Toward Automated Defect Extraction From Bias Temperature Instability Measurements

Dominic Waldhoer, Christian Schleich, Jakob Michl, Bernhard Stampfer, Konstantinos Tselios, *Graduate Student Member, IEEE*, Eleftherios G. Ioannidis, Hubert Enichlmair, *Member, IEEE*, Michael Waltl, *Member, IEEE*, and Tibor Grasser, *Fellow, IEEE*

***Abstract*— Defects in the gate oxide give rise to bias temperature instability (BTI), which is considered a serious threat to the device reliability of ultrascaled MOSFETs. Extrapolating the device degradation over the operational lifetime, therefore, requires detailed knowledge about the distributions of defects causing BTI. Typically, BTI degradation is modeled by calibrating a *predefined* defect parameter distribution, such as normally distributed defect bands, by employing measure-stress-measure (MSM) sequences at various temperatures. Here, we present the *Effective Single Defect Decomposition* (ESiD), a novel method for a *semiautomated extraction* of defect parameters from MSM experiments *which does not require any prior assumptions about their distributions*. This technique decomposes the MSM sequences into contributions from dominant effective single-defects and constructs a defect parameter distribution that reproduces the experimental data with high accuracy. We validate this new method by comparing its results to density functional theory (DFT) predictions and single-defect characterizations.**

***Index Terms*— Bias temperature instability (BTI), charge trapping, defect characterization, measure-stress-measure (MSM), nMOS, pMOS.**

I. Introduction

THE *Bias Temperature Instability* (BTI) [1] refers to a shift of the threshold voltage ΔV_{th} during device operation. Based on our current understanding, this shift is primarily caused by charge trapping and defect creation in the oxide [2] or the interfacial layer (IL) [3], and is one of the limiting reliability challenges for modern MOSFET technologies.

Manuscript received March 26, 2021; revised May 26, 2021 and June 18, 2021; accepted June 20, 2021. Date of publication July 2, 2021; date of current version July 23, 2021. The financial support by the Austrian Federal Ministry for Digital and Economic Affairs, the National Foundation for Research, Technology and Development and the Christian Doppler Research Association is gratefully acknowledged. The review of this article was arranged by Editor R. Wang. *(Corresponding author: Dominic Waldhoer.)*

Dominic Waldhoer, Christian Schleich, Bernhard Stampfer, Konstantinos Tselios, and Michael Waltl are with the Christian Doppler Laboratory for Single-Defect Spectroscopy in Semiconductor Devices, Institute for Microelectronics, TU Wien, 1040 Vienna, Austria (e-mail: waldhoer@iue.tuwien.ac.at).

Jakob Michl and Tibor Grasser are with the Institute for Microelectronics, TU Wien, 1040 Vienna, Austria.

Eleftherios G. Ioannidis and Hubert Enichlmair are with ams AG, 8141 Premstätten, Austria.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TED.2021.3091966.

Digital Object Identifier 10.1109/TED.2021.3091966

Predicting BTI degradation over the device lifetime, therefore, relies on detailed knowledge about the involved defects and their charge trapping behavior under varying temperatures and gate biases. In ultrascaled devices, in which charge trapping of individual defects is directly observable, such information can be obtained by single-defect characterization techniques like *random telegraph noise* (RTN) analysis [4] or *time-dependent defect spectroscopy* (TDDS) [5]. However, these methods are not applicable to large-area devices or material systems exhibiting larger defect densities like SiC/SiO_2. In such cases, typically recovery traces $\Delta V_{\text{th}}(t)$ after degradation at accelerated stress conditions are measured using extended measure-stress-measure sequences (eMSM) [6]. Applying a calibrated device model including a physical description of the charge transfer between defect and charge reservoir, defect parameters like trap levels or relaxation energies can be extracted by tuning the model to match the observed recovery traces [7]–[9]. Note that using a physics-based defect model is key here in order to extrapolate the BTI degradation from the accelerated stress conditions back to typical biases and temperatures during operation.

However, due to the amorphous nature of the gate dielectric, the defect parameters are widely distributed. Since ΔV_{th} is the collective response of a whole defect ensemble, an appropriate distribution function has to be used when extracting defect parameters from eMSM sequences. The energetic parameter distribution of the defects is commonly assumed to be Gaussian, where mean and standard deviation serve as fitting parameters for the model. It has been demonstrated that this approach can match experimental degradation [7], [10] well and also provides reasonable agreement with theoretical trap levels [11] and activation energies [12] obtained from density functional theory (DFT) calculations. However, the fit relaxation energies are often far too high to be considered physical when compared to typical values for probable defect candidates [13].

As shown in Fig. 1, the parameter optimization for a chosen distribution function can lead to spurious solutions, with large parts of the defect band being inactive even under severe stress conditions. Although these inactive defects do not affect the ΔV_{th} prediction, they distort the statistics of relevant defects and lead to higher apparent mean values. Likely examples for such a misinterpretation of data, resulting in unphysically high extracted parameters, can be found in the literature, see e.g., [7]–[10]. This issue also has to be taken into account

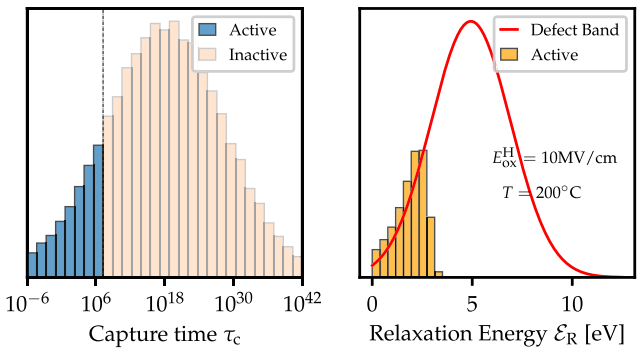

Fig. 1. Extracted τ_c and $\mathcal{E}_R$ distribution for electron traps in SiO_2 as published in [7]. Even under severe conditions, only a small fraction of the distribution is active ($\tau_c < 10^7$ s), leading to significant deviations from the assumed Gaussian distribution when only considering the active defect set.

when comparing experimentally extracted defect parameters to theoretical studies.

Furthermore, due to the highly nonlinear dependence of the degradation response on the defect parameters, a good initial guess and frequent manual intervention are required during the parameter optimization process. This becomes particularly cumbersome and time-consuming when dealing with advanced gate stacks, several different processes splits or material systems like SiC/SiO_2 where multiple defect bands are involved [14].

