JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

1/2005

Advanced silicon devices and technologies for ULSI era

Special issue edited by Andrzej Jakubowski and Lidia Łukasiak

Part 2

Editorial Board

Editorial Advisory Board

ISSN 1509-4553

Circulation: 300 copies

Sowa - Druk na życzenie, www.sowadruk.pl, tel. 022 431-81-40

JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

Preface

Despite the fact that a range of limitations are beginning to appear as CMOS technology is being raised to ever higher levels of perfection, it is anticipated that silicon will be the dominant material of the semiconductor industry for at least the first half of the 21st century. The forecast for microelectronics development published in 2001 by SIA (Semiconductor Industry Association) reaches ahead to the years 2010–2016. Moreover, a comparison with former SIA forecasts indicates that they become more aggressive (that is more optimistic) with time.

While the development of silicon microelectronics in the past could be attributed mostly to the reduction of the feature size (progress in lithography), today it relies more on new material solutions, such as SOI, SON, SiGe or SiC. The combination of this trend with continuous miniaturization provides the opportunity of improving IC functionality and speed of operation.

Telecommunications and information technology are arguably the most powerful drivers behind microelectronics product development nowadays. Thus plenty of new applications are being created for fast analog and RF circuits, as well as for information processing ones. It is clear that with the anticipated $f_{\max} = 130$ GHz and $f_T = 65$ GHz to be reached by RF bipolar transistors in 2005, according to the 2001 issue of ITRS (it is interesting to note that in 1999 ITRS predicted only $f_{\max} = 50$ GHz and $f_T = 40$ GHz for 2005), a lot of effort must be put into the development of appropriate material, processing, characterization and modeling. Such an outstanding progress, however, will not happen without increased speed offered by new materials solutions. SiGe-base HBTs may serve as an example here – a device with $f_T = 350$ GHz has already been reported in the literature.

High-speed isn't, however, everything. Portable wireless products push, for obvious reasons, for low-power solutions. This trend requires new architectural solutions (e.g., channel thinning), and in consequence, new material, such as SOI (or its possible successor SON – silicon-on-nothing), where current driveability is considerably higher than in conventional MOSFETs.

In this issue the Reader will find the second part of the papers and lectures presented during the 6th Symposium Diagnostics & Yield: Advanced Silicon Devices and Technologies for ULSI Era, which took place in Museum of Earth, Warsaw, Poland on June 22–25, 2003. A number of the papers are devoted to the most important issues concerning semiconductor technology (challenges faced by CMOS technology, gate-dielectric fabrication, wafer-cleaning problems, fabrication of nanostructures, micromachined sensors) and advanced materials, such as SOI and SiGe. Several papers address the problems of characterization of semiconductor structures and nanostructures (DC and noise analysis, photoelectric measurements, scanning probe microscopy), as well as those of IC testing and assembly. The subject of modeling (both devices and circuits) and parameter extraction is covered, too.

We hope the Readers will find these Proceedings useful and interesting.

Organization of this Symposium would not be possible without the support of Polish Committee for Scientific Research (research projects no. 4T11B03523, 8T11B07519, 8T11B07419). The organizers acknowledge also the contribution of the Committee of Electronics and Telecommunications of Polish Academy of Sciences, and the NEXUS Microsystems Association. The organizers are also grateful to the National Institute of Telecommunications for making this journal available for the publication of the Symposium papers.

Andrzej Jakubowski – General Chairman
Lidia Łukasiak – Programme Chairman
Guest Editors

Invited paper

Challenges in scaling of CMOS devices towards 65 nm node

Małgorzata Jurczak, Anabela Veloso, Rita Rooyackers, Emmanuel Augendre, Sofie Mertens, Aude Rotschild, Marc Scaekers, Richard Lindsay, Anne Lauwers, Kirklen Henson, Simone Severi, Ivan Pollentier, and An de Keersgieter

Abstract—The current trend in scaling transistor gate length below 60 nm is posing great challenges both related to process technology and circuit/system design. From the process technology point of view it is becoming increasingly difficult to continue scaling in traditional way due to fundamental limitations like resolution, quantum effects or random fluctuations. In turn, this has an important impact on electrical device specifications especially leakage current and the circuit power dissipation.

Keywords— CMOS devices, gate dielectrics, shallow junctions, silicide, gate stack, lithography, gate patterning, silicon recess, device integration.

1. Introduction

The traditional scaling approach, which has been the base of the semiconductor industry for the last 30 years is beginning to show the fundamental limits of the materials building the modules of a planar CMOS transistor. Significant efforts are being devoted to the introduction of new materials that will replace the existing ones to further extend the device scaling process.

However, even though the new materials and device architectures are being investigated, the gate length scaling below 60 nm has a significant impact on the electrical properties of the devices. The continuous thinning of the gate dielectric layers, the increasing channel doping and the aggressive, abrupt junctions, required to control the short channel effects, start to significantly affect circuit power dissipation.

Nitrided oxides, gate predoping, offset spacers, spike anneal, tilted pocket implants, ultra low energy implants are a few examples of the process steps that were introduced in the 0.13 μm technology generation to enable device gate length scaling. These previous "improvements" will be able to prolong scaling down to 65 nm node. New solutions to device architecture, device substrate, gate stack will have to be introduced to maintain 12% increase in device performance forecasted by ITRS roadmap [1].

This paper gives an overview of the main challenges being faced in the front-end of line (FEOL) technology development for 65 nm node. This includes the gate stack, gate patterning, junction and silicide process.

2. Gate stack

Aggressive scaling of CMOS devices puts severe constrains on the gate dielectric. In order to meet requirements for drive current and off-state leakage set up by ITRS roadmap it is essential to simultaneously scale further gate oxide thickness and limit the gate leakage current. In addition, mobility degradation has to be minimized and gate oxide reliability still has to meet the 10-year lifetime requirements. Despite low gate leakage current as compared to nitrided oxides (Fig. 1), high-k dielectrics are still not ready for device integration due to reliability problems, high mobility degradation and instability of the threshold voltage.

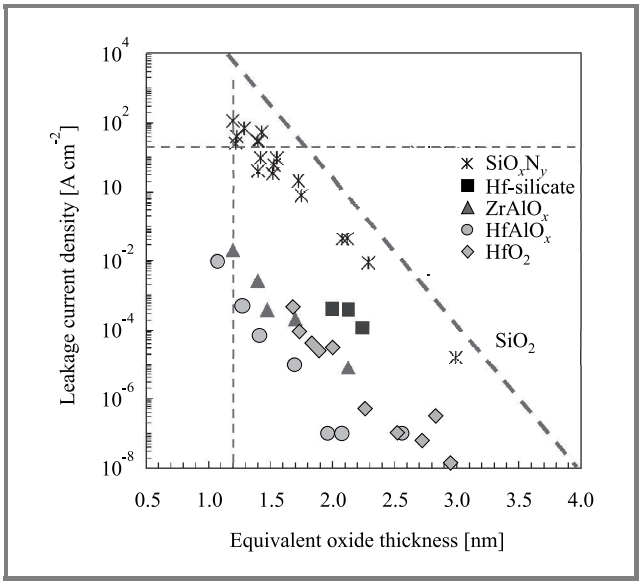


Fig. 1. Comparison of different gate dielectrics with respect to gate leakage and EOT.

Therefore, there is strong motivation to extend the use of oxynitrides to its ultimate limits. Oxynitrides fabricated with a new plasma nitridation techniques such as decoupled plasma nitridation (DPN) [2] or slot plane antenna (SPA) seem to be the best candidates for 65 nm high performance and general purpose CMOS devices (Fig. 1) at least until high-k materials reach their maturity for the gate dielectric applications. Using plasma nitridation it is possible to introduce more nitrogen than with furnace nitridation techniques into thin oxide layer and thus increase the dielectric constant k and scale equivalent oxide thickness without reducing significantly physical oxide thickness [3].

