Y(\omega)}{\Delta \ln(\omega)} = \frac{1 + j}{2} \left( \omega \lambda q^2 \Delta N_t L_G W_G \right),$$ \hspace{1cm} (11)

where the term $\lambda \Delta \ln(\omega)$, which appears during the derivation, is sufficiently small to be dropped from the expression. In (11), $L_G$ and $W_G$ are the gate length and width of the transistor, respectively. Examples for $N_t$ calculated from the measured $y_{12}$ are presented in Fig. 7 alongside a calculation of $N_t$ based on the $g_m$-f dispersion as in [7]. The values for $N_t$ calculated from $y_{11}$ exhibited a similar behavior as the ones calculated from $y_{12}$. For the calculations according to (11), first, the elements surrounding $y_{11}$ and $y_{12}$ in the small-signal model in Fig. 2 have to be subtracted until only the contribution of the $y_{gd}$ or $y_{gs}$ remains. The resulting uncertainties, indicated by the shaded areas in Fig. 7, reveal that an exact quantitative resolution of $N_t$, e.g. as a function of the gate voltage and thus of energy, is limited by the measurement accuracy of the $s$ parameters and the subsequent deconvolution of the $y$-parameters. For the largest parts of the bias range, however, the average values for $N_t$ calculated from the different components are within an order of magnitude of each other and they are in good agreement with results.
that both we and others have observed with different characterization techniques [2, 40–43]. The results are thus sufficiently accurate to verify the physical origin of the model assumptions.

8. Conclusions

The effects of traps related to the gate oxide are clearly discernible in the RF $y$-parameters of III-V MOSFETs. The inclusion of traps in $y_{12}$ and $y_{12}$ is thus essential to accurately describe a small-signal model for these devices. Although the resulting effects on design parameters such as $h_{21}$ and MSG are small, the effects have to be identified in the first place to be able to draw this conclusion with certainty. The calculation of the trap density $N_t$ based on the presented model can be limited by the measurement accuracy, but it yields values for $N_t$, which are in agreement with values calculated by other techniques. This demonstrates that the model parameters reasonably reflect physical assumptions.

Acknowledgements

We thank J. Svensson for growing the nanowires for the vertical nanowire MOSFETs, F. Lindelöw and J. Mo for providing us with samples of lateral nanowire MOSFETs and of planar MOSFETs, respectively, and S. Andrić for helpful discussions. This work was supported in part by the Swedish Foundation for Strategic Research and the Swedish Research Council, and in part by the European Union through the H2020 Program IN-SIGHT under Grant 688784.

Appendix

Here we provide the complete expressions for the $y$-parameters for the model in Fig. 2. For convenience, we repeat the expressions for all parameters, which enter into the $y$-parameters:

$$y_{gx} = \omega \times g_{gx,0} + j\omega \left(C_{gx,0} - C_{gx,0}\ln(\omega/\omega_0)\right),$$

$$y_{xp} = \omega \times g_{xp,0} + j\omega \left(C_{xp,0} - C_{xp,0}\ln(\omega/\omega_0)\right) + g_{xl},$$

$$g_{m} = g_{mi} \left[1 + \gamma_1 \ln(\omega/\omega_0) + j\omega(1 + \gamma_2 \ln(\omega/\omega_0))\right] - j\omega \left(C_{m,0} - C_{m,0}\ln(\omega/\omega_0)\right),$$

$$g_{d} = \frac{g_{m0}}{1 + j\omega\tau_1},$$

$$R_s = \frac{1}{1.4 g_{mi}} \quad \text{and} \quad R_f = \frac{1}{1.4 g_{mi} \times C_{gd,0}/C_{gx,0}}.$$

With these, the intrinsic $y$-parameters (i.e. after subtracting the resistances $R_s$, $R_1$, and $R_f$) are

$$y_{11} = \frac{i_1}{y_{12}} \left|_{V_{12}=0} = y_{gp} + \frac{y_{gs}}{1 + y_g R_1} - y_{12},\right.$$

$$y_{12} = \frac{i_2}{y_{21}} \left|_{V_{12}=0} = -y_{gp} - \frac{y_{gd}}{1 + y_g R_1},\right.$$  

$$y_{21} = \frac{i_2}{y_{21}} \left|_{V_{12}=0} = g_{m} \frac{1 + y_g R_1}{1 + y_g R_1} + y_{12} - g_{11} - \frac{g_{22}}{1 + y_g R_1},\right.$$  

$$y_{22} = \frac{i_2}{y_{22}} \left|_{V_{12}=0} = g_{ds} + j\omega C_{gd} - y_{12} + g_{11},\right.$$  

References

[1] A. Tessmann, A. Leuther, F. Heinz, F. Bernhardt, L. John, H. Massler, L. Czornomaz, T. Merkle, 20-nm In$_{0.53}$Ga$_{0.47}$As MOSFET, IEEE Electron Device Letters 33 (3) (2012) 369–371. doi:10.1109/LED.2011.2181323.

[2] J. Franco, V. Putcha, A. Vais, S. Sioncke, N. Waldron, D. Zhou, G. Reza, P. J. Roussel, G. Groeseneken, M. Heyns, N. Collaert, D. Linten, T. Grasser, B. Kaczer, Characterization of oxide defects in InGaAs MOS stacks for high-mobility n-channel MOSFETs, IEEE Transactions on Electron Devices 60 (2) (2013) 776–781. doi:10.1109/TED.2012.2231867.

[3] B. Kaczer, T. Grasser, J. Roussel, J. Martin-Martinez, R. O’Connor, B. J. O’Sullivan, G. Groeseneken, Ubiquitous relaxation in BTI stressing—new evaluation and insights, in: 2008 IEEE International Reliability Physics Symposium, 2008, pp. 20–27. doi:10.1109/RELPHY.2008.4558858.

[4] R. Engel-Herbert, Y. Hwang, S. Steiner, Comparison of methods to quantify interface trap densities at dielectric/III-V semiconductor interfaces, Journal of Applied Physics 108 (12) (2010) 124101. doi:10.1063/1.3520431.

[5] H. Riel, L.-E. Wernersson, M. Hong, J. A. del Alamo, III–V compound semiconductor transistors—from planar to nanowire structures, MRS Bulletin 39 (8) (2014) 668–677. doi:10.1557/mrs.2014.137.

[6] V. Putcha, J. Franco, A. Vais, S. Sioncke, B. Kaczer, D. Linten, G. Groeseneken, On the apparent non-arrhenius temperature dependence of charge trapping in III-V/high-k MOS stack, IEEE Transactions on Electron Devices 65 (9) (2018) 3689–3696. doi:10.1109/TED.2018.2851189.

[7] S. Johannson, M. Berg, K. Persson, E. Lind, A high-frequency transconductance method for characterization of high-k border traps in III-V MOSFETs, IEEE Transactions on Electron Devices 60 (2) (2013) 776–781. doi:10.1109/TED.2012.2231867.
