Modelling of Charge Trapping/De-trapping Induced Voltage Instability in High-k Gate Dielectrics

Sahar Sahhaf, Kris De Brabanter, Robin Degraeve, Johan A. K. Suykens, Bart De Moor and Guido Groeseneken

***Abstract*—Investigation of trapping/de-trapping induced voltage-instabilities does not only demand accurate measurements but also establishing a precise methodology for extracting the exact magnitude of the voltage-shifts in the hysteresis curves is indispensable. Particularly, in DC measurements where the induced voltage-shifts are small, an excellent accuracy of the analysis method is required. Therefore, in this paper we develop a new methodology that with excellent agreement models the complete measured I_d-V_g hysteresis curves using least squares support vector machines (LS-SVM). Further, we apply this model and formulate an optimization problem resulting in the maximum trap-induced voltage-shifts in the entire hysteresis curves. Also for the first time, we quantify the induced error on these extracted maxima. Finally, we illustrate the applicability of the introduced methodology by profiling the initially present and stress-induced defects in a 1 nm SiO_2/3 nm HfSiO dielectric stack.**

***Index Terms*—trapping and de-trapping induced voltage-instability, maximum voltage-shift, oxide defects, Hf-based dielectric.**

I. Introduction

HIGH-K -based dielectrics (e.g. Hf-based) are shown to be potential candidates to replace SiO_2 as gate dielectric [1]. However, one of the most challenging problems of these high-k materials is their large amount of defects [2] which can lead to trapping and de-trapping induced voltage-instabilities. To investigate the build up or loss of charge, often voltage shifts in the I_d-V_g hysteresis curves are studied. For such a study, one should not only perform a correct measurement but also the necessity of an accurate analysis of the measured data is of equal importance. There are many ameliorations of hysteresis measurement techniques (e.g. [3]–[5]). However, not many publications dealing with the accuracy of the analysis techniques can be found.

Different models have been proposed trying to translate the change in the drain current to a threshold voltage-shift. But each model has its own drawbacks. An example is the compact model (three-parameter equation from SPICE) [6] applied in OTF (on the fly technique) which cannot describe the turning point of the current near the threshold. We illustrate this in Section II.

S. Sahhaf, R. Degraeve and G. Groeseneken are with IMEC, Kapeldreef 75, B-3001 Leuven, Belgium

Also S. Sahhaf and G. Groeseneken are with the Department of Electrical Engineering (ESAT), Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium

K. De Brabanter, J. Suykens and B. De Moor are with the Department of Electrical Engineering (ESAT) Research Division SCD, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium

Some techniques [4], [7], [8] estimate the ΔV as a shift between I_d-V_g traces measured during the rise and fall of the pulse bias. Note that the approximation applied in [4] to estimate the ΔV is valid when certain conditions are met e.g. no significant mobility decrease due to a high vertical field in the channel caused by high gate bias.

One should keep in mind that first, the inaccuracies of the ΔV extraction procedures can add up to the errors introduced during the measurement (e.g. propagation delay in the cables due to an inaccurate measurement configuration or parasitic components of the device and etc... [4], [9]). Second, in slow ramp measurements, due to the pronounced de-trapping compared to fast measurements (μs ramp rates as e.g. in pulsed I_d-V_g [3]), the magnitudes of the hysteresis (ΔV) are very small and therefore large analysis-induced errors are not tolerated. Based on these two reasonings, we define the goal of this work as reducing the errors on the extracted ΔV values by increasing the accuracy of the analysis procedure.

The paper is laid out as follows: In Section II we introduce a modelling technique based on least squares support vector machines (LS-SVM) [10]. This technique is applied to model the I_d-V_g hysteresis and we show that LS-SVM is capable of detecting the nonlinear form of the hysteresis also in the curved parts (around threshold). Based on this underlying model, we formulate a constrained optimization problem resulting in the extraction of maximum voltage-shift in the entire hysteresis curve. In Section III, we use the methodology to study the spatial and energy profile of traps in 1 nm SiO_2/3 nm HfSiO. In this study we investigate charge trapping monitored by I_d-V_g characteristics where we deliberately choose a DC hysteresis measurement. The challenge is to find the maximum voltage-shift knowing that the magnitude of the shifts are extremely small. In this way, we demonstrate the accuracy of the new methodology. Finally, in Section IV we conclude the paper.

