Electrical methods for estimating the correlation length of insulator thickness fluctuations in MIS tunnel structures

S.E. Tyaginov a,b,*, M.I. Vexler b,c, A. El Hdiy d, K. Gacement, V. Zaporojtchenko e

a Institut für Mikroelektronik, TU Vienna, Gußhausstraße 27-29, A-1040 Vienna, Austria
b Ioffe Physical Technical Institute, Polytechnicheskaya Str. 26, 194021 St.-Petersburg, Russia
c Institut für Elektronische Bauelemente und Schaltungstechnik, TU Braunschweig, PF 3329, D-38023 Braunschweig, Germany
d LMEN(EA), UFR Sciences BP 1039, F-51687 Reims cedex 2, France
e Technische Fakultät der Universität Kiel, Kaiserstraße 2, D-24143 Kiel, Germany

A R T I C L E   I N F O
Available online 31 August 2011

Keywords:
 Silicon dioxide
 Thin insulating film
 MIS tunnel structure
 Thickness fluctuation
 Current–voltage characteristic
 Soft breakdown
 Gate oxides
 Alternative dielectrics
 Silicon
 Aluminium

A B S T R A C T

The possibilities of experimental extraction of the correlation length of insulator thickness fluctuations from the data of electrical measurements on thin metal-insulator-semiconductor (MIS) structures are discussed. The procedure of statistical treatment of currents flowing in a random selection of MIS tunnel diodes is developed enabling the estimation of such a length. Another proposed technique is based on the quantitative analysis of soft-breakdown-related current jumps down occurring under high-voltage stress. The novel methods were tested using Al/SiO2/Si structures and shown to yield the value of correlation length close to that given by a straightforward ”covariant” method applied to the thickness profiles of the same oxide films.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Tunnel-thin films of different dielectric materials attract now more attention than ever before. To a great extent, although not exclusively, such a wave of interest was raised by development of MOSFET-based electronics. This development relies on scaling the linear dimensions L of individual Metal-Oxide-Semiconductor Field-Effect Transistors, which is known to imply also an appropriate reduction of the gate oxide thickness [1]. Evidently, the smaller this thickness is, the more critical are its fluctuations along the structure, especially if we consider the tunnel currents. In the last decade, the complex of problems related to insulator thickness variations in ultra-small MOSFETs and in MIS tunnel devices have been concretized in the literature (e.g. [2–5]). One of the most important tasks in this field is certainly the experimental diagnostics of fluctuation parameters.

Usually, for a local film thickness d, Gauss distribution with the nominal thickness value d n and the standard deviation s d is adopted [4–6]. However, spatial fluctuations of thickness cannot occur arbitrarily abruptly. To account for this, an additional parameter – the correlation length of insulator thickness variations l – is introduced, which is to be understood as a minimal distance between two points where the local thicknesses are considered independent. A more precise definition of l will be given below. The length l, like s d, is an overall quality indicator for the dielectric layer. Ideally, l should be very short, so that the condition L > l could be satisfied diminishing the spread between the device curves, whatever s d is.

In this paper, we propose two methods to estimate the parameter l from the electrical measurements. One of
them uses the experimental data on the statistical current spread, together with the dependencies for the standard deviation of current vs. the $I/\lambda$ ratio calculated in our work [7]. The second method is based on the effect of current drop during the stress in MIS structures with a finite $\sigma_d$. Such a drop eventually occurs under a high stressing bias [8] and is assumed to arise from the current offset in the thinnest device fraction of a size $\sim \lambda^2$. These methods are supposed to be applicable for any tunnel-thin dielectric films, not only and even not primarily to the gate oxides in MOSFETs.

After a brief sample description, we first discuss the definition of the correlation length $\lambda$ and then present and verify the above-mentioned methods for its estimating. It should be emphasised that our goal is just to demonstrate the novel techniques for studying the thickness fluctuations, and not to diagnose the quality of specific devices.

2. Samples

In this work, Al/SiO$_2$/Si MIS (i.e. MOS) tunnel structures are used as testing bench for the suggested measurement procedures.

These structures were fabricated on the boron-doped p-Si ($N_d \sim 10^{18}$–$10^{19}$ cm$^{-3}$) and phosphorus-doped n-Si ($N_d \sim 10^{16}$ cm$^{-3}$) wafers. Exact impurity concentration is of minor importance for our purposes. Thin oxide layers with the averaged thickness $d_n = 2.7$ nm were grown in a dry mixture of $O_2$ (20%) and $N_2$ at 700°C. The size of Aluminium contacts was $\varnothing = 200$, 400, 1000 $\mu$m (“large-area” devices) or 10 × 10, 10 × 20 $\mu$m$^2$ (“small-area” devices).