To overcome these issues, we present the *Effective Single Defect Decomposition* (ESiD), a novel defect extraction method based on a nonnegative linear least square (NNLS) algorithm, which decomposes the ΔV_{th} degradation into effective single-defect contributions. This method allows for a semiautomatic extraction of physical defect parameters from experimental eMSM recovery traces which *requires only a little manual effort and does not depend on predefined distribution functions*. We employ this method to characterize both electron and hole traps in large-area SiON devices. It will be shown that the defect parameters obtained with our method are in good agreement with theoretical DFT studies of hydrogen-related defects like the hydroxyl-E' center [15] or the hydrogen bridge (HB) [16] in SiO_2, emphasizing the important role of hydrogen in reliability physics. We will further verify our method by comparing the extracted defect distributions with previously published single-defect characterization data obtained on similar but scaled devices using TDDS.

II. Experimental Setup

For the purpose of this work, we extract electron and hole trap distributions from pMOS/negative BTI (NBTI) and nMOS/positive BTI (PBTI) measurements on large-area MOSFET devices ($W \times L = 10 \times 10$ μm) fabricated within a commercially available 180 nm SiON technology. The temporal recovery of device degradation after stress has been monitored with our custom-built low-noise measurement tool [17] using an eMSM scheme with stress times in the range of $t_{str} = 100$ μs to 10 ks and recovery times up to $t_{rec} = 10$ ks. The ΔV_{th} readout has been performed at $I_D = 100$ nA·W/L with a minimal delay after stress of 100 μs. In order to accurately account for the bias- and temperature dependence of BTI, the experiments were performed at three stress fields $|E_{ox}^{H}| = 5$, 6.5, and 8 MVcm^{-1}, and at

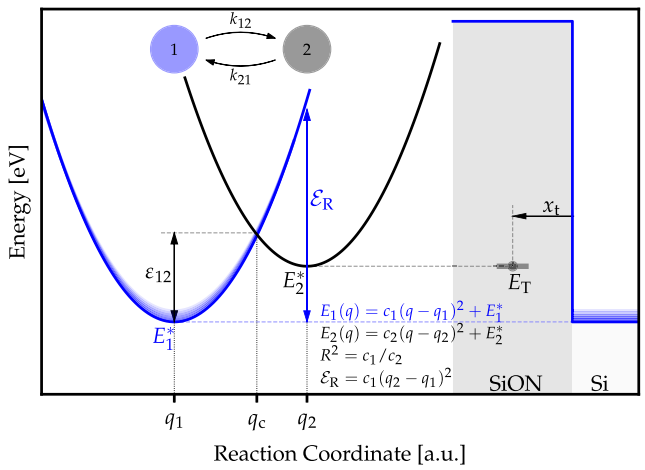

Fig. 2. Configuration coordinate diagram and the describing parameters of a charge transition within NMP theory. The defect states are modeled as harmonic oscillators. In the classical limit, the bias-dependent intersection point of the two parabolas determines the reaction barrier for a charge transfer between defect and device channel.

three temperatures $T = 50$, 100, and 150 °C. The ΔV_{th} response has been extracted by converting ΔI_D according to the initial $I_D(V_G)$ characteristics of the unstressed device [6]. For each device, low-frequency capacitance–voltage (CV) curves at $f = 100$ kHz were recorded and used to calibrate the electrostatics in our device models.

III. Physical Model

To model the impact of oxide defects on the threshold voltage V_{th}, we employ our previously developed 1-D device reliability simulator *Comphy* [7]. The simulator consists of an analytic device model based on a surface-potential description [18], which provides quantities like the electric field E_{ox} and the carrier concentrations n, p in the channel as a function of V_G and T. These data are subsequently used to calculate the charge transfer rates between oxide defects and channels within a nonradiative multiphonon (NMP) defect model [19]. Here, the charged and uncharged defect states of an effective two-state model are described as quantum mechanical harmonic oscillators forming a Markov chain as shown in Fig. 2. Within this model, the charge transfer is then treated as a nonadiabatic transition between the two potential energy surfaces (PES) [13]. In the classical limit of NMP theory, the transition rates are solely determined by the classical barriers $\varepsilon_{12/21}$ given by the bias-dependent intersection point q_c. Using this approximation, the charge capture and emission times for a pMOS are given by [19]

$$\tau_{c/e}^{-1} = p\sigma v_{th}\vartheta \exp\left(-\varepsilon_{12/21}/k_B T\right). \quad (1)$$

Here, ϑ denotes a Wentzel-Kramers–Brillouin (WKB) tunneling factor accounting for the distance between defect and charge reservoir, v_{th} is the thermal velocity of the carriers within the channel. The capture cross section was assumed to be $\sigma = 10^{-15}$ cm^2 throughout this work [20], [21]. Within the classical approximation, the scaling of the q-axis is immaterial and thus the defect behavior is completely determined by four parameters: the *relaxation energy* $\mathcal{E}_R = S\hbar\omega$, which is related to the Huang–Rhys factor S [22]; the *curvature ratio* between the two PESs $R^2 = c_1/c_2$; the *trap level* E_T, which determines the energetic alignment between the defect and the band edges;

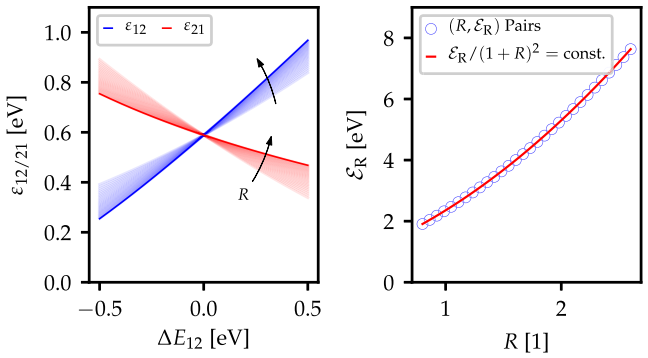

Fig. 3. Charge transition barriers for multiple $(R, \mathcal{E}_R)$ pairs (left). As long as these pairs are along the curve $\mathcal{E}_R/(1+R)^2 = \text{const.}$ (right), the corresponding barriers are experimentally nearly indistinguishable. Therefore, in order to extract defect parameters experimentally, one has to fix either R or $\mathcal{E}_R$.