Continued thinning of gate dielectrics and an increase of the interface-state density created by plasma nitridation bring many concerns related to reliability issues. Thorough characterization shows, however, that a careful optimization of the gate dielectric process enables maintaining the maximal operation voltage, at which the device can be operational for 10 years, above the supply voltage [4]. This is possible since oxide reliability is dependent rather on physical oxide thickness than on EOT.

The gate activation, which has direct impact on the drive current, is more and more affected by the reduced thermal budged required in order to control short channel effects. Additional gate implantation called gate predoping is mandatory to maintain low gate depletion. Metal gate seems to be a solution to this problem, however, due to the unsuitability of the work function and highly complex integrity it will not be ready for the 65 nm node.

3. Gate patterning

In order to cope with progressive scaling of CMOS devices new lithography techniques are being investigated. The 157 nm lithography is considered as the leading candidate for semiconductor device manufacturing at the 65 nm technology node [5]. However, development of 157 nm lithography requires considerable effort and effective solutions to resist and reticle materials, lenses, CD metrology, etc. This list of challenges raises fears that the 157 nm lithography will not be ready for the 65 nm node. Therefore, other methods such as resist or hard mask trimming were developed to support lithography in the scaling progress.

In the 65 nm node, where the gate length is in the range of 65–25 nm, depending on applications, the key process step developments are lithography, dry etch, and CD metrology. With respect to lithography, the emphasis is put on the development of thin resist processes (using 193 nm lithography) in combination with alternating phase shift mask (AltPSM) optical extensions. With this combination the lines down to 65 nm with good CD control can be printed (Fig. 2a). Further reduction of the gate length is obtained by subsequent resist or hard mask trimming (Fig. 2b). Trimming techniques allow obtaining gate lengths down to a few nanometers.

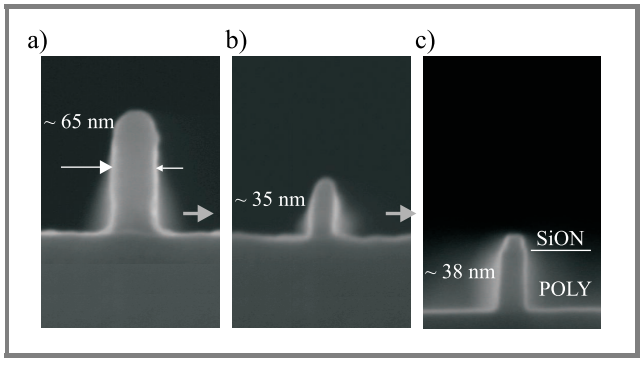


Fig. 2. Visualization of the trim and gate etch: (a) showing the X-section of features after litho; (b) etch trim; (c) gate etch and strip.

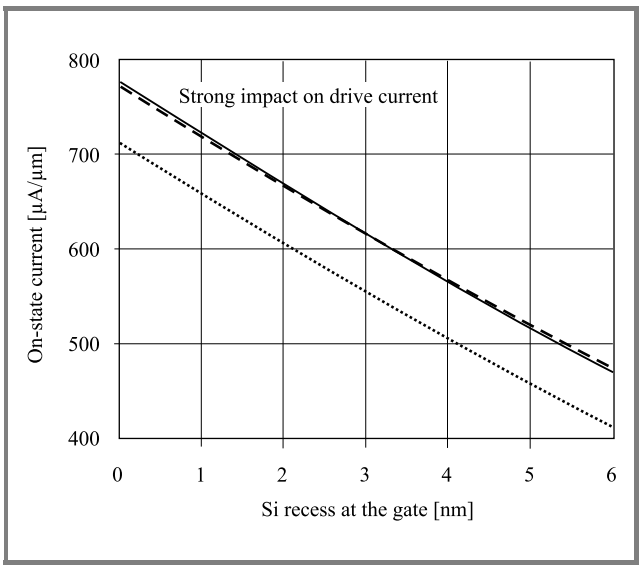


Fig. 3. Impact of Si recess created at the gate patterning on the drive current (simulation results).

An important issue when scaling devices and the gate oxide thickness is silicon recess in source and drain areas which occurs during overetch step in gate electrode patterning. Even small, a few nanometers consumption of Si in the substrate has considerable impact on device performance and leads to increase in the off-state leakage and series resistance (Fig. 3).

4. Junction

Approaching the 65 nm node ITRS requirements for ultrashallow junctions (USJ) has become a great challenge, es-

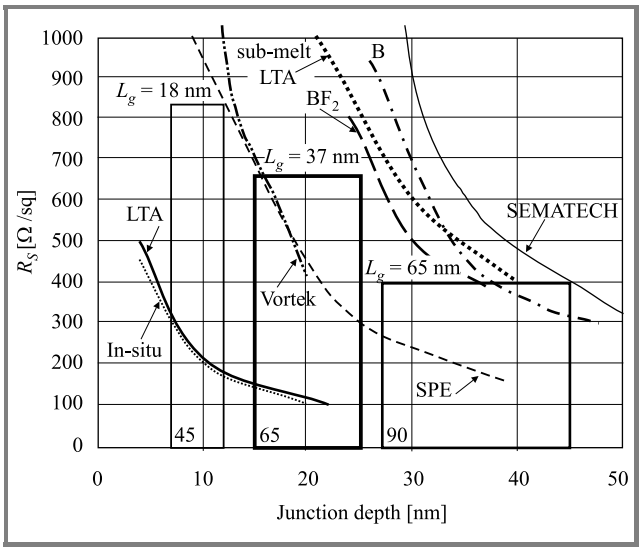


Fig. 4. Summary of p^+/n junctions obtained with different techniques. The boxes indicate the requirements for different technology generations (literature results).

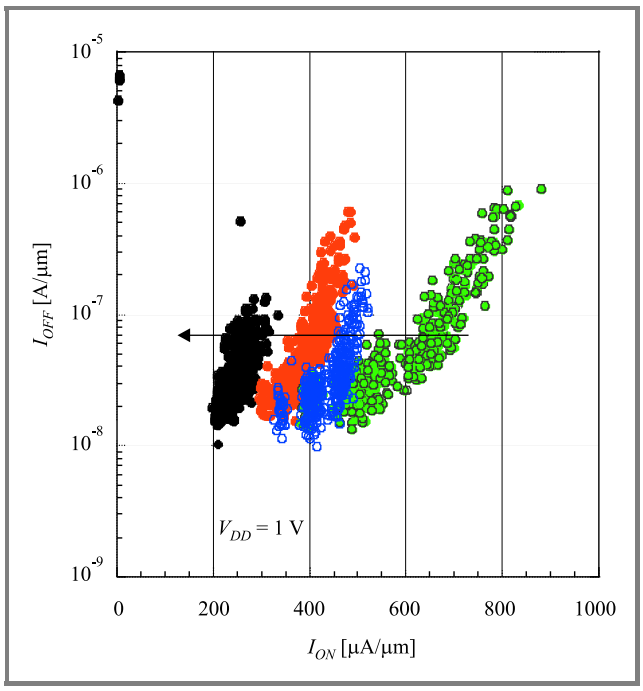


Fig. 5. Drive current versus off-state leakage for NMOS devices with different doses in extensions. The arrow indicates the reduction in the doping concentration in the junctions.