For each oxidised wafer, thickness deviation $\sigma_d$ was carefully measured, using the atomic force microscope (AFM) or, in some cases, the transmission electron microscope. Any of our fabrication cycles yielded SiO$_2$ films with a deviation $\sigma_d$ lying between 0.20 and 0.30 nm, i.e. noticeably exceeding the values typical for the modern silicon technology [1]. However, for this study, large dispersion of $d$ makes a benefit, as it warrants a strong effect of thickness variations on the device characteristics. All our results shown in the paper refer to the samples with $\sigma_d = 0.28$ nm.

Despite excessive oxide thickness fluctuations the behaviour of MOS elements fabricated in a same technological cycle has not shown any anomalies. For instance, for large-area diodes, the hot-electron-injection induced luminescence at positive substrate voltage was observed and the MOS tunnel emitter transistors exhibited rather high gain. So there is no reason for suspicions of irrelevancy of the structures for testing purposes, which might arise in the case of a huge, from a modern viewpoint, $\sigma_d$ value.

3. Definition and “covariant” determination of the parameter $\lambda$

The most direct way for obtaining information about the thickness fluctuations is to mathematically process the recorded dependency $d(x)$ of a thickness on the coordinate in the plane of a semiconductor/insulator interface. Particularly, the covariance between local thickness at the points $x$ and $x+l$ as a function of $l$ is written as

$$\text{cov}(l) = \frac{\langle (d(x) - d_n)(d(x+l) - d_n) \rangle}{\sigma_d^2}$$

(1)

It is evident that $\text{cov}$ should tend to 1 in the limit of $l \to 0$ due to a full correlation between the thicknesses at two neighbouring points on an infinitesimal distance. Oppositely, if $l$ is very large, $d(x)$ and $d(x+l)$ are independent quantities, the mean value of their product in (1) transforms to the product of mean values, each of them being equal to 0.

In principle, the parameter $\lambda$ could be defined as a value of $l$ at which the covariance (1) is reduced to some small value $\text{cov}_{\text{crit}}$ taken on agreement. However, it is often assumed that $\text{cov}(l)$ obeys Gauss ($\sim \exp(-l^2/\lambda^2)$) or exponential ($\sim \exp(-l/\lambda)$) law [2], and, in order to get $\lambda$, the measured dependency $\text{cov}(l)$ is fitted near $l=0$ by one of these functions. This approach yields, of course, different results for the two laws. Furthermore, $\text{cov}(l)$ and $\lambda$ may be sensitive to how the direction $x$ is set; if a possible planar anisotropy is not considered, an angular averaging may be additionally made.

In common, there is so far no consensus in some details related to the extraction of $\lambda$, and, perhaps, it would be more correct to call this parameter the “characteristic length” (as we did in [7]) instead of the “correlation length” of thickness fluctuation in order to avoid conflicting with formal mathematical definitions. But anyway, the “covariant” determination of the correlation (characteristic) length $\lambda$ is straightforward and elucidates the physical meaning of this value.

In practise, the preliminary measurement of the necessary dependencies $d(x)$ may be very complicated and time-consuming. For our samples, we succeeded in obtaining $d(x)$ — to within a constant term – using the AFM. Namely, we first measured the topography of our oxide layer and then the topography of silicon accessed to after etching the oxide. Since the surface in the second case was rather smooth, we assumed the thickness fluctuations to be completely reflected by the film relief. Fig. 1a represents the height profile $h(x)$ along some direction recorded with oxide, and, in order to extract the parameter $\lambda$, we can calculate a covariance handling with $h(x)$ as if it was $d(x)$ and using $\langle h \rangle$ at the place of $d_n$ in (1). The resulting function $\text{cov}(l)$ is shown in Fig. 1b. It is easy to deduce that $\lambda \sim 40–70$ nm. Some uncertainty is due to a freedom with $\text{cov}_{\text{crit}}$ in definition, as it was said above. Note also that no substantial anisotropy in the Si/SiO$_2$ interface plane was revealed for our samples.

4. Statistical method for the estimation of $\lambda$

Along with or instead of the microscopy-aided diagnostics, the electrical measurements can be used for a study of spatial fluctuations of insulator thickness.

So, the standard deviation of thickness $\sigma_d$ may be rather precisely estimated on the slope of current–voltage curves of large-area MIS structures (see in [7]). Such structures may be easily fabricated in the same cycle
at the same wafer for which the length $\lambda$ is to be determined.