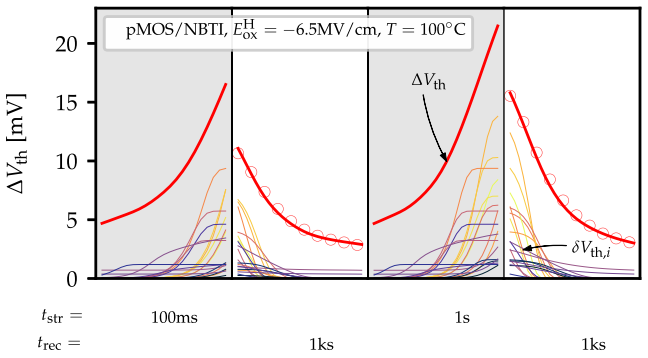

Fig. 4. Decomposition of an eMSM sequence into single defect contributions δV_{th} (Scale: $\times 300$) using our ESiD method. As can be seen, the MSM signal is composed of many defects spanning a wide range of time scales.

and the *trap location* x_t, which influences the WKB tunneling rate as well as the bias dependence. Note that this model only describes charge trapping in preexisting oxide defects. Other effects like the creation of interface defects during stress have been successfully described as an effective first-order reaction within an empirical double-well model [23], [24] mathematically similar to (1). Therefore, also such processes can be described with artificial oxide defects within the NMP model, although the associated parameters carry no physical meaning. This will be exploited here to model the (quasi)-permanent BTI component P within NMP theory. The corresponding fictitious defects can then be identified based on their large time constants.

In previous studies on defect characterization, the energetic parameters E_T and $\mathcal{E}_R$ have been assumed to follow a Gaussian distribution, while a homogeneous spatial distribution up to a maximum depth, typically 0.6 nm, was used [7]. Together with a single value for the curvature ratio R and the total trap density N_T, a defect band is therefore characterized by the parameter tuple

$$\mathcal{P} = \left(\langle E_T\rangle, \sigma_{E_T}, \langle \mathcal{E}_R\rangle, \sigma_{\mathcal{E}_R}, R, x_t, N_T\right). \tag{2}$$

This tuple is then optimized against the experimental degradation using local optimization schemes. However, the restriction to specific distributions can lead to artificial defect parameters.

Furthermore, such a model description should be checked for cross correlation between different parameters in order to ensure a unique parameter set. In the case of the two-state NMP model, as given above, we observe a nonlinear correlation between the parameters $\mathcal{E}_R$ and R. A Taylor expansion of the capture barrier ε_{12} with respect to the energy difference ΔE_{12} between the two states yields [25]

$$\varepsilon_{12} = \frac{\mathcal{E}_R}{(1+R)^2} + \frac{R}{1+R}\Delta E_{12} + \mathcal{O}\left(\Delta E_{12}^2\right). \tag{3}$$

As can be seen in Fig. 3, the capture and emission barriers stay approximately the same within a wide range of R when the zeroth-order term is kept fixed. Therefore, different pairs $(R, \mathcal{E}_R)$ along this curve are expected to give similar degradation responses, leading to a nonunique parameter set to explain the experimental data. Note that, while the slope of the individual barriers shows a weak dependence on R, this difference becomes undetectable when dealing with a whole defect band instead of individual traps. In an earlier work [7], a rather high curvature ratio of $R = 2.59$ together with mean relaxation energy of 7.95 eV have been reported for hole traps in SiON. However, such values seem unphysical and disagree with theoretical DFT studies on defects in amorphous silica, which typically predict values in the range $R = 0.8 - 1.2$ and $\mathcal{E}_R = 2 - 4$ eV [13]. Since this correlation prohibits the simultaneous extraction of both $\mathcal{E}_R$ and R from typical eMSM sequences, we limit our model to the linear electron-phonon coupling regime [26] given by $R = 1$ in order to be consistent with the theoretical results. Using our novel extraction method described below, we will show that such a defect parametrization can accurately reproduce the measurement data while also providing defect parameters in good agreement with DFT.

IV. Effective Single Defect Decomposition (ESiD)

Instead of fitting the parameter tuple $\mathcal{P}$ to experimental data, we directly exploit the fact that the macroscopically measurable $\Delta V_{th}(t)$ results from trapping and detrapping of a whole defect ensemble. In the limit of low defect concentrations, the charge state of one defect does not influence the behavior of another defect via electrostatic interaction. In this case, the total degradation can be expressed as a linear superposition of the form

$$\Delta V_{th}(t) = \int_{\Omega} N(\boldsymbol{p})\delta V_{th}(t;\boldsymbol{p})d\boldsymbol{p} \tag{4}$$

where $\delta V_{th}(t;\boldsymbol{p})$ denotes the contribution to ΔV_{th} of a single defect with parameters $\boldsymbol{p} = (E_T, \mathcal{E}_R, x_t)$, see also Fig. 4. All these contributions are then weighted according to a distribution function $N(\boldsymbol{p})$ and integrated over the parameter space $\boldsymbol{\Omega}$. Expressed in this language, the extraction of defect parameters from experiments is equivalent to finding a suitable distribution $N(\boldsymbol{p})$ over $\boldsymbol{\Omega}$ which matches the observed degradation. Although the required assumption of noninteracting defects might seem too restrictive, in practice the defect–defect interaction is mostly negligible in the context of BTI. However, the validity of this assumption can be verified retrospectively by using the obtained $N(\boldsymbol{p})$ in a self-consistent simulation and comparing the results to the simulation with noninteracting defects. For the data published here, the maximal error was found to be below 4%.

While previous works assumed $N(\boldsymbol{p})$ to be Gaussian along the energetic dimensions $E_T, \mathcal{E}_R$ and uniform along x_t,

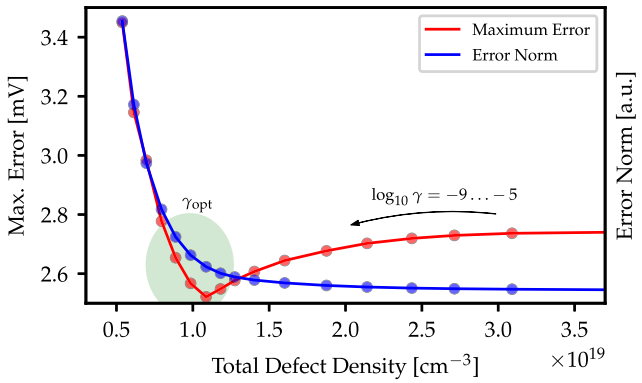

Fig. 5. Impact of the regularization parameter γ. For small γ, the solution requires an unphysically large defect density, whereas too large values lead to quickly growing approximation errors. The optimal value for γ lies between these two detrimental regions.