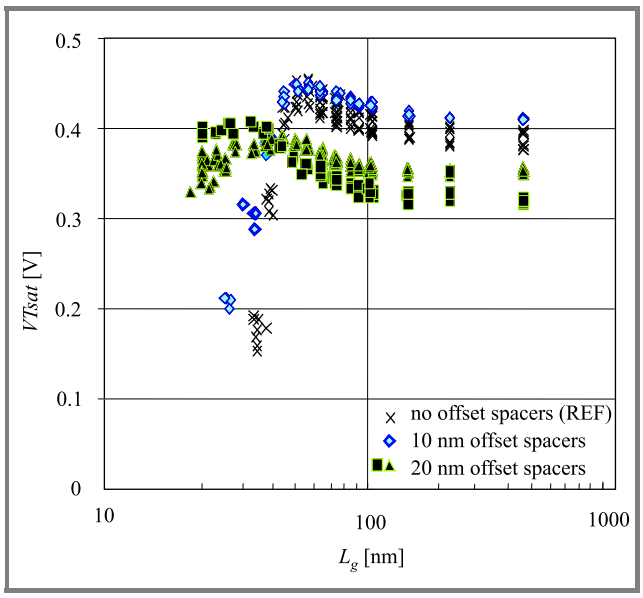


Fig. 6. Impact of offset spacers on control of short channel effects.

pecially for p^+ dopants. With boron, traditionally used for p^+/n junction formation, implanted alone and annealed with conventional methods, the 90 nm node specifications can hardly be reached (Fig. 4). Simple lowering of implantation dose reduces the junction depth but on the other hand leads to high series resistance and degradation of the on-state current (Fig. 5). The literature data shows that only with very sharp spike anneal (high rump-up and rump down using, e.g., Levitor or Vortek tool) the improvement in the trade-off between junction resistance and junction depth can be achieved [6]. Even better results can be obtained with a co-implantation with other elements and/or pre-amorphization implants in conjunction with spike anneal. The combination of Ge pre-amorphization, F co-implantations and fast ramp-up spike anneal appeared to be sufficient to fulfill the 65 nm node requirements [7]. For the next generation other alternative routes of forming the junctions such as solid phase epitaxial re-growth (SPER) [8] or laser anneal (LTA) are investigated.

Introduction of offset spacers allows relaxing the requirements for the extensions without aggravating short channel effects (Fig. 6). This, however, puts very severe constrains on their width control since any small variations in the spacer width may result in huge variation in VT and off-state leakage.

5. Silicide

The scaling of the gate length (below 40 nm) goes in parallel with the reduction in the HDD junction depth. Shallower junctions, in turn, determine more constrained requirements for the silicides. Conventionally used CoSi seems to reach its ultimate limits. For the gate length below 40 nm gate sheet resistance increases dramatically (Fig. 7), which is related to silicide instability and silicide cracking. No improvement is observed when Co is alloyed with Ni. Besides, high Si consumption during the silicide process leads to unacceptable increase in the junction leakage. Therefore, other materials such as NiSi or PtSi were considered as potential candidates for 65 nm technology node [9].

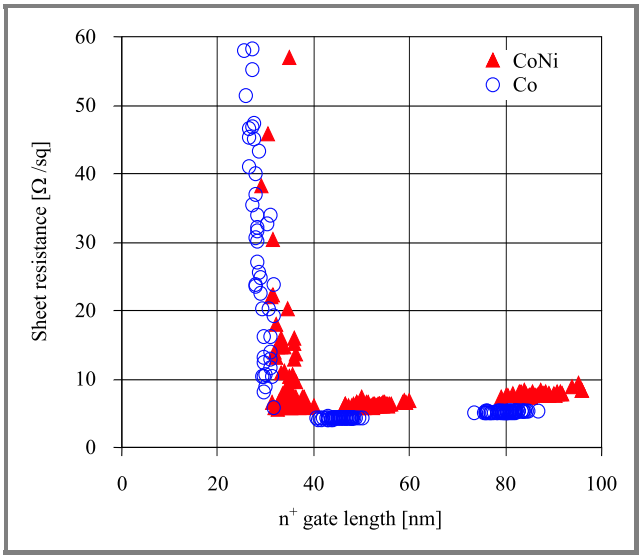


Fig. 7. Gate sheet resistance of Co and CoNi silicided gate as a function of the gate length.

For Ni-silicided poly gates low sheet resistance is obtained down to the narrowest line widths (Fig. 8). The trade-off between junction leakage and silicide sheet resistance is

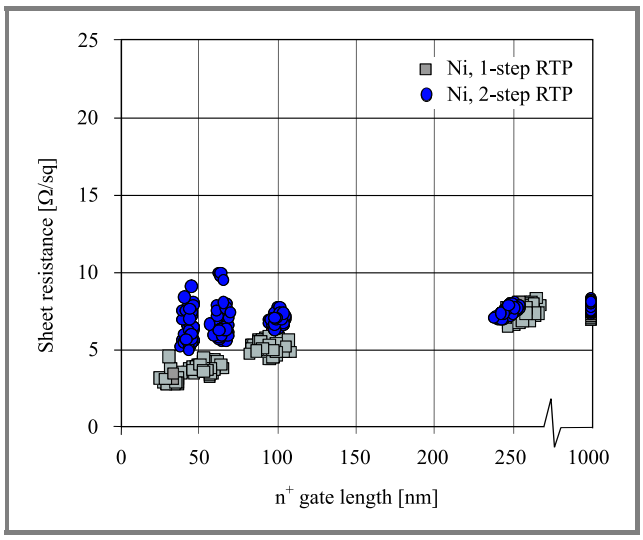


Fig. 8. Gate sheet resistance of Ni silicided gate processed with different annealing steps as a function of the gate length.

in favour of Ni-silicide. Besides, NiSi is known to have smaller contact resistance, which has significant role in total resistance of the junctions in sub-50 nm devices.

6. Conclusions

Conventional bulk scaling reaches its physical limitations. The 65 nm CMOS appears to be the last technology node, where the conventional planar transistor architecture in conjunction with standard modules will be applied. In order to maintain the trend in performance improvement, next technology generations will probably have to deal with more exotic solutions such as strained silicon, metal gate integration, high-*k* dielectrics (mainly for low power applications) and eventually double gate architectures.

References

[1] International Technology Roadmap for Semiconductions, 2001, http://public.itrs.net/Files/2001ITRS/Home.htm

[2] F. N. Cubaynes, J. Schmitz, C. van der Marel, H. Snijders, A. Veloso, and A. Rothschild, "Plasma nitridation optimization for sub-1 A gate dielectrics", in *ECS Symp*., Paris, France, 2003.

[3] A. Veloso, F. N. Cubaynes, A. Rothschild, S. Mertens, R. Degraeve, R. O'Connor, C. Olsen, L. Date, M. Schaekers, C. Dachs, and M. Jurczak, "Ultra-thin oxynitride gate dielectrics by pulsed-RF DPN for 65 nm general purpose CMOS applications", in *Conf. ESSDERC'03*, Lisbon, Portugal, 2003.

[4] R. O'Connor, R. Degraeve, B. Kaczer, A. Veloso, G. Hughes, and G. Groeseneken, "Weibull slope and voltage acceleration of ultra-thin (1.1–1.45 nm EOT) oxynitrides", in *Conf. INFOS'03*, Stará Lesná, Slovakia, 2003.

[5] L. Van den Hove, M. Goethals, K. Ronse, M. Van Bavel, and G. Vandenberghe, "Lithography for sub-90 nm applications", in *IEDM Tech. Dig.*, San Francisco, USA, 2002, pp. 3–8.