For the finite parameters $\sigma_d$ and $\lambda$, the current should vary from sample-to-sample. Statistical characteristics of such variations for any value of the $L/\lambda$ ratio are to be found within our model [7]. It is assumed that the local thickness is not changed on a distance shorter than $\lambda$, i.e. $d$ remains constant within each cell $\lambda \times \lambda$, see Fig. 2. Such cells are suggested to cover the whole substrate. Current densities in each section are calculated using the "local" models developed for MIS structures with $d = \text{const}$ (see e.g. [9] and references therein).

Fluctuations of any other barrier parameters, than thickness, are ignored since the measured $I$–$V$ curves of very large-area devices had been satisfactorily reproduced by simulation with the given $\sigma_d$ and the ordinary, not fitted, values for the band offsets, permittivity etc. Particularly, the value of 3.15 eV was taken for the Si/SiO$_2$ conduction band disconti-

unity and $\epsilon$, see Fig. 3. One can see

![Image]

**Fig. 1.** (a) AFM measurements: height profile of SiO$_2$ surface; (b) covariance between local thicknesses in two points situated at a certain distance $l$ as a function of $l$.

![Image]

**Fig. 2.** Fragmentation of the device area by the grid with the fixed cell size $\lambda$. Within each cell, $d$ is assumed to be constant.

Below the sequential steps to perform within this statistical technique, are listed again:

1. Recording of the $I$–$V$ characteristics on a large random set of MIS capacitors of area $S=L^2$.
2. Extraction of the $\langle I/S \rangle$, $\sigma_{IS}$ (at some $V$) and $\mu$ values relying on the data of item 1.
3. Determination of the $d_n$, $\sigma_d$ parameters using the $I$–$V$ curve of some large-area capacitor.
4. Calculation of the $\mu = \mu(L/\lambda)$ dependence with the $\sigma_d$ value from the item 3 ($V$ and $d_n$ are also used here but have no substantial impact on the final result)
   a. simulation of the dependence of the current density $j$ on the thickness $d$ by usual MIS models,
   b. generation of the distribution $f_j$ of the variate $j$ using the normal law for the local thickness $d$ and the results of item 4a,
   c. calculation of the probability density $f_{j\mu}$ for the $I/S$ value, assuming that the device includes $n$ cells, as an $n$-fold convolute of the distribution function for $j$.
   d. definition of the possible $n = n_i$ values ($i = 1, 2, 3$) and their weights $p_i$, treating the $L/\lambda$ ratio as an argument and considering a shift of the device edge, respectively, to cell edges,
   e. production of the distribution function $f_{IS}$ of the $I/S$ quantity doing a superposition of functions $f_{j\mu}$ weighted with the factors $p_i$ for each $L/\lambda$,
   f. calculation of dependences $\sigma_{IS}(L/\lambda)$ and $\mu(L/\lambda)$ values using the function $f_{IS}$.

Estimation of $\lambda$ disposing at the experimental value of $\mu$ and at the calculated $\mu(L/\lambda)$ curve.

Calculated dependencies of $\mu$ vs. $L/\lambda$ for three values of $\sigma_d$ and two voltages $V$, are presented in Fig. 3.
that the curves corresponding to the same \( \sigma_d \), but to the different voltages lie close to each other. If we know \( L, \mu \) and \( \sigma_d \) for some device, the correlation length \( \lambda \) can be reasonably estimated, even if the wafer doping parameters, \( d_n \) value or bias conditions in experiment and in simulation are not identical. It should be admitted that the parameter \( \lambda \) extracted within this model of squares with fixed thickness and the parameter introduced in previous section may not be completely equivalent. However, there is no reason for worrying, as one should take in mind the existing uncertainty in the definition of \( \lambda \).

In inset of Fig. 3, a histogram of currents measured for a selection of our MIS diodes with the smallest area \( (10 \times 10 \mu m^2 \), i.e. \( L \sim 10 \mu m \) ) is shown, in order to illustrate the discussed method. The experimental value of \( \mu \) approximately equals 0.18. Using Fig. 3, we obtain the \( L/\lambda \) ratio of 200 and therefore, \( \lambda \sim 50 \) nm, which is in a satisfactory agreement with the results given by a straightforward method of Section 3. For the structures with twice larger \( (10 \times 20 \mu m^2 \) area, the parameter \( \mu \) was slightly less, as it qualitatively should be because the length \( \lambda \) does not depend on the device size.