II. Modelling the Hysteresis Curve

In this section we describe the principles of least squares support vector machines (LS-SVM) [10] used in our proposed methodology. LS-SVM allows to handle complex and unknown model structures (not known a priori) in the data which can be an obstacle for classical parametric models since they require a pre-specified structure.

A. Least Squares Support Vector Machines

When the underlying physical model (to be estimated) is complex and unknown to the user, a technique called

least squares support vector machines (LS-SVM) [10] can be applied. The full derivation of the method is beyond the scope of this paper, therefore we only provide the basic formulation. Further information regarding this topic can be found in [10].

Given a data set $\{(X_k, Y_k), k = 1, \ldots, n\}$, then the estimated LS-SVM model is given by

$$y(x) = \sum_{k=1}^{n} \alpha_k K(x, X_k) + b, \tag{1}$$

with $K(X_k, X_l) = \exp(-(X_k - X_l)^2/\sigma^2)$ and σ a tuning parameter called the bandwidth of the kernel. The values for α en b are found by solving the linear system

$$\left[\begin{array}{c|c} 0 & 1_n^T \\ \hline 1_n & \Omega + \frac{1}{\gamma} I_n \end{array}\right] \left[\begin{array}{c} b \\ \hline \alpha \end{array}\right] = \left[\begin{array}{c} 0 \\ \hline Y \end{array}\right], \tag{2}$$

with $Y = (Y_1, \ldots, Y_n)^T$ the outputs, $1_n = (1, \ldots, 1)^T$, Ω a positive definite (symmetric) matrix of the following form

$$\Omega = \left[\begin{array}{ccc} K(X_1, X_1) & \cdots & K(X_1, X_n) \\ \vdots & \ddots & \vdots \\ K(X_n, X_1) & \cdots & K(X_n, X_n) \end{array}\right].$$

The tuning parameters σ (bandwidth of the kernel) and regularization constant γ in the linear system (2) can be found by cross-validation [11]. Most commonly used is leave-one-out cross-validation (LOO-CV). In LOO-CV, the original sample is randomly partitioned into n subsamples. Of the n subsamples, $n-1$ subsamples are used as training data, i.e. solve the system (2) to obtain α and b. The single subsample, not used as training data, is employed to test the model (i.e. evaluate the trained LS-SVM model on this subsample, see (1)). The cross-validation process is then repeated n times (folds), with each of the n subsamples used exactly once as validation data. The n results from the folds are then averaged to produce a single estimation. The lowest value according this criterion gives the optimal σ and γ. Fig. 1 illustrates the principle of LOO-CV.

The applicability of the LS-SVM technique is widely extended and is capable of detecting underlying linear/non-linear models regardless of the form of the data. We have used the StatLSSVM toolbox [12] to model the hysteresis curves. Finally, we summarize the construction of the LS-SVM model in Algorithm 1.

Algorithm 1 Construction of an LS-SVM model

1: Given a data set $\{(X_k, Y_k), k = 1, \ldots, n\}$, perform LOO-CV to obtain optimal σ and γ
2: Solve the linear system (2) given the optimal σ and γ
3: Evaluate the LS-SVM model (1) in the points of interest.

As an illustration, we apply LS-SVM to model the I_d-V_g hysteresis curve (up trace) obtained from a DC hysteresis measurement (raw data) as shown in Fig. 2. For the comparison, the data is also modelled using a simple compact model (SPICE level-1) [6] which is the basis of the on-the-fly technique (OTF) and is formulated as

$$I_d = \beta \frac{V_g - V_T}{1 + \theta(V_g - V_T)}, \tag{3}$$

with I_d the drain current, β depending on the effective mobility and θ describing the curvature of the I_d-V_g curve. We remark that in (3), to obtain a proper fit, only data where $V_g > V_T$ should be used. Note that the inaccuracy of this model increases around the turning point of the hysteresis near the threshold. It is obvious that the advantage of the LS-SVM model is that it indeed describes the full hysteresis including the curved parts (around threshold) with a high level of accuracy.

B. Extracting the Maximum voltage-Shift in the Entire Hysteresis Curve

In this section we explain the procedure to find the maximum voltage-shift ($\Delta V = V_2 - V_1$) in the hysteresis based on a constraint optimization problem using the LS-SVM models describing the up and down I_d-V_g traces.