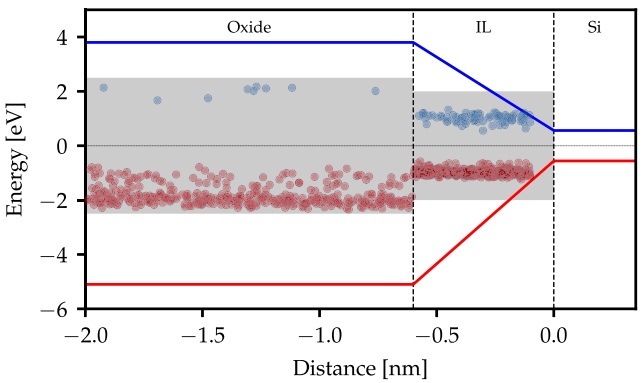

Fig. 6. Samples drawn from the extracted defect distributions within the band diagram. The shaded areas indicate the search regions used in ESiD. The highest concentration of both electron (blue) and hole (red) traps is located within the IL.

we employ a NNLS estimator to retrieve $N(\boldsymbol{p})$ from measurements without dictating the resulting distribution. In order to do so, we first select a physically reasonable search region within the parameter space Ω, e.g., by choosing a certain energy region for E_T and $\mathcal{E}_R$, and discretize it using a uniform grid of parameter points $\boldsymbol{p}_i$. Next, we employ the Comphy modeling framework to calculate the single defect responses at every grid point and collect them in the *response matrix* $(\boldsymbol{\delta V})_{ji} = \delta V_{\text{th}}(t_j; \boldsymbol{p}_i)$. Note that the ith column of $\boldsymbol{\delta V}$ represents the concatenated theoretical responses of *all* recorded MSM sequences for the defect parameters $\boldsymbol{p}_i$. Similarly, we construct a column vector $(\boldsymbol{\Delta V})_j = \Delta V_{\text{th}}(t_j)$ called the *observation vector*.

In this discretized form, a naive NNLS estimator for the distribution function can be expressed as

$$\boldsymbol{N} = \underset{\hat{\boldsymbol{N}} \geq \boldsymbol{0}}{\arg\min} \left\| \boldsymbol{\delta V} \cdot \hat{\boldsymbol{N}} - \boldsymbol{\Delta V} \right\|_2^2 \tag{5}$$

where the column vector $(\boldsymbol{N})_i = N(\boldsymbol{p}_i)$ is a discrete representation of the sought distribution function $N(\boldsymbol{p})$. Note that the enforced nonnegativity constraint is essential here, since N ought to represent a physical defect density, which can only take semipositive values. However, decomposing the measured ΔV_{th} into single-defect contributions is mathematically an *ill-posed* problem, leading to numerically unstable solutions, which frequently exhibit unphysical total defect densities and a spiky distribution function. The reason for this unwanted behavior lies in the fact that the estimator (5) tries to minimize the error to the measurement data as much as possible without any regard to the total density required or the smoothness of the solution. To mitigate this, we regularize the problem by introducing a *Tikhonov* term [27] to the estimator

$$\boldsymbol{N} = \underset{\hat{\boldsymbol{N}} \geq \boldsymbol{0}}{\arg\min} \left\| \boldsymbol{\delta V} \cdot \hat{\boldsymbol{N}} - \boldsymbol{\Delta V} \right\|_2^2 + \gamma^2 \left\| \hat{\boldsymbol{N}} \right\|_2^2. \tag{6}$$

The effect of this additional term can be seen in Fig. 5. While for small values of the regularization parameter γ, the approximation error is very small, the corresponding parameter distribution requires a large defect density. Increasing γ leads to a significant reduction of the defect density, while the error to the measurement data is only weakly affected. However, if γ is too large, the problem becomes overregularized, resulting in a steep error increase. According to the *L-curve* criterion [28] the optimal value for γ lies between these two regimes near the observed "corner." In order to solve (6) efficiently, the regularized estimator can be recast to the standard NNLS form by introducing

$$\boldsymbol{N} = \underset{\hat{\boldsymbol{N}} \geq \boldsymbol{0}}{\arg\min} \left\| \boldsymbol{A} \cdot \hat{\boldsymbol{N}} - \boldsymbol{b} \right\|_2^2 \quad \text{with} \tag{7}$$

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{\delta V} \\ \gamma\, \boldsymbol{I} \end{bmatrix} \quad \text{and} \quad \boldsymbol{b} = \begin{bmatrix} \boldsymbol{\Delta V} \\ \boldsymbol{0} \end{bmatrix} \tag{8}$$

and $\boldsymbol{I}$ being a unit matrix of proper dimensions. In our implementation, the estimator given by (7) is solved iteratively by increasing γ and invoking a fast NNLS algorithm introduced in [29] at each step until the "corner" is detected.

Although the formulation of defect parameter extraction as an NNLS optimization problem allows in principle to search the parameter space without any restrictions, we reintroduce the concept of defect bands by defining oxide regions with homogeneous defect parameters over x_t. This allows us to identify such a band by a simple 2-D $(E_T, \mathcal{E}_R)$ map while also accounting for different chemical environments in the bulk oxide and the IL. In this sense, ESiD can be seen as an extension of (τ_c,τ_e)-maps [30], a technique for extracting the distribution of time constants from repeated measurements with varying stress times. Instead of (τ_c,τ_e), ESiD extracts the more fundamental distribution $(E_T, \mathcal{E}_R)$, which allows to calculate (τ_c,τ_e) under arbitrary conditions.

V. Results

We employ the proposed ESiD extraction algorithm to extract defect parameter distributions for the hole and electron traps in SiON from NBTI measurements on pMOS devices and PBTI on nMOS devices. In order to account for the chemically different environments in the oxide and IL, we define two disjunct and independent search regions for the ESiD extraction. The used parameters for the search regions as well as the extracted total defect densities are summarized in Table I. Based on experimental [31] and theoretical [32] studies on the Si/SiO_2 interface, the IL thickness was assumed to be 0.6 nm. Note that the precise value of the assumed IL thickness is only of minor importance for the resulting defect distribution. The same holds true for the capture cross section σ, which enters the transition rates (1) only linearly, whereas all other parameters have an exponential impact.