The same routine of the determination of \( \lambda \) was also performed with the data on statistical current scattering borrowed from another work ([11], Fig. 6c there) where the value of \( \mu \) is 0.102. Assuming that \( L=2.3 \) \( \mu m \) (square root from the gate area in [11]) and that \( \sigma_d \sim 0.2 \) nm, we have \( L/\lambda \sim 50 \) and estimate \( \lambda \sim 40 \) nm. Naturally, the coincidence of correlation lengths in the present study and in [10] is a pure occasion, but nevertheless it is a good sign that such an important fluctuation parameter as \( \lambda \) in our samples does not depart dramatically from what could be deduced from the independently published data.

5. Breakdown method for estimation of \( \lambda \)

The procedure of estimating \( \lambda \) based on the statistical analysis of experimental \( I-V \) curves is not suitable for large-area devices. The matter is that the sample-to-sample variations of the current (at a fixed \( V \)) arising from the effect of \( \lambda \) should be very small in this case. Moreover, some technological defects in the samples can become responsible for a current spread. Therefore, another technique for the determination of \( \lambda \) is proposed here in order to cover also the case of \( L \gg \lambda \).

In presence of a noticeable thickness deviation \( \sigma_d \) device parts with the smallest local thicknesses \( d \) provide a major contribution to the total current (Fig. 4a). Within our model of squares, this means that the current is crowded in one \( \lambda \times \lambda \) cell. However, the simplifying assumption of a fixed thickness inside each section is impractical with respect to the thinnest cells. More convenient is to estimate a fraction of current \( \xi \) flowing through the thinnest spot as: \( \int_0^\xi j(d)\Gamma^+(d)\,dd/\int_0^{\infty} j(d)\Gamma^+(d)\,dd \), where \( \Gamma^+ \) is the renormalized for \( d>0 \) Gaussian distribution and the upper limit in integral \( D \) is taken from the condition: \( \eta = \int_0^D \Gamma^+(d)\,dd = 1/N \) with \( N=\sim(L/\lambda)^2 \sim S/\lambda^2 \). Such an estimation has been performed in [8] and, referring to it, for our samples having \( \sigma_d = 0.28 \) nm, 1% of total current should flow through only \( 10^{-8} (\eta=10^{-8}) \) fraction of the device area \( S \). If, for some reason, the thinnest cell is excluded from the current transport, an abrupt reduction of \( I \) (at a fixed bias) is expected. This would enable to estimate the value of \( \lambda \).

In the paper [8], we reported that, under certain conditions, soft breakdown in an MOS tunnel structure may cause an abrupt decrease of current. Such a current jump-down has been observed on many different Al/SiO2/Si samples and also on MIS tunnel structures with the epitaxially-grown CaF2 insulator layer [12]. For a simple explanation of this effect, it was assumed [8] that the soft breakdown spot is a device section whose local current–voltage characteristic is ohmic, in contrast to the undamaged sections, where the characteristic is tunnel (approximately exponential, Fig. 4b).

**Fig. 3.** Dependency \( \mu = \mu(I/V) \) calculated for the thickness deviation values \( \sigma_d = 0.1, 0.2, \) and \( 0.3 \) nm and two voltages \( V=1.0 \) and \( 3.0 \) V. The circle marks the parameter of \( \mu \) obtained from the data on statistical scattering of current in our samples, shown in the inset as a histogram, and the square—the parameters \( \mu \) from the work [11].

**Fig. 4.** (a) Current fraction and corresponding areal fraction in the presence of considerable \( \sigma_d \). (b) \( I-V \) characteristics of the damaged spot and of the same area before the insulator has been damaged. (c) Abrupt current decrease after the stress; \( \xi = \Delta I/I_0 = \left|I-I_0\right|/I_0 \), where \( I_0 \) and \( I \) are the currents before and just after the jump-down, respectively.
Under moderate voltages, the total current increases after the stress due to the contribution of the broken zone. However, the tunnel resistance rapidly decreases with $|V|$ and – for any tunnelling parameters – finally becomes lower than the resistance of the damaged area. In other words, the tunnel exponent should anyway intersect the linear $I-V$ characteristic at some bias $|V|=V^*$ (Fig. 4b). This explanation is also valid if a more general power-law approach for the conductivity of a breakdown spot ($I \sim V^x$ where $x$ is a parameter) is adopted.

It is the thinnest device cell that is first broken down during the electrical overload of a device. At $|V| > V^*$, an abrupt reduction of current after the stress should be observed due to the exclusion of a cell with the minimal $d$ from current transport. One can speak about the “exclusion”, since the tunnel exponent increases with $|V|$ much faster than any power-law characteristic.