We explain this problem with the graphical illustration shown in Fig. 3. Denote by f and g, the estimated LS-SVM models for the up and down traces respectively with the given data. The problem of finding the maximum shift ΔV is formulated as follows

$$\left\{\begin{array}{l} \max\limits_{V_1, V_2} \Delta V = V_2 - V_1 \\ \\ \text{s.t.} \quad f(V_1) = g(V_2) \end{array}\right. \tag{4}$$

The maximization problem (4) can be efficiently solved by using TOMLAB optimization software available at http://tomopt.com/tomlab/.

We apply the technique explained above to the real DC measured data. We remark that the applicability of this approach is not limited to a certain regime (linear or sub-threshold regime), but we search the maximum voltage-shift in the entire I_d-V_g hysteresis without loosing accuracy in the curved part of the hysteresis near the threshold.

Fig. 4 shows the extracted maximum hysteresis for I_d-V_g sweeps with increasing top V_g obtained by solving the optimization problem (4). For the comparison, we have also extracted ΔV using fixed current thresholds. As illustrated, the extracted ΔV values depend on the choice of the threshold current level. It is obvious that the real maximum ΔV is found only by solving (4).

For the first time, the errors on the extracted ΔV values from I_d-V_g sweeps with increasing top V_g (measured on fresh and stressed devices) are quantified as shown in Fig. 5. Note that each of these shown errors consists of two components: 1) error introduced by the LS-SVM model, 2) error caused by the optimization algorithm. This again clarifies the necessity of a model which estimates the underlying form of the data with a low error.

III. Traps Profiling Based on I_d-V_g Hysteresis Measurements

To study the defect profile in 1 nm SiO_2/3 nm HfSiO, we investigate charge trapping monitored by I_d-V_g characteristics with $V_d = 0.1V$ when sweeping from accumulation towards inversion and back, while successively increasing the maximum gate voltage $V_{g,\max}$. We apply sufficiently negative gate bias i.e. $-1V$ to obtain complete charge recovery after each sweep. Instead of measuring I_d-V_g as fast as possible to ensure maximum hysteresis, we deliberately choose a "slow" sweep (20 ms at each V_g-step of $0.02V$ using an Agilent 4156) and we wait 1s at $V_{\text{g,max}}$ before measuring the down ramp of the sweep. Due to these long trapping times, also defects which are located spatially deep in the bulk of the stack can trap electrons.
In such a slow ramp measurement, de-trapping during the down trace is much more pronounced compared to fast measurements with μs ramp rates as e.g. in pulsed I_d-V_g [3] and surely the traps in the interfacial layer and at the SiO_2/high-k interface are already de-trapped and therefore not detectable.

Instead, it is guaranteed that the measured hysteresis is caused by defects present in the central part of the high-k layer. Importantly, the magnitudes of the remaining hysteresis (ΔV) are very small. As the ΔV values are small, it is important to use an accurate procedure for extracting the ΔV in order to reduce the analysis-induced errors as much as possible.

All measured I_d-V_g sweeps (on fresh samples and after constant voltage stress (CVS)) with increasing top V_g are modelled by LS-SVM (explained in Section II) and further the maximum voltage-shifts are extracted from the developed optimization procedure (see (4)). The maximum hysteresis (ΔV) for slow I_d-V_g sweeps with increasing $V_{g,\max}$, is a measure for trap density in the central part of the high-k (as explained before) and is shown in Fig. 6 for fresh and stressed devices.

The fresh curve reveals a continuous increase of the hysteresis and therefore of the trap density in the high V_g range (shallow energy level range [13] indicating the existence of an initial defect band in this range. After stress, the maximum trap density-increase (corresponding with maximum change in tangent-slope between fresh and post-stress) is observed at low V_g range (deep energy levels [13]). Also note that after stress, the hysteresis curves indicate no significant increase in trap density at high V_g range (shallow energy levels), since the tangent-slopes are almost parallel.

IV. Conclusion

In this work, we proposed a new methodology for characterizing the trap-induced voltage-shifts in the I_d-V_g hysteresis curves. The goal of this work was to reduce the inaccuracies of the extracted ΔV as 1) they can escalate the measurement errors and 2) in case of slow-rate measurements, due to the small magnitude of the voltage-shifts, analysis-induced errors are not tolerated. In order to achieve this, we used a modelling technique called least squares support vector machines (LS-SVM) capable of handling complex and a priori unknown model structures in the data with a quantified low error.