[6] R. Lindsay, B. Pawlak, J. Kittl, K. Henson, C. Torregiani, S. Giangrandi, R. Surdeanu, W. Vandervorst, A. Mayur, J. Ross, S. McCoy, J. Gelpey, P. Stolk, and K. Maex, "A comparison of spike, flash, SPER and laser annealing for 45 nm CMOS", in *MRS Symp*., Paris, France, 2003.

[7] B. J. Pawlak, R. Lindsay, R. Surdeanu, X. Pages, W. Vandervorst, and K. V. D. Jeugd, "The role of F with Ge pre-amorphisation in forming PMOS junctions for the 65 nm CMOS technology node", in *ECS Symp*., Paris, France, 2003.

[8] R. Lindsay, B. J. Pawlak, J. Kittl, K. Henson, S. Giangrandi, R. Duffy, R. Surdeanu, W. Vandervorst, P. Stolk, and K. Maex, "Leakage optimisation of ultra-shallow junctions formed by solid phase epitaxial regrowth (SPER)", *Proc. J. Vac. Sci. Technol*, vol. B, Febr. 2003.

[9] A. Lauwers, J. Kittl, A. Akheyar, M. Van Dal, O. Chamirian, M. de Potter, R. Lindsay, and K. Maex, "Silicide scaling: Co, Ni or CoNi?", in *ECS Symp*., Paris, France, 2003.

Małgorzata Jurczak received M.Sc. and Ph.D. degrees in electrical engineering from the Warsaw University of Technology (WUT), Poland, in 1991 and 1997. In 1991 she joined WUT, where she worked on modelling of MOS SOI devices. In 1994 she was with NMRC, Cork, Ireland, and in 1997 with Kyung Hee University, Seoul, Korea. In 1998 she joined CNET Grenoble, France Telecom. She was involved in development of 0.18 and 0.12 μm CMOS process and alternative approaches for sub-0.1 μm CMOS. In 2000 she joined IMEC, Leuven, Belgium. In the years 200–2003 she was the IMEC coordinator of the JDP program with Philips on device process integration for 90 nm and 65 nm CMOS. From 2003 she has been the group leader of the CMOS Device Implementation Projects (45 nm CMOS, FUSI gate, strained Si, HK+MG integration, FD SOI, MUGFET). She is also a project manager of IMEC Industrial Affiliation Program (IIAP) on ultimate devices EMERALD. She has participated in several European projects: ACE, ULTRA, MEDEA, HUNT. She was the project coordinator of the European IST ARTEMIS project. Currently she is in the Technical Advisory Committee of the IST Integrated Project NanoCMOS. She holds 15 French and American patents and authored and co-authoured more than 60 papers. She has been a member of scientific committee of IEDM (in 2002 and 2003) and ESSDERC conferences (2000, 2004 and 2005). She received the Best Paper Award at ESSDERC'2000 Conference and the Paul Rappaport Award for the best paper published in IEEE TED in 2000.
e-mail: malgorzata.jurczak@imec.be
IMEC
Kapeldreef st 75
3001 Leuven, Belgium

Anabela Veloso, Rita Rooyackers, Emmanuel Augendre, Sofie Mertens, Aude Rotschild, Marc Scaekers, Richard Lindsay, Anne Lauwers, Kirklen Henson, Simone Severi, Ivan Pollentier, An de Keersgieter
IMEC
Kapeldreef st 75
3001 Leuven, Belgium

Invited paper

Gate dielectrics: process integration issues and electrical properties

Udo Schwalke

Abstract—In this work we report on the process integration of crystalline praseodymium oxide (Pr_2O_3) high-*k* gate dielectric. Key process steps that are compatible with the high-*k* material have been developed and were applied for realisation of MOS structures. For the first time Pr_2O_3 has been integrated successfully in a conventional MOS process with n^+ polysilicon gate electrode. The electrical properties of Pr_2O_3 MOS capacitors are presented and discussed.

Keywords— high-k dielectrics, CMOS, Pr_2O_3, process integration, resist removal, wet chemical cleaning, wet chemical etching, RIE.

1. Introduction

The scaling of gate dielectric thickness is a major challenge for future ULSI CMOS technologies. Only by taking advantage of in-situ gate stack processing and advanced CMOS process architectures, excellent CMOS device characteristics could be achieved with ultra-thin (1.6 nm) SiO_2 gate dielectric [1]. However, the gate oxide thickness has to be scaled down aggressively much further into the sub-1 nm regime within the next decade, according to the ITRS roadmap [2]. At such a thickness, unacceptably high direct tunneling leakage current will flow through the gate oxide. The exponential increase in tunneling current with decreasing film thickness represents a fundamental scaling limit for SiO_2. In order to overcome this barrier, the use of alternative gate dielectrics with a higher permittivity (high-*k*) is urgently needed. Due to the high-*k* values, gate dielectrics with sub-1 nm equivalent oxide thickness can be realised and show acceptable low leakage currents. Unfortunately, most of these amorphous materials are not sufficiently stable [3], change phase [4] and electrical properties at temperatures much lower than required for CMOS processing. Accordingly, the most desirable direct substitution of SiO_2 in a given CMOS process is not possible.

Very recently it was shown that crystalline praseodymium oxide (Pr_2O_3) films have excellent dielectric properties [5]. Besides an effective dielectric constant of $K_{Pr_2O_3} \approx 30$ and low leakage currents ($< 10^{-8}$ A/cm^2 @ 1 V & EOT$=$ 1.4 nm), Pr_2O_3 films were found to be thermally stable up to 1000°C for 15 s RTA anneals in N_2 ambient. Because of these attractive properties, Pr_2O_3 appear as a suitable high-*k* replacement for SiO_2.

Although these first results are very promising, a successful CMOS process integration of crystalline high-*k* Pr_2O_3 gate dielectric requires a substantial amount of process development. Besides thermal processing, especially wet chemical etching and cleaning procedures as well as reactive ion etching (RIE), processes compatible with the new material have to be developed. In addition, process-damage effects on the high-*k* dielectric needs to be evaluated and minimised.

In this work we report on a successful attempt to integrate Pr_2O_3 directly into a given n^+ polySi gate MOS technology. The electrical properties of the MOS devices are presented and will be discussed.

2. Praseodymium oxide growth and wafer preparation

For the process integration experiments Pr_2O_3 was grown on hydrogen terminated 3″ Si(100) wafers in a multichamber molecular beam epitaxy (MBE) system [5]. The MBE is equipped with e-beam evaporators for Si and praseodymium oxide. Pr_2O_3 layers were grown on preimplanted p-type Si(100) wafers at temperatures between 625 and 725°C using a commercially available ceramic Pr_6O_{11} source. The layers were subsequently covered in-situ with 50 nm undoped polySi, since unprotected Pr_2O_3 layers are not stable against air [6]. The polySi protected Pr_2O_3 wafers were transferred to the CMOS process fab of the Institut für Halbleitertechnik (IHT) at the Technische Universität Darmstadt for process development. After RCA clean and a brief HF-dip a 250 nm in-situ phosphorous doped polySi layer was deposited at approximately 750°C. The structures obtained in this way were mainly used for subsequent process development.