In accordance with previous findings [33], our extraction algorithm shows that most of the active defects in both n- and

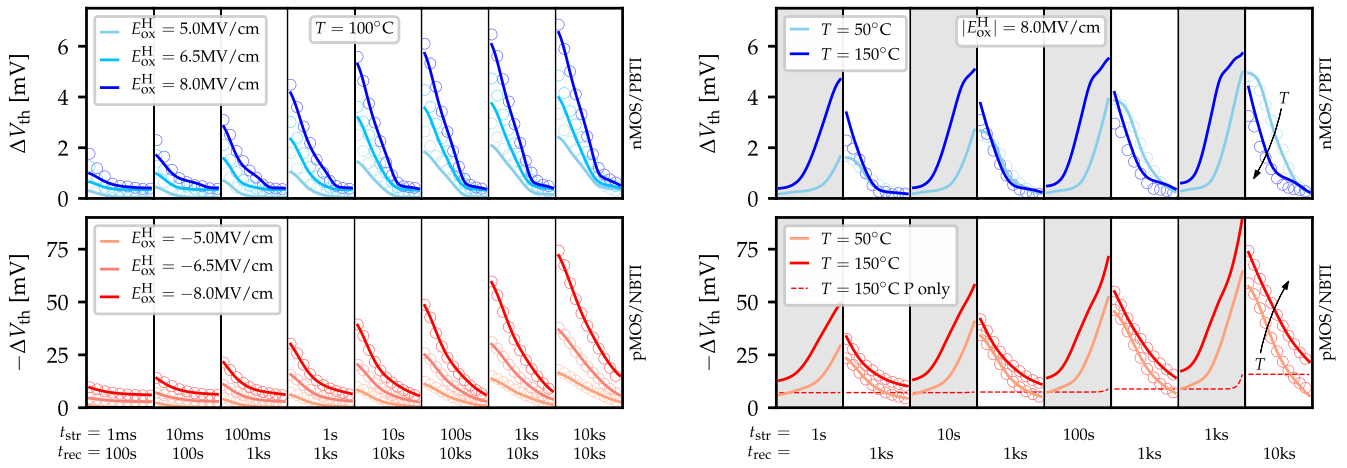

Fig. 7. Experimental and simulated eMSM sequences using the extracted defect bands for varying gate bias (left) and temperature (right). The first data point is measured 100 μs after stress. The extracted (quasi)-permanent component (P) is separately shown as a dashed line for the pMOS in the right plot. As can be seen, there is a very good agreement with the measured ΔV_{th} degradation over the entire measured temperature and bias range. Note the apparent negative temperature activation for PBTI in the right figure due to fast electron traps.

TABLE I
DEFECT BANDS

	E_T [eV]	$\mathcal{E}_R$ [eV]	x_t [nm]	N_T [cm^{-3}]
Hole traps (IL)	-2.0–0.0	0.1–5.0	0.0–0.6	8.0×10^{18}
Hole traps (oxide)	-2.5–0.0	0.1–5.0	0.6–2.5	2.8×10^{18}
Electron traps (IL)	0.0–2.0	0.1–5.0	0.0–0.6	1.4×10^{18}
Electron traps (oxide)	0.0–2.5	0.1–5.0	0.6–2.5	7.0×10^{16}
Grid spacing	0.05	0.125	0.1	–

pMOS are located in the IL, as depicted in Fig. 6. However, the total active defect density in nMOS devices is lower compared to pMOS by a factor of seven in good agreement with earlier investigations [23], [34]. Although these findings have been well established, note that they naturally emerge from simple eMSM recovery traces using our algorithm. Compared to other reports utilizing Gaussian parameter distributions [7]–[9], our results show significantly lower defect densities, ranging from 5 to 15 times less depending on the device type. As already explained in Fig. 1, this might be an artifact of choosing a specific parameter distribution function, where the observed degradation only stems from the distribution tail, leading to an apparent higher defect density. One advantage of our novel extraction approach is that inactive defects within the provided measurement window are automatically discarded since they cannot contribute to the superposition given in (4). However, as shown in Fig. 7, our extracted parameter set describes all recorded MSM traces very well, despite the much lower defect density. As can be seen, both the dependence on the gate bias and the temperature activation of BTI degradation is well captured by our model.

Most notably, the temperature activation in the PBTI/nMOS case shows a peculiar feature due to fast electron traps. As shown in Fig. 7 (right), the measured ΔV_{th} degradation actually decreases with higher temperatures for longer stress times. As pointed out already in [35], simpler BTI characterization schemes like the use of single effective activation energies would predict unphysical negative activation energy for the charge transfer process in this scenario. The degradation during stress is indeed larger for higher temperatures, however, some traps become too fast and already recover during the delay time between stress and recovery phase. These traps, therefore, fall out of the measurement window and cause this apparent decrease in degradation.

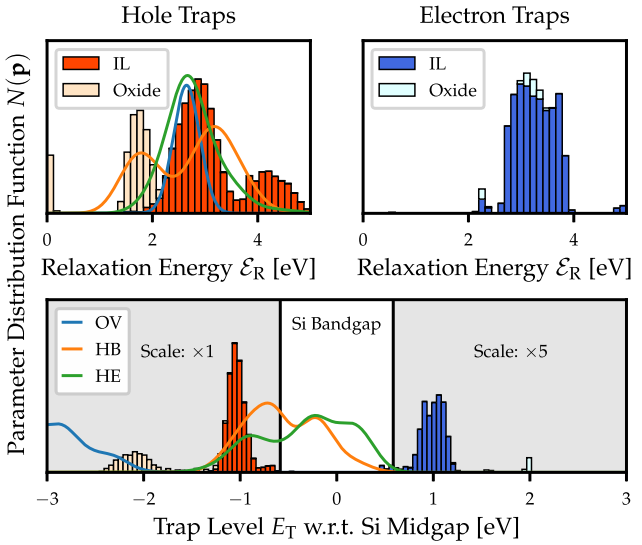

Fig. 8. Extracted defect bands and a comparison to theoretical defect candidates [13] for hole traps in SiO_2. As can be seen, the hydrogen-related defects are in good agreement with our obtained IL hole traps, whereas the OV defect could act as a deeper trap in the oxide.