We applied LS-SVM and modelled the entire I_d-V_g hysteresis curves without loosing accuracy in the curved part around V_T. Based on this model, we formulated a constrained optimization problem resulting in very accurate extraction of maximum hysteresis.

In the final step, we used our methodology to investigate the defects profile in a 1 nm SiO_2/3 nm HfSiO stack. As the magnitude of the voltage-shifts are small, the necessity of an extraction procedure with high accuracy was again emphasized. In summary, we have identified traps present in the center part of the high-k layer. We profiled these traps energetically: On fresh samples, the defect density is constant (not zero) at deep energy levels and increases in the shallow energy range, while traps generated by CVS reach a maximum density at deep energy levels.

Acknowledgment

The authors would like to thank Ph.J. Roussel for useful discussions. Also This work is part of IMEC's Industrial Affiliation Program, funded by IMEC's core partners: Intel, Texas Instruments, Micron, Infineon, NXP, ST, Matsushita, TSMC, Samsung and Elpida. Support from IMEC's p-line for processing and AMSIMEC for electrical characterization is greatly acknowledged. Research supported by Onderzoeksfonds K.U.Leuven/Research Council KUL: GOA/11/05 Ambiorics, GOA/10/09 MaNet, CoE EF/05/006 Optimization in Engineering (OPTEC) en PFV/10/002 (OPTEC), IOF-SCORES4CHEM, several PhD/post-doc & fellow grants; Flemish Government: FWO: FWO: PhD/postdoc grants, projects: G0226.06 (cooperative systems and optimization), G0321.06 (Tensors), G.0302.07 (SVM/Kernel), G.0320.08 (convex MPC), G.0558.08 (Robust MHE), G.0557.08 (Glycemia2), G.0588.09 (Brain-machine) research communities (WOG: ICCoS, ANMMM,MLDM); G.0377.09 (Mechatronics MPC), IWT: PhD Grants, Eureka-Flite+, SBO LeCoPro, SBO Climaqs, SBO POM, O&O-Dsquare, Belgian Federal Science Policy Office: IUAP P6/04 (DYSCO, Dynamical systems, control and optimization, 2007-2011), IBBT, EU: ERNSI; FP7-HD-MPC (INFSO-ICT-223854), COST intelliCIS, FP7-EMBOCON (ICT-248940), Contract Research: AMINAL, Other: Helmholtz, viCERP, ACCM. BDM is a full professor at the Katholieke Universiteit Leuven, Belgium. JS is a professor at the Katholieke Universiteit Leuven, Belgium.

References

[1] Y. Kim, G. Gebara, and M. Freiler *et al.*, "Conventional n-channel MOSFET devices using single layer HfO_2 and ZrO_2 as high-k gate dielectrics with polysilicon gate electrode," *IEDM Techn. Dig.*, pp. 455–458, 2001.

[2] A. Kerber, E. Cartier, L. Pantisano, R. Degraeve, T. Kauerauf, Y. Kim, A. Hou, G. Groeseneken, H. Maes, and U. Schwalke, "Origin of the threshold voltage instability in SiO_2/HfO_2 dual layer gate dielectrics," *IEEE Electron Device Lett.*, vol. 24, no. 2, pp. 87–89, 2003.

[3] A. Kerber, E. Cartier, L. Pantisano, M. Rosmeulen, R. Degraeve, T. Kauerauf, G. Groeseneken, H. Maes, and U. Schwalke, "Characterization of the V_t-instability in SiO_2 HFO_2 gate dielectrics," *Proceedings 41st Annual IEEE International Reliability Physics Symposium*, pp. 41–45, 2003.

[4] C. D. Young, Y. Zhao, D. Heh, R. Choi, B. H. Lee, and G. Bersuker, "Pulsed I_d-V_g methodology and its application to electron-trapping characterization and defect density profiling," *IEEE Transactions on electron devices*, vol. 56, no. 6, pp. 1322–1329, June 2009.

[5] B. Kaczer, V. Arkhipov, R. Degraeve, N. Collaert, G. Groeseneken, and M. Goodwin, "Disorder-controlled-kinetics model for negative bias temperature instability and its experimental verification," *IRPS*, pp. 381–387, 2005.