3. Development of key process steps

3.1. Wet chemical cleaning and etching

Wet chemical processing is extensively used in CMOS manufacturing, for cleaning, etching and resist removal. However, little is known on the compatibility of these standard procedures with high-*k* materials, especially Pr_2O_3. In order to obtain first-hand information, blanket Pr_2O_3 test samples were subjected to various wet chemical treatments. The results of wet chemical cleaning and etching

studies are summarized in Table 1. Pr_2O_3 remains stable in DI-water, HF-dip and the alkaline component of the RCA clean, and no change in the film thickness was observed by means of ellipsometry. However, Pr_2O_3 was found to dissolve in the standard RCA clean and appears unstable

Table 1
Wet chemical cleaning and etching

Procedure	Effect on Pr_2O_3
HF-dip	Appears stable
Buffered HF	Unstable
RCA clean	Soluble
RCA clean, alkaline component only	Stable
H_3PO_4	Soluble
HCl	Soluble
DI-water	Stable

in buffered HF solution. HCl solution of 0.75% was found to etch Pr_2O_3 with a high etch rate of 28 nm/min. Obviously, acid containing solutions, in particular HCl, are suitable to remove Pr_2O_3 films selectively from the silicon surface.

3.2. Resist removal

Resists used for lithography have to be removed after etching or ion implantation. However, several wet chemical methods for resist removal contain acids and are, therefore, expected to dissolve Pr_2O_3. In fact, such a behavior was observed as summarized in Table 2. Only organic solvents, like acetone or AZ100 remover, are suitable for resist stripping when Pr_2O_3 films are present.

Table 2
Resist removal

Procedure	Effect on Pr_2O_3
Acetone	Stable
H_2SO_4/H_2O_2	Soluble
HNO_3 (100%)	Soluble
AZ100 wet remover	Stable
O_2-plasma ashing	Change in film properties

Besides wet chemical treatments, resist ashing in O_2 plasma is a commonly used procedure. A clear change in film properties was observed after O_2 plasma treatment, especially the etch rate in buffered HF was found to increase compared to Pr_2O_3 reference samples. Most likely oxygen radicals produced in the plasma are incorporated in the praseodymium oxide film and/or may diffuse through and react with the Si-interface. In order to obtain further information we deposited aluminum dots on these samples and performed *CV* measurements. We observed a reduction in the dielectric constant and an increase in the threshold voltage. Both results suggest that an interfacial oxide is formed as a result of the O_2 plasma treatment, similar to the film degradation of unprotected Pr_2O_3 when exposed to air for sufficiently long times [6].

3.3. Anisotropic reactive ion etching

The development of a highly selective RIE process is another major issue for a successful process integration of praseodymium oxide in the polySi gate CMOS process.
In order to avoid damaging of the silicon substrate, the polySi gate etch has to stop at the dielectric. We found that our standard SF_6/Cl_2 process fulfills these requirements. The selectivity of the polySi-RIE against Pr_2O_3 is higher than 300 so that the process window is sufficiently large to allow for the necessary over-etching.

4. Results and discussion

4.1. Process integration of Pr_2O_3

After the appropriate RCA clean and a brief HF-dip a 250 nm in-situ phosphorus doped polySi layer was deposited on the 50 nm polySi-covered Pr_2O_3 wafers at approximately 750°C. Subsequently, the standard lithography was applied using the gate mask. The n^+ polySi gate stack is defined by anisotropic reactive ion etch using SF_6/Cl_2.

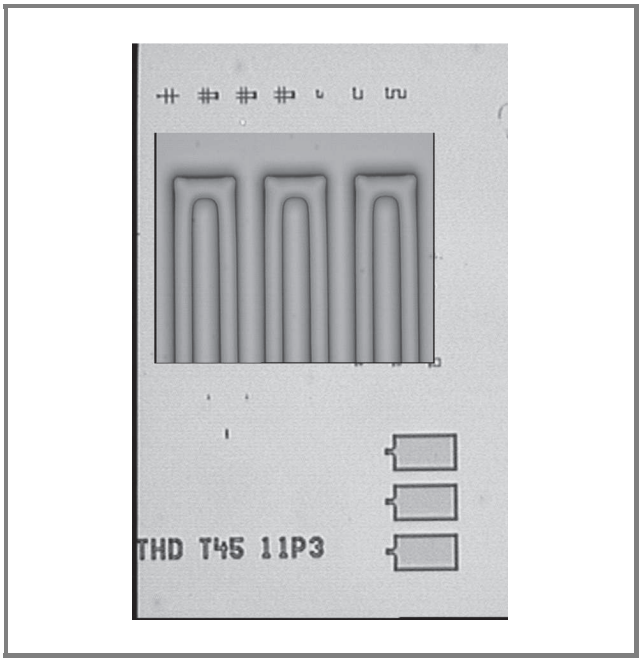


Fig. 1. Microscope image of n^+ polySi/Pr_2O_3 gate test structures after RIE. The inset shows an enlarged part of a meander structure.

In Fig. 1 a part of the test chip at the gate level is shown after the polySi gate etch. The inset in Fig. 1 shows an enlargement of a polySi meander, which is well defined. No polySi residues or damage to the Pr_2O_3 dielectric in the open areas has been observed. Also, within the uncertainty of the thickness measurements no thinning of

the praseodymium oxide was detected, confirming the high selectivity of the RIE process.

4.2. HF-CV characteristics

After wet resist strip and cleaning, the gate stack was briefly annealed at 900°C, 10 s in N_2 using RTA in order to activate the gate doping. Subsequently, electrical measurements on n^+ polySi/Pr_2O_3/Si capacitors were performed. In Fig. 2 an example of the HF-*CV* characteristic is shown.

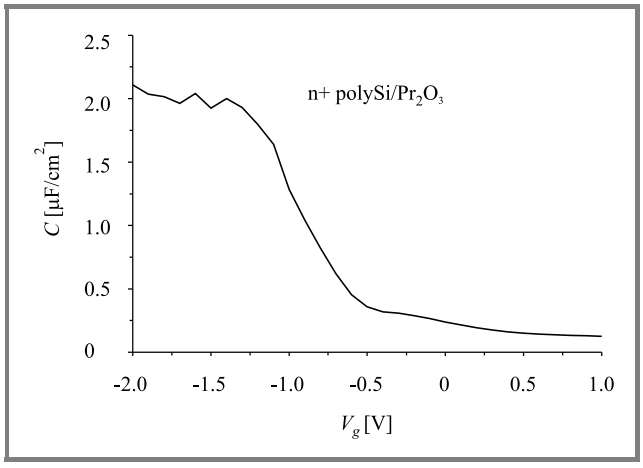


Fig. 2. HF-*CV* curve measured at 10 kHz to minimise serial resistance effects.

No evidence for hysteresis effects was noted when changing polarity of the sweep voltage within this bias range. From the accumulation capacitance of the HF-*CV* curves the effective dielectric constant was calculated $K_{Pr_2O_3} = 36$, using the physical film thickness values measured by ellipsometry. The corresponding equivalent oxide thickness EOT = $d_{\text{Film}} * (K_{SiO_2}/K_{Pr_2O_3})$ was determined as EOT = 1.8 nm, in this case without QM corrections.