The proposed method for estimation of $\lambda$ includes the following steps. First, it is necessary to apply the constant–voltage stress (CVS) at a sufficiently high bias until the soft breakdown occurs. Then, one can find the relative current decrease, i.e. the value of $\xi = (I_0-I)/I_0$ where $I_0$ and $I$ denote the currents before and just after the jump-down (Fig. 4c). Then it is possible to calculate – for a known deviation $\sigma_d$ – the areal fraction $\eta$ corresponding to the obtained current fraction $\xi$ (see Fig. 4c). The final step is the estimation of a linear size of the damaged spot, which effectively equals to the value of $\lambda$, namely $\lambda \sim l_{SB} = (\eta S)^{1/2}$.

In Fig. 5, the most typical behaviour of a current through the MOS tunnel structure during the CVS is presented. The values of $I$ are much larger than in Fig. 3 because of larger area here: $\Omega = 400 \mu m^2$, $S = 1.26 \times 10^{-3} \text{cm}^2$. The inset of Fig. 5 shows the evolution of the dependencies $\Delta I/I_0$ on $V$ after multiple stresses. This inset demonstrates the main tendency: at relatively moderate $|V| < V^*$, there is a growth of current after the stress, whereas for $|V| > V^*$ the considerable decrease in current occurs. For the case presented in the main Fig. 5, the relative decrease of current is $\xi \sim 10^{-2}$. Therefore, as follows from a discussion at the beginning of the section, for our samples we should take $\eta = 10^{-8}$ and it is easy to find $\lambda \sim (10^{-8} S)^{1/2} \approx 40 \text{nm}$.

Breakdown occurrence is known [13] to depend on the gate material. For demonstration, the samples with Al electrodes have been taken. For MOSFET electronics, more relevant would be the poly-Si gated devices, but, as said at the end of Introduction, we do not imply MOSFETs as an exclusive or even main application for our methods of estimating $\lambda$.

However, still as a link to the field-effect transistors, it may be not useless to mention that the applied CVS corresponds to the uniform hot carrier stress (e.g. [14]) of a MOSFET when its source, drain and substrate are grounded. This uniform stress is in contrast to the channel hot carrier stress [13] where the damage is localised closer to the drain region and therefore the effect of thickness deviation may be masked.

To check the independence of this result of the area $S$, we performed similar measurements with differently-sized large-area MOS devices and became sure that for the same oxidised wafer, $\xi$ (and, therefore, $\eta$) are changed approximately inversely proportional to $S$, as it should be for $\Delta I \propto I_0$. We also examined this technique attracting samples with other values of $\sigma_d$ varying within 0.2 and 0.3 nm and obtained $\lambda$ lying in the range of 30–60 nm with the most probable value of $\sim 40 \text{nm}$. Here we do not perform any statistical analysis of this spread, because – as already mentioned – the definition of $\lambda$ is the subject of agreement.

For verification, it is important that values of $\lambda$ obtained for our oxide films within this breakdown method and within the basic “covariant” method (Section 3) are in a quite good agreement.

6. Conclusion

In this work, we have proposed two methods for estimating the correlation length $\lambda$ of the insulator thickness fluctuations in MIS tunnel structures based on the electrical measurements. The first method implies the statistical treatment of current spread in a device series and is applicable if the expected value of $\lambda$ is larger than or comparable to the device size $L$. The second method presumes the analysis of current changes after the electrical stress in the large-area structures (the case $L \gg \lambda$).

The knowledge of the parameter $\lambda$ is important, because $\lambda$ – together with $\sigma_d$ – is a quality indicator for a thin insulator film. But in this paper we neither aimed to use the measurements of correlation lengths for the control of sample fabrication nor strived to attain possibly small values of $\lambda$. The goal was just to demonstrate new techniques.

These techniques were quite successfully examined on $\text{SiO}_2$ layers with a noticeable thickness deviation $\sigma_d$ in $\text{Al/AlO}_x/\text{Si}$ MOS tunnel diodes. We intentionally chose this system as a testing bench because the tunnelling in it is reliably parameterized. However, one may suppose that the accuracy of our methods will be worse than that of the microscopy-aided diagnostics for very small $\sigma_d$, which is just the case of modern MOSFET technology. For this reason, the new methods are potentially more interesting and useful for tunnel-thin films of insulating materials for which the overall technology level is inferior to that of $\text{SiO}_2$ (as e.g. high-k gate
oxides). The main advantage of the proposed techniques for estimating $\lambda$ is their simplicity.

Acknowledgements

This work is financially supported by the Austrian Science Fund FWF under the Project no. P18316-N13.

References