[6] H. Reisinger, U. Brunner, W. Heinrigs, W. Gustin, and C. Schlünder, "A comparison of fast methods for measuring NBTI degradation," *IEEE Transactions on Device and Materials Reliability*, vol. 7, no. 4, pp. 531–539, December 2007.

[7] C. Leroux, J. Mitard, G. Ghibaudo, X. Garros, G. Reimbold, B. Guillaumor, and F. Martin, "Characterization and modeling of hysteresis phenomena in high dielectrics," *IEDM Tech. Dig.*, pp. 737–740, 2004.

[8] G. Bersuker, J. P. Zeitzoff, B. H. Lee, R. Choi, G. Brown, and C. D. Young, "Mobility evaluation in transistors with charge-trapping gate dielectrics," *Appl. Phys. Lett.*, vol. 87, no. 4, p. 042905, July 2005.

[9] C. Shen, M.-F. Li, X. P. Wang, Y.-C. Yeo, and D.-L. Kwong, "A fast measurement technique of MOSFET I_d-V_g characteristics," *IEEE Electron Device Lett.*, vol. 27, no. 1, pp. 55–57, January 2006.

[10] J. A. K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, *Least Squares Support Vector Machines*. Singapore: World Scientific, 2002.

[11] P. Burman, "A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing methods," *Biometrika*, vol. 76, no. 3, pp. 503–514, 1989.

[12] K. De Brabanter, J. A. K. Suykens, and B. De Moor, "Nonparametric regression via StatLSSVM," http://www.esat.kuleuven.be/sista/statlssvm, K.U.Leuven, Tech. Rep. 11-134, 2011.

[13] S. Sahhaf, R. Degraeve, M. Cho, K. De Brabanter, Ph. J. Roussel, M. B. Zahid, and G. Groeseneken, "Detailed analysis of charge pumping and I_d-V_g hysteresis for profiling traps in SiO_2/HfSiO(N)," *Microelectronic Engineering*, vol. 87, no. 12, pp. 2614–2619, 2010.

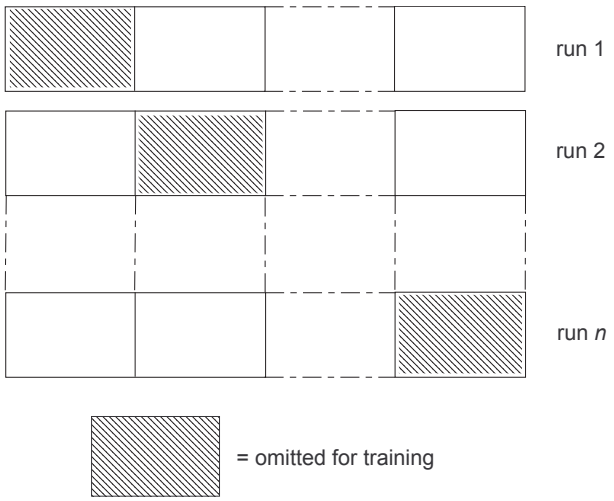


Fig. 1.