4.3. Leakage characteristics

In Fig. 3 an example of the current-voltage characteristics is shown. For gate voltages below $|-3|$ V, which correspond to an equivalent field strength of approximately 22 MV/cm, the measured leakage currents are around 10 pA, which is given by the detection limit of the measurement system. The upper limit of the current density in this bias range is estimated to be $\approx 10^{-6}$ A/cm^2, in good agreement with previous results on simple gold-dot capacitors [5]. For further increased gate bias, a tunneling current starts to appear. Directly after the first sweep, a second sweep was performed. As evident from the shift in the *I-V* characteristics, significant charge trapping has occurred. The dielectric breakdown takes place at a gate bias of $|-7.5|$ V, which corresponds to a breakdown field of approximately 41 MV/cm, also in agreement with previous observations of gold-dot capacitors. However, in contrast

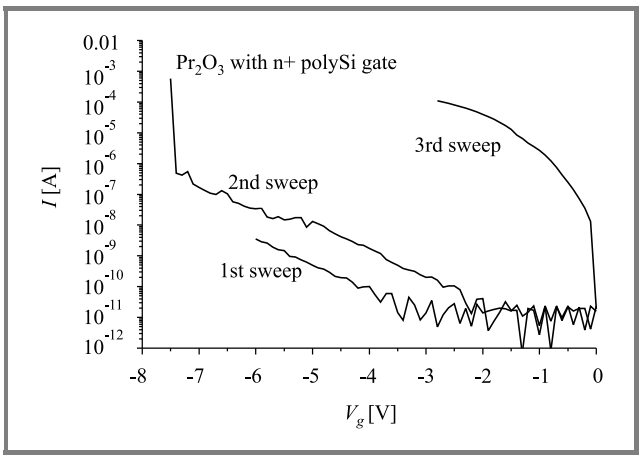


Fig. 3. Example of the current-voltage characteristics of the n^+ polySi/Pr_2O_3 MOS capacitors.

to previous findings, the breakdown was found irreversible, as evident from the 3rd sweep shown in Fig. 3.

4.4. Interface trap density and oxide charge

Admittance measurements were performed to obtain the equivalent parallel conductance G that can be used to get information on the interface trap density D_{it}. For the virgin MOS capacitor a value of $D_{it} = 3.5 \cdot 10^{11}$ /cm^2 eV is deduced from the peak conductance shown in Fig. 4. Compared to high quality device-grade SiO_2 gate oxides, this value corresponds to an increase of at least 2 orders of magnitude and is expected to degrade MOSFET device characteristics. Clearly, a substantial amount of interface engineering and optimization is needed in order to achieve the same quality level as for SiO_2.

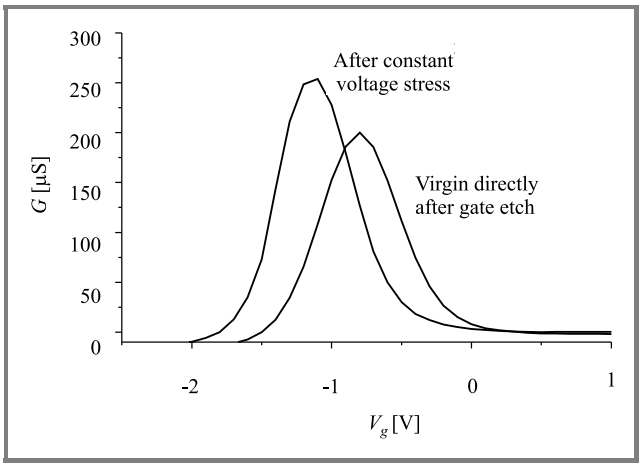


Fig. 4. Results of conductance measurements performed on virgin and electrically stressed devices.

When the Pr_2O_3 capacitor was stressed with 35 MV/cm for 10 s, an increase in interface trap density to $D_{it} = 6 \cdot 10^{11}$ /cm^2 eV was observed (Fig. 4). In addition, the peak position is shifted to further negative values, indicating

the build-up of positive trapped charge on the order of $4 \cdot 10^{12}$ q/cm^2. In this respect, the Pr_2O_3 dielectric behaves similarly to SiO_2, although the detailed mechanisms are probably quite different and require further investigation.

5. Conclusion

The crystalline high-*k* Pr_2O_3 material was found to be compatible with n^+ polySi gate CMOS processing when appropriate etching and cleaning procedures are used and thermal processing is restricted to RTA. We therefore conclude that a complete re-engineering of the CMOS manufacturing process may not be necessary in this case. However, the electrical characteristics of the Pr_2O_3 MOS structures reveal very high values of the interface trap density. Clearly, a substantial amount of interface engineering and optimization is needed in order to achieve the same quality level of state-of-the-art device-grade SiO_2. This is especially important in view of the fact that an enormous knowledge exists about the SiO_2/Si system used and continuously optimised in semiconductor manufacturing for more than 30 years.

References

[1] U. Schwalke, A. Gschwandtner, G. Innertsberger, and M. Keber, "Through-the-gate-implanted ultrathin gate oxide MOSFETs with corner parasitics-free shallow-trench-isolation", *IEEE EDL*, vol. 20, no. 7, 1999, p. 363.

[2] The International Technology Roadmap for Semiconductor, Semiconductor Industry Association, 1999, http://public.itrs.net/Files/2003ITRS/Home2003.htm

[3] A. Kerber, E. Cartier, R. Degraeve, L. Pantisano, Ph. Roussel, and G. Groeseneken, "Strong correlation between dielectric reliability and charge trapping in SiO_2/Al_2O_3 gate stacks with TiN electrodes", in *VLSI Symp. Technol.*, Honolulu, USA, 2002.

[4] P. J. Chen, E. Cartier, R. J. Carter, T. Kauerauf , C. Zhao, J. Petry, V. Cosnier, Z. Xu, A. Kerber, W. Tsai, E. Young, S. Kubicek, M. Caymax, W. Vandervorst, S. De Gendt, M. Heyns, M. Copel, P. Bajolet, and J. Maes, "Thermal stability and scalability of Zr-aluminate-based high-*k* gate stacks", in *VLSI Symp. Technol.*, Honolulu, USA, 2002.

[5] H. J. Osten, J. P. Liu, P. Gaworzewski, E. Bugiel, and P. Zaumseil, "High-*k* gate dielectrics with ultra-low leakage current based on praseodymium oxide", *Tech. Dig. IEDM*, 2000, pp. 653–656.

[6] H. J. Osten, E. Bugiel, J. Dąbrowski, A. Fissel, T. Guminskaya, J. P. Liu, H. J. Müssig, and P. Zaumseil, "Epitaxial praseodymium oxide: a new high-*k* dielectric", in *Proc. Int. Worksh. Gate Insulat.*, Tokyo, Japan, 2001, pp. 100–103.

Udo Schwalke was awarded a Research Fellowship from the Alexander-von-Humboldt-Foundation for his Ph.D. thesis in 1984. During 1984–1986, he was appointed Caltech Research Fellow at the California Institute of Technology, USA. In 1987, he joined the Siemens AG, R&D Microelectronics in Munich, Germany. From 1990–1992 he was responsible for the 64 Mb DRAM device design at IBM/Siemens, USA. Subsequently, at Infineon Technologies AG (the former Siemens Semiconductor Group) he worked on novel process architectures and device reliability issues. Since August 2001, Dr. Schwalke is Professor in electrical engineering and the Managing Director of the Institute of Semiconductor Technology at the Technical University of Darmstadt. He has authored and co-authored more than 60 papers, holds several patents and is a member of IEEE and ECS.

e-mail: schwalke@iht.tu-darmstadt.de
Institut für Halbleitertechnik (IHT)
Technische Universität Darmstadt
Schlossgartenstr. 8
64289 Darmstadt, Germany

Invited paper

Understanding of wet and alternative particle removal processes in microelectronics: theoretical capabilities and limitations

François Tardif, Adrien Danel, and Olivier Raccurt

Abstract—A 2 orders of magnitude range of van der Waals interactions is considered here to take the majority of the variety of shapes and materials of actual particles into account. Comparing these interactions with the repulsive forces generated by electrostatic charges, drag, surface tension, shock waves, high accelerations and aerosol particles, the intrinsic capabilities and limitations of the different cleaning processes can be predicted. Three kinds of particle-removal processes have been identified – universal processes capable of removing all particle sizes and types, even from patterned wafers, processes that present the same theoretical ability but are actually limited by the accessibility of the particles, and finally cleanings that are not able to remove all particle sizes.