A. Defect Distributions

Since we have fixed the parameter $R = 1$ and averaged over x_t within a defect band, the defining characteristic of a defect band is its $(E_T, \mathcal{E}_R)$ distribution. These distributions are presented as stacked histograms in Fig. 8 for all four defect bands used during the extraction. As can be seen, most electron and hole traps are located within the IL layer approximately 1.0 eV above and below the Si midgap, respectively. While hole traps are still present in the deeper oxide layer, electron traps are almost entirely absent in this region. This finding explains the more pronounced quasi-permanent BTI component (P) in pMOS devices compared to their nMOS counterparts. Note, however, that P cannot be fully described by charge trapping in preexisting oxide defects alone since other effects like defect creation and hydrogen

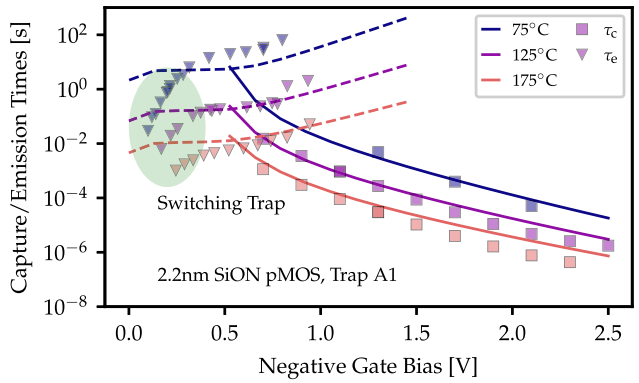

Fig. 9. Application of our Comphy/NMP framework to capture and emission times obtained from single defects using TDDS [38]. Note that most features are well described within the two-state NMP model, however, the switching trap behavior would require additional states to be modeled accurately.

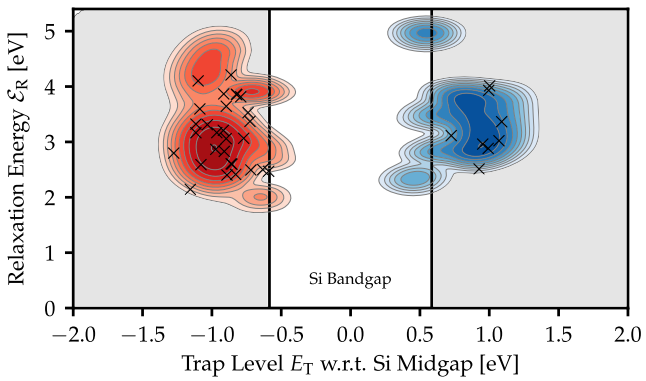

Fig. 10. Heat map of the $(E_T, \mathcal{E}_R)$ distribution for the extracted IL trap bands using our ESiD method together with parameters from single defects obtained with TDDS [38], [40]. As can be seen, both distributions overlap nicely, confirming the physicality of the extracted defect bands.

kinetics also play an important role [36], [37]. The rigorous treatment of these mechanisms is beyond the scope of this work, instead, P is represented by artificial defects within the used NMP framework. In our simulation we consider defects, which can only trap charges but not emit them again within the experimental window ($\tau_e > 5 \times t_{\mathrm{rec}}^{\mathrm{max}}$) even at the highest temperature, to be responsible for P. The collective ΔV_{th} response of these defects is shown as a dashed line in Fig. 7 (right). However, the corresponding defect parameters are unphysical and therefore not shown in Fig. 8.

The recoverable BTI component, however, is well described by the NMP model and its extracted defect parameters allow the identification of likely responsible defect types by comparison to DFT predictions. As shown in Fig. 8, the $(E_T, \mathcal{E}_R)$ distribution of the extracted IL hole traps is in good agreement with theoretical results [13] for the well-known HB defect and the recently popularized hydroxyl-E′ center (HE) [15]. Regarding the energetic location of the defect bands, it has to be kept in mind that E_T predictions commonly show errors of up to ± 0.5 eV within standard DFT. Hence, the shown theoretical distributions should be considered as physically reasonable ranges rather than precise distributions of defects. However, the oxygen vacancy (OV), which has been previously suspected [3] to be responsible for BTI, shows a far too deep trap level to account for most of the observed degradation. Since the OV has been identified by electron spin resonance (ESR) measurements during NBTI stress [39], we speculate that the deeper oxide defects extracted from our data can be attributed to OVs.

B. Comparison to Single Defect Characterizations

In order to validate the defect parameters extracted from eMSM measurements using our ESiD method, we also extracted single defect parameters from earlier datasets [38], [40] on ultrascaled SiON devices otherwise similar to the large-area devices used in this work. Using the same Comphy/NMP modeling framework, we extracted the $(E_T, \mathcal{E}_R)$ parameters from the recorded capture and emission times. As shown exemplarily in Fig. 9, the used two-state NMP model can describe most of the defect characteristics rather well. One exception is the *switching trap* behavior, referring to a drop of the emission time for low gate bias, which requires the introduction of additional metastable defect states [13]. However, this feature is not expected to be important in the context of large-area devices with a large ensemble of defects being involved, provided the devices are not switched into accumulation for pulsed measurements [41]. Fig. 10 shows that the resulting single-defect parameters agree excellently with the extracted parameters for the IL layer from our eMSM measurements. Note that TDDS typically requires a smaller experimental window ($\tau_c, \tau_e \leq 100$ s) compared to eMSM since many repeated recovery traces have to be analyzed in order to obtain sensible mean capture/emission times. Therefore it is expected that TDDS primarily reflects the properties of the faster IL layer instead of the deeper oxide. Our findings show that the ESiD algorithm indeed provides physical defect bands which can be used for the identification of defects and can act to some degree as a suitable substitute for the much more involved single-defect characterization.

VI. Conclusion

In this work, we demonstrated that defect extractions from MSM sequences using BTI compact models can suffer from artificial inactive defects by assuming a specific parameter distribution as well as from ambiguous defect distributions due to parameter cross-correlations within the model. Both effects limit the physical interpretability of the obtained defect bands. Furthermore, the manual optimization of defect bands is an error-prone and laborious task, especially for new materials or advanced gate stacks. To overcome these issues, we proposed a novel semiautomated extraction method, the ESiD, which is based on a regularized NNLS estimator. This method decomposes the measured ΔV_{th} degradation into effective single-defect contributions and provides an optimal distribution function for the defect parameters. We showed that this new approach can accurately model the observed BTI degradation, while only a vaguely defined parameter search region has to be specified by the user. Contrary to previous defect extractions, we showed that our obtained defect bands are physically sound by comparison to DFT predictions and single-defect data obtained with TDDS. Our method is therefore expected to also provide physical defect parameters in cases like SiC-based power devices, which do not meet the requirements to perform TDDS.