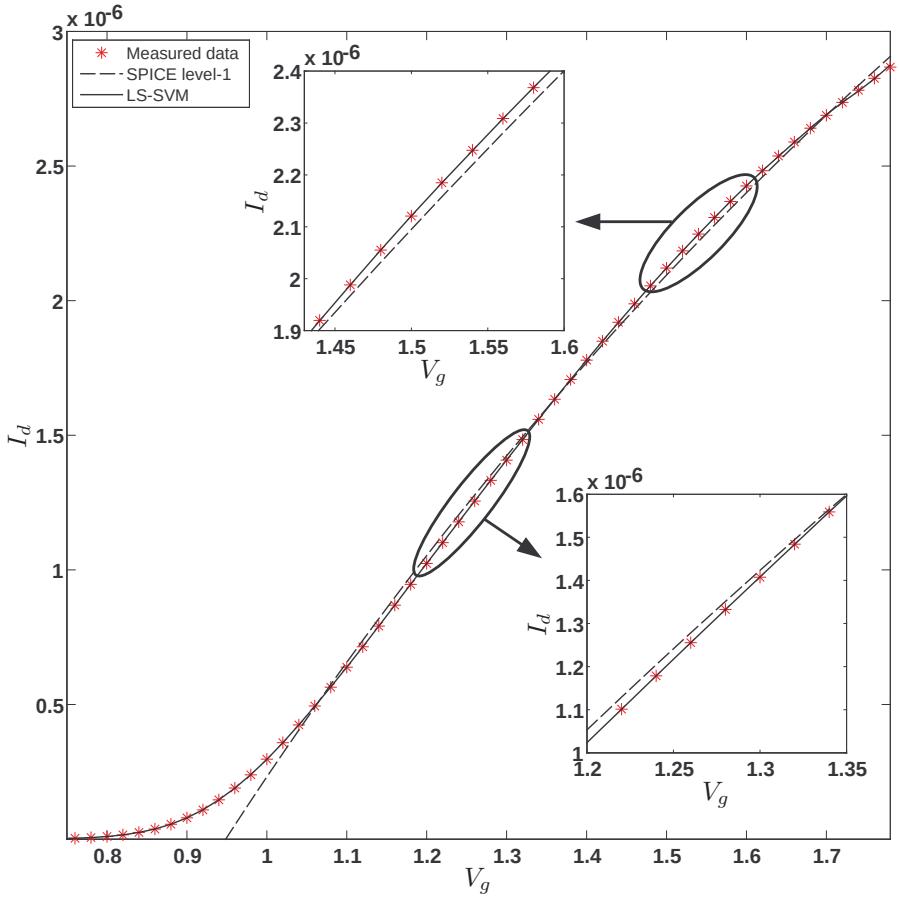


Fig. 2.

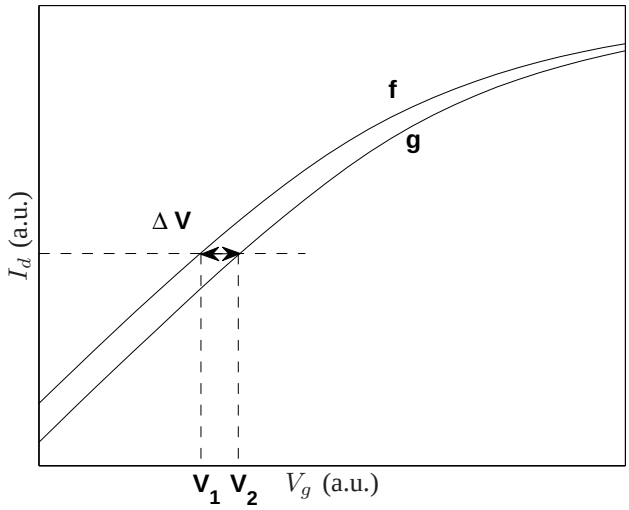


Fig. 3.

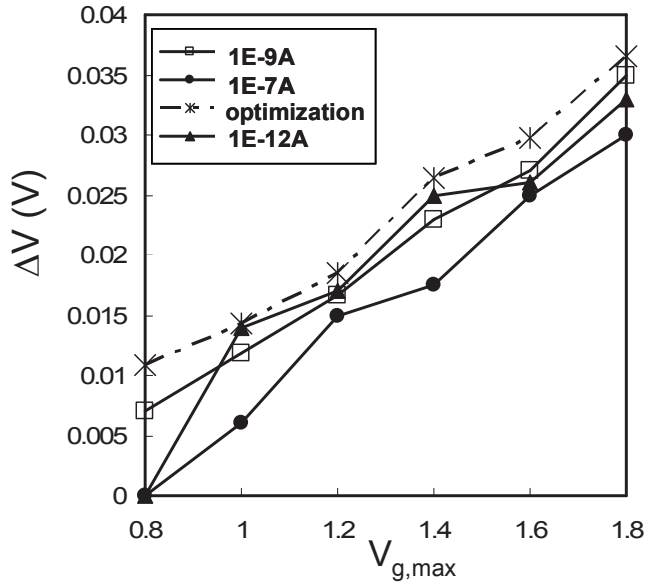

0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
ΔV (V)
1E-9A
1E-7A
optimization
1E-12A
0.8
1
1.2
1.4
1.6
1.8
V g,max

Fig. 4.