Keywords— cleaning, SC1, particle removal.

1. Introduction

The continuous increase of IC integration density requires a reduction of both device dimensions and the corresponding amount of the material used. Consequently, the concentration of contaminants affecting the fabrication yield of very competitive microelectronic manufacturing is becoming smaller and smaller. Therefore, cleaning performance has to be continuously improved to remove the ultimate traces of contamination, such as particles, metals, organics, bases and anions.

According to the *International Technology Roadmap for Semiconductors* [1], smaller and smaller particle sizes will have to be eliminated as the device dimensions decrease: 50 nm as from 2004 and 10 nm in 2016.

Particles will probably be the most challenging type of contamination in the near future as the removal mechanism in the conventional Standard Clean 1 (SC1) [2] process is mainly based on a controlled consumption of the layer under the particle. But this consumption will be rapidly prohibited as the accuracy of the device dimensions (implantation, silicon on insulator, etc.) are now approaching the under-etching thicknesses required for particle removal.

In this work, fundamental particle-substrate interactions due to van der Waals, drag and surface tension forces, as well as electrostatic charges, are used to understand the intrinsic capabilities and limitations of the wet SC1 process and emerging new techniques such as explosive evaporation, high velocity sprays, acoustic waves, laser and cryogenic techniques.

2. Forces acting on particles

The main forces likely to be exerted on fine particles are calculated in Table 1. It can be seen from this table that the main four parameters that drive the particle adhesion/removal mechanisms are the electrostatic, van der Waals, capillary and drag forces [3].

Table 1
Orders of magnitude of the different forces acting on a 100 nm spherical particle in a solution with the density of 1

Forces	Order of magnitude (N)	Proportionality (R: Radius)
Van der Waals	10^{-7}	R
Electrostatic	10^{-8}	–
Capillary	10^{-8}	R
Drag (water, 10 m/s)	10^{-9}	R
Gravitation	10^{-16}	R^3
Archimedes	10^{-17}	R^3
Hydrostatic	10^{-21}	R^3

2.1. Capillary forces

The surface tension γ_{lg} is due to the cohesion between the molecules of the media and tends to minimize the interfacial areas. It represents a force per unit of interfacial length. In the case of the reference spherical particle, the maximum capillary force is obtained when the liquid wets the particle material perfectly and the gas/liquid interface is acting on the whole particle perimeter (see Fig. 1):

$$F_\gamma = 2\pi R\gamma_{lg}\,. \tag{1}$$

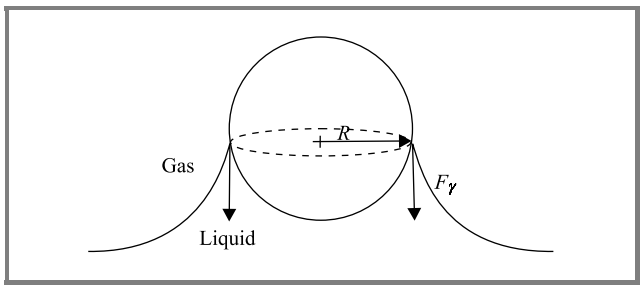


Fig. 1. Schematic of the maximum capillary force acting on a particle at a liquid/gas interface (case of contact angle = 0).

2.2. Drag forces

The viscosity of the moving fluids induces a drag force on the particles. In the case of a spherical particle of diameter D placed in a flow velocity V_p, this force is given with a good approximation [4] by the Stokes law up to the Reynolds number of ten[1]. In the case of a particle deposited on a surface, an additional constant of 1.7 accounting for the effect of the surface must be added [5]:

$$F_{Drag} = 1.7 \cdot 3\pi\,\mu\, D V_p\,, \tag{2}$$

where μ represents the fluid viscosity: 10^{-3} kg/m·s for water at 20°C.

This force is theoretically able only to push the particle in parallel to the surface. It can be expected that an asperity on the particle or on the substrate will transform this tangential force to a lift-off momentum [6]. In the case of non-spherical particles, the drag forces are generally higher.

2.3. Van der Waals forces

The particle-substrate attractions due to the van der Waals forces result from the dipole/dipole interactions between their constitutive molecules. These forces are so high that the particles are generally flattened on the substrate. The integration of the interactions between all the constitutive volume elements of a spherical particle or an infinite flat particle both on a perfectly flat substrate are given by (3) and (4), respectively:

$$F_{vdW} = \frac{AR}{6h^2}\,, \tag{3}$$

$$F_{vdW} = \frac{AS}{6\pi h^3}\,, \tag{4}$$

where: R – particle radius, h – particle-substrate distance (the minimum distance equal to the Lennard-Jones distance of $h_0 = 0.4$ nm for the considered materials), A – the Hamaker constant (depends on the particle, substrate materials and on the nature of the media: interaction transmission), S – facing particle and substrate surface.

The Hamaker constants for different particle/substrate materials can be calculated using the data and the formula given by Israelachvili [7]. The results listed in Table 2 show that van der Waals forces for typical particle materials on SiO_2 substrates can vary by one order of magnitude depending on the media.

Table 2
Hamaker constants in water and in air calculated from reference [7]

Media	Al_2O_3/SiO_2	SiO_2/SiO_2	PSL/SiO_2
Water	$1.6 \cdot 10^{-20}$ J	$6.5 \cdot 10^{-21}$ J	$1.0 \cdot 10^{-20}$ J
Air	$9.6 \cdot 10^{-20}$ J	$6.3 \cdot 10^{-20}$ J	$7.5 \cdot 10^{-20}$ J

The difference between van der Waals attractions acting on the reference rigid and spherical particle and on actual particles is investigated here. The consequence of the non ideality of the actual particles: flattening, non specific shape, roughness, partially embedded, etc., can finally be considered as an additional flat surface in contact with the substrate.

In this work, the difference between the ideal and rigid sphere and the actual non-ideal particle is arbitrarily expressed by the fraction f equal to the surface in contact divided by the maximum surface that a particle of the same dimension is able to present πR^2:

$$f = \frac{S}{\pi R^2}\,. \tag{5}$$

In this way, f is null for an ideal particle and reaches 1 for a particle presenting the maximum surface in contact.

For a quasi-spherical particle presenting a small flat contact area, the total van der Waals forces can be considered as the sum of the contributions from the non-deformed particle and the contact surface [8] with the latter dominating the contribution from the spherical particle. Therefore, the ratio $\mathfrak{R}$ between the van der Waals attraction of the actual and the ideal (rigid and spherical) particles can be approximated by the equation:

$$\mathfrak{R} = \frac{\dfrac{AR}{6h^2} + \dfrac{AS}{6\pi h^3}}{\dfrac{AR}{6h^2}} = 1 + f\,\frac{R}{h}\,. \tag{6}$$

The results have been plotted in Fig. 2 for particle sizes ranging from 10 to 150 nm. It can be seen that the van der Waals forces increase rapidly with the non-ideality of the particles and can be more than 2 orders of magnitude higher than those for the ideal particle. This effect decreases with the particle size. Finally, due to a large variation range

[1]For a 100 nm particle, the Reynolds number of 10 corresponds to a water flow velocity of 100 m/s!

of the Hamaker constant and considerable impact of the particle, van der Waals interactions can vary to a very large extent for the different existing particle types.