REFERENCES

[1] D. K. Schroder, "Negative bias temperature instability: What do we understand?" *Microelectron. Rel.*, vol. 47, no. 6, pp. 841–852, Jun. 2007, doi: 10.1016/j.microrel.2006.10.006.

[2] D. S. Ang, S. Wang, G. A. Du, and Y. Z. Hu, "A consistent deep-level hole trapping model for negative bias temperature instability," *IEEE Trans. Device Mater. Rel.*, vol. 8, no. 1, pp. 22–34, Mar. 2008, doi: 10.1109/TDMR.2007.912275.

[3] J. P. Campbell, P. M. Lenahan, C. J. Cochrane, A. T. Krishnan, and S. Krishnan, "Atomic-scale defects involved in the negative-bias temperature instability," *IEEE Trans. Device Mater. Rel.*, vol. 7, no. 4, pp. 540–557, Dec. 2007, doi: 10.1109/TDMR.2007.911379.

[4] E. Simoen and C. Claeys, "Random telegraph signal: A local probe for single point defect studies in solid-state devices," *Mater. Sci. Eng. B*, vols. 91–92, pp. 136–143, Apr. 2002, doi: 10.1016/S0921-5107(01)00963-1.

[5] T. Grasser, H. Reisinger, P.-J. Wagner, and B. Kaczer, "Time-dependent defect spectroscopy for characterization of border traps in metal-oxide-semiconductor transistors," *Phys. Rev. B, Condens. Matter*, vol. 82, no. 24, Dec. 2010, Art. no. 245318, doi: 10.1103/PhysRevB.82.245318.

[6] B. Kaczer *et al.*, "Ubiquitous relaxation in BTI stressing—New evaluation and insights," in *Proc. IEEE Int. Rel. Phys. Symp.*, Apr. 2008, pp. 20–27, doi: 10.1109/RELPHY.2008.4558858.

[7] G. Rzepa *et al.*, "Comphy—A compact-physics framework for unified modeling of BTI," *Microelectron. Rel.*, vol. 85, pp. 49–65, Jun. 2018, doi: 10.1016/j.microrel.2018.04.002.

[8] G. Rzepa, F. Schanovsky, and M. Karner, "From gate oxide characterization to TCAD predictions: Exploring impact of defects across technologies," in *Proc. Silicon Nanoelectron. Workshop (SNW)*, Jun. 2019, pp. 1–2, doi: 10.23919/SNW.2019.8782971.

[9] B. J. O'Sullivan *et al.*, "Defect profiling in FEFET Si:HfO_2 layers," *Appl. Phys. Lett.*, vol. 117, no. 20, Nov. 2020, Art. no. 203504, doi: 10.1063/5.0029072.

[10] D. Claes, J. Franco, N. Collaert, D. Linten, and M. Heyns, "Positive bias temperature instability of HfO_2-based gate stacks at reduced thermal budget for future CMOS technologies," *J. Appl. Phys.*, vol. 128, no. 10, Sep. 2020, Art. no. 104101, doi: 10.1063/5.0006110.

[11] D. Waldhoer, A.-M. El-Sayed, Y. Wimmer, M. Waltl, and T. Grasser, "Atomistic modeling of oxide defects," in *Noise in Nanoscale Semiconductor Devices*, T. Grasser, Ed. Cham, Switzerland: Springer, 2020, pp. 609–648, doi: 10.1007/978-3-030-37500-3_18.

[12] D. Waldhoer, Y. Wimmer, A. M. El-Sayed, W. Goes, M. Waltl, and T. Grasser, "Minimum energy paths for non-adiabatic charge transitions in oxide defects," in *Proc. IEEE Int. Integr. Rel. Workshop (IIRW)*, Oct. 2019, pp. 1–5, doi: 10.1109/IIRW47491.2019.8989889.

[13] W. Goes *et al.*, "Identification of oxide defects in semiconductor devices: A systematic approach linking DFT to rate equations and experimental evidence," *Microelectron. Rel.*, vol. 87, pp. 286–320, Aug. 2018, doi: 10.1016/j.microrel.2017.12.021.

[14] V. V. Afanasev, M. Bassler, G. Pensl, and M. Schulz, "Intrinsic SiC/SiO_2 interface states," *Phys. Status Solidi (A)*, vol. 162, no. 1, pp. 321–337, 1997, doi: 10.1002/1521-396X(199707)162:1<321::AID-PSSA321>3.0.CO;2-F.

[15] A.-M. El-Sayed, M. B. Watkins, T. Grasser, V. V. Afanas'ev, and A. L. Shluger, "Hydrogen-induced rupture of strained Si–O bonds in amorphous silicon dioxide," *Phys. Rev. Lett.*, vol. 114, p. 115503, Mar. 2015, doi: 10.1103/PhysRevLett.114.115503.

[16] P. E. Blöchl and J. H. Stathis, "Hydrogen electrochemistry and stress-induced leakage current in silica," *Phys. Rev. Lett.*, vol. 83, no. 2, pp. 372–375, Jul. 1999, doi: 10.1103/PhysRevLett.83.372.

[17] M. Waltl, "Ultra-low noise defect probing instrument for defect spectroscopy of MOS transistors," *IEEE Trans. Device Mater. Rel.*, vol. 20, no. 2, pp. 242–250, Jun. 2020, doi: 10.1109/TDMR.2020.2988650.

[18] S. M. Sze, *Physics of Semiconductor Devices*. Hoboken, NJ, USA: Wiley, 2007, p. 200, doi: 10.1002/0470068329.

[19] T. Grasser, "Stochastic charge trapping in oxides: From random telegraph noise to bias temperature instabilities," *Microelectron. Rel.*, vol. 52, no. 1, pp. 39–70, Jan. 2012, doi: 10.1016/j.microrel.2011.09.002.