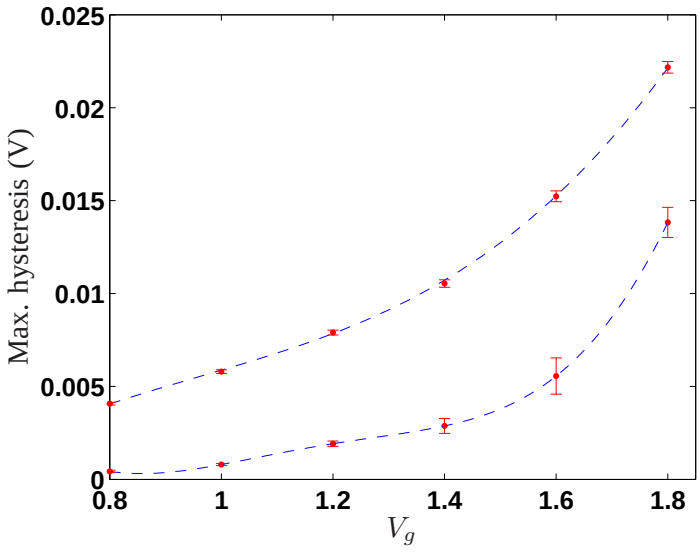


Fig. 5.

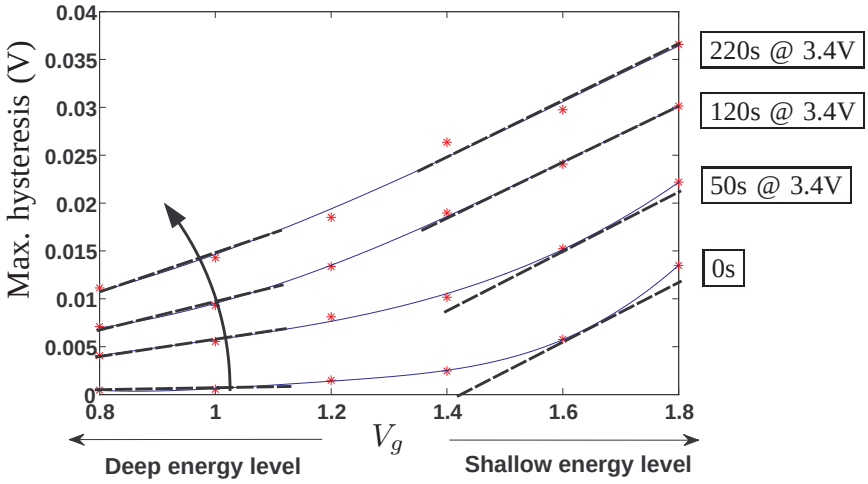

Max. hysteresis (V)
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0.8
1
1.2
1.4
1.6
1.8
V_g
Deep energy level
Shallow energy level
220s @ 3.4V
120s @ 3.4V
50s @ 3.4V
0s

Fig. 6.

Fig. 1 In leave-one-out cross-validation n runs are done where in each run a different part of the data set is omitted. The n results from the folds are then averaged to produce a single estimation. The lowest value according to this criterion gives the optimal σ and γ.

Fig. 2 LS-SVM estimate (full line) and Spice level-1 estimate (dashed line) of the up trace of the hysteresis. Two sections of the estimated models are enlarged to clearly visualize the distinction between LS-SVM and Spice level-1. Spice level-1 cannot describe data near the threshold.

Fig. 3 Illustration of the maximum voltage-shift in a hysteresis curve. ΔV denotes the maximum shift of the hysteresis curves, f and g represent the resulting LS-SVM models of the up and down traces.

Fig. 4 Extracted hysteresis by solving the optimization problem (4) to obtain the maximum hysteresis, and by measuring ΔV at fixed current thresholds as indicated in the legend.

Fig. 5 Extracted maximum hysteresis for I_d-V_g sweeps (fresh and stressed devices) with increasing top V_g with corresponding errors. The errors consists of two components i.e. error introduced by the LS-SVM model and error caused by the optimization algorithm.

Fig. 6 Extracted maximum hysteresis from I_d-V_g sweeps (using LS-SVM and optimization procedure) with increasing top V_g in fresh and stressed devices. The fresh curve reveals a continuous increase of voltage-shift in the high voltage range indicating the existence of an initial defect band in energetically shallow levels. After stress, the maximum trap density-increase (corresponding with maximum change in tangent-slope between fresh and post-stress) is observed at low voltage range corresponding with energetically deep levels.