[20] T. Tewksbury, "Relaxation effects in MOS devices due to tunnel exchange with near-interface oxide traps," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, USA, 1992. [Online]. Available: http://dspace.mit.edu/handle/1721.1/13238

[21] B. Ruch, M. Jech, G. Pobegen, and T. Grasser, "Applicability of Shockley–Read–Hall theory for interface states," *IEEE Trans. Electron Devices*, vol. 68, no. 4, pp. 2092–2097, Apr. 2021, doi: 10.1109/TED.2021.3049760.

[22] K. Huang, A. Rhys, and N. F. Mott, "Theory of light absorption and non-radiative transitions in *F*-centres," *Proc. Roy. Soc. London A, Math. Phys. Sci.*, vol. 204, no. 1078, pp. 406–423, 1950, doi: 10.1098/rspa.1950.0184.

[23] V. Huard, M. Denais, and C. Parthasarathy, "NBTI degradation: From physical mechanisms to modelling," *Microelectron. Rel.*, vol. 46, no. 1, pp. 1–23, Jan. 2006, doi: 10.1016/j.microrel.2005.02.001.

[24] V. Huard, "Two independent components modeling for negative bias temperature instability," in *Proc. IEEE Int. Rel. Phys. Symp.*, May 2010, pp. 33–42, doi: 10.1109/IRPS.2010.5488857.

[25] T. Grasser, H. Reisinger, P.-J. Wagner, F. Schanovsky, W. Goes, and B. Kaczer, "The time dependent defect spectroscopy (TDDS) for the characterization of the bias temperature instability," in *Proc. IEEE Int. Rel. Phys. Symp.*, May 2010, pp. 16–25, doi: 10.1109/IRPS.2010.5488859.

[26] C. S. Kelley, "Moments of semiclassical and classical absorption and emission band shapes of impurities in solids," *Phys. Rev. B, Condens. Matter*, vol. 20, no. 12, pp. 5084–5089, Dec. 1979, doi: 10.1103/PhysRevB.20.5084.

[27] A. N. Tikhonov, *Numerical Methods for the Solution of Ill-Posed Problems*. Dordrecht, The Netherlands: Springer, 1995, doi: 10.1007/978-94-015-8480-7.

[28] D. Calvetti, S. Morigi, L. Reichel, and F. Sgallari, "Tikhonov regularization and the L-curve for large discrete ill-posed problems," *J. Comput. Appl. Math.*, vol. 123, no. 1, pp. 423–446, 2000, doi: 10.1016/S0377-0427(00)00414-3.

[29] R. Bro and S. D. Jong, "A fast non-negativity-constrained least squares algorithm," *J. Chemometrics*, vol. 11, no. 5, pp. 393–401, 1997, doi: 10.1002/(SICI)1099-128X(199709/10)11:5<393::AID-CEM483>3.0.CO;2-L.

[30] T. Grasser *et al.*, "Analytic modeling of the bias temperature instability using capture/emission time maps," in *IEDM Tech. Dig.*, Dec. 2011, pp. 27.4.1–27.4.4, doi: 10.1109/IEDM.2011.6131624.

[31] D. A. Muller, T. Sorsch, S. Moccio, F. H. Baumann, K. Evans-Lutterodt, and G. Timp, "The electronic structure at the atomic scale of ultrathin gate oxides," *Nature*, vol. 399, no. 6738, pp. 758–761, Jun. 1999, doi: 10.1038/21602.

[32] Y. Yamashita *et al.*, "Direct observation of site-specific valence electronic structure at the SiO_2/Si interface," *Phys. Rev. B, Condens. Matter*, vol. 73, no. 4, Jan. 2006, Art. no. 045336, doi: 10.1103/PhysRevB.73.045336.

[33] Y. Maneglia and D. Bauza, "Extraction of slow oxide trap concentration profiles in metal–oxide–semiconductor transistors using the charge pumping method," *J. Appl. Phys.*, vol. 79, no. 8, pp. 4187–4192, 1996, doi: 10.1063/1.361786.

[34] J. P. Campbell, P. M. Lenahan, A. T. Krishnan, and S. Krishnan, "Location, structure, and density of states of NBTI-induced defects in plasma nitrided pMOSFETs," in *Proc. IEEE Int. Rel. Phys. Symp. 45th Annu.*, Apr. 2007, pp. 503–510, doi: 10.1109/RELPHY.2007.369942.

[35] M. W. Feil, K. Puschkarsky, W. Gustin, H. Reisinger, and T. Grasser, "On the physical meaning of single-value activation energies for BTI in Si and SiC MOSFET devices," *IEEE Trans. Electron Devices*, vol. 68, no. 1, pp. 236–243, Jan. 2021, doi: 10.1109/TED.2020.3036321.

[36] T. Grasser *et al.*, "Gate-sided hydrogen release as the origin of 'permanent' NBTI degradation: From single defects to lifetimes," in *IEDM Tech. Dig.*, Dec. 2015, pp. 20.1.1–20.1.4, doi: 10.1109/IEDM.2015.7409739.

[37] T. Grasser *et al.*, "Implications of gate-sided hydrogen release for post-stress degradation build-up after BTI stress," in *Proc. IEEE Int. Rel. Phys. Symp. (IRPS)*, Apr. 2017, pp. 6A-2.1–6A-2.6, doi: 10.1109/IRPS.2017.7936334.

[38] T. Grasser *et al.*, "On the microscopic structure of hole traps in pMOSFETs," in *IEDM Tech. Dig.*, Dec. 2014, pp. 21.1.1–21.1.4, doi: 10.1109/IEDM.2014.7047093.

[39] J. T. Ryan, P. M. Lenahan, T. Grasser, and H. Enichlmair, "Observations of negative bias temperature instability defect generation via on the fly electron spin resonance," *Appl. Phys. Lett.*, vol. 96, no. 22, May 2010, Art. no. 223509, doi: 10.1063/1.3428783.

[40] M. Waltl, B. Stampfer, G. Rzepa, B. Kaczer, and T. Grasser, "Separation of electron and hole trapping components of PBTI in SiON nMOS transistors," *Microelectron. Rel.*, vol. 114, Nov. 2020, Art. no. 113746, doi: 10.1016/j.microrel.2020.113746.

[41] T. Grasser *et al.*, "Advanced characterization of oxide traps: The dynamic time-dependent defect spectroscopy," in *Proc. IEEE Int. Rel. Phys. Symp. (IRPS)*, Apr. 2013, pp. 2D.2.1–2D.2.7, doi: 10.1109/IRPS.2013.6531957.