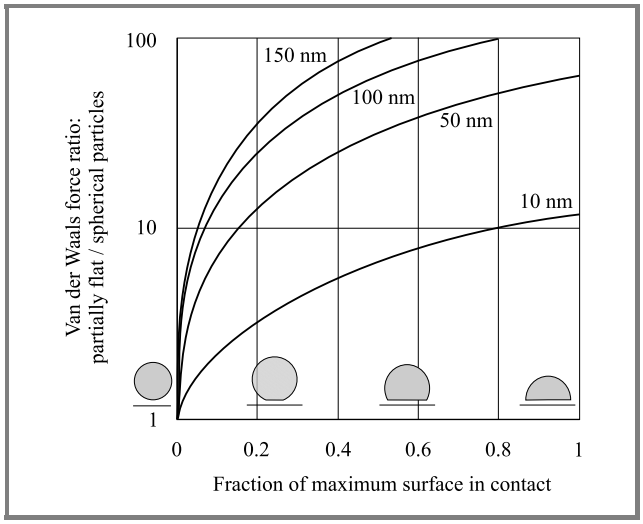


Fig. 2. Ratio $\Re$ between van der Waals forces calculated for the reference rigid sphere and particles presenting a flat surface in contact. This surface is given in percentage of the area of the sphere cross section which represents the maximum contact surface for a given particle size.

In this work, we arbitrarily chose to consider the van der Waals forces present in the system consisting of a rigid sphere of Al_2O_3 on SiO_2 substrate because its Hamaker constant is the highest of those listed in Table 2. In order to account for a shift from ideality, a range of the Hamaker constants of up to 2 orders of magnitude above the value mentioned above is considered. This approach enables a very wide majority of actually observed particles to be covered but clearly does not take into account the extreme cases of flat-shape particles made of materials exhibiting very high Hamaker constants.
The substrate roughness is not considered here as it generally decreases the contact areas leading to lower van der Waals interactions.

2.4. Force originating from electrostatic charges

Material surfaces usually present electrostatic charges that originate from ionization or dissociation of functional surface groups in chemical equilibrium with H^+ ions from the media (pH). In a liquid, a large number of charges are available close to the surface[2]. Ions carrying charge of the sign opposite to that of the surface (the counter ions) are immediately attracted to the surface, masking its surface potential until apparent neutrality is reached. As shown in Fig. 3, the surface charge density is characterized by the Zeta potential. This potential is due to the contributions from the particle charges and the retinue of counter ions sufficiently attached to the particle surface when it moves against the liquid (shear layer), e.g., under the influence of an electric field.
The thickness of the diffuse layer results from the competition between the electrostatic attraction exerted on the counter ions that build up at the charged surface and their re-diffusion to the bulk solution. A high temperature and a low ionic strength, therefore, enhance the diffuse layer thickness.

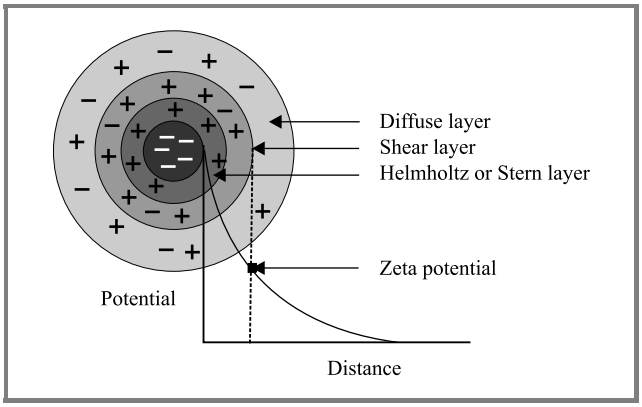


Fig. 3. Representation of a negatively charged particle dipped in an electrolyte.

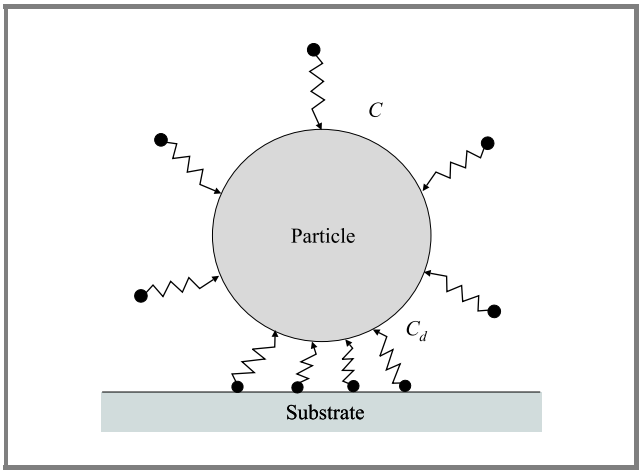


Fig. 4. Representation of a negative particle dipped in an electrolyte.

In fact, at very small particle/substrate distances, the major interactions caused by the charged surfaces are not due directly to the electrostatic forces but to the entropic contribution [9]. As shown in Fig. 4, the counter ion concentration is very high in between the particle and the substrate due to the overlap of the particle and substrate diffuse layers. This leads to the differential pressure between the top and bottom of the particle, expressed below:

$$F_e = (C_d - C)\,\mathrm{k}T\,, \tag{7}$$

where: k – the Boltzmann constant, T – absolute temperature, C – concentrations.

[2]Even in the case of ultra pure water, there is a sufficient reservoir of H^+ and OH^- of 10^-7 mole/L.

Figure 5 shows the evolution of the Zeta potential versus pH for Si and SiO_2 as substrate materials and one of the most electropositive particle materials – alumina.

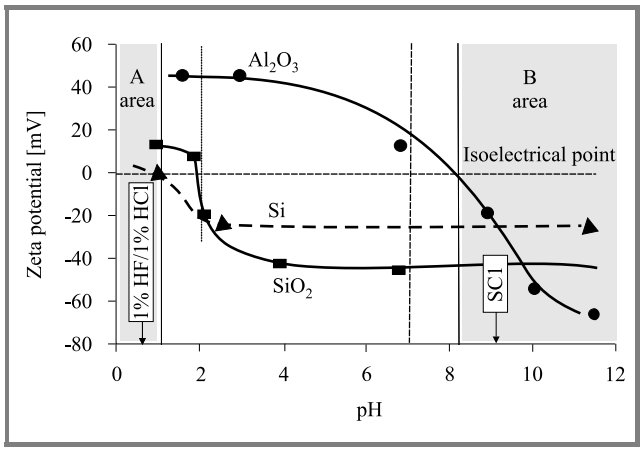


Fig. 5. Evolution of Al_2O_3, Si and SiO_2 Zeta potentials versus pH.

In order to facilitate particle removal and to prevent any re-deposition it is to be hoped that charges gathered on all particle types and on the substrate are of the same sign leading to electrostatic repulsion. This condition is fulfilled in both A and B areas.
The forces originating from electrostatic charges can be calculated by solving the Poisson-Boltzmann equation [7]. They are favored by:

– high absolute values of surface potential (same electrical sign), the area B is generally better in this respect than the area A;

– high ionic force (densification of the diffuse layer increasing the entropic force);

– high temperature (kT factor in Eq. (7)).

3. Particle removal mechanisms

Theoretical performances of conventional and prospective particle cleaning processes are discussed here.

3.1. Cleaning by etching and electrostatic repulsion

This cleaning mechanism consists in separating the particle from the substrate by consuming the substrate, the particles, or both, until the repulsive forces of electrostatic origin exceed the van der Waals forces. This means that the pH has to be adjusted in the area A or B. As shown in the example of Fig. 6, electrostatic forces decrease more slowly with the distance than van der Waals interactions. Therefore a liberation distance always exists regardless of the particle size and charge (theoretically 1.6 nm in the case of an ionic force of 0.5 M in Fig. 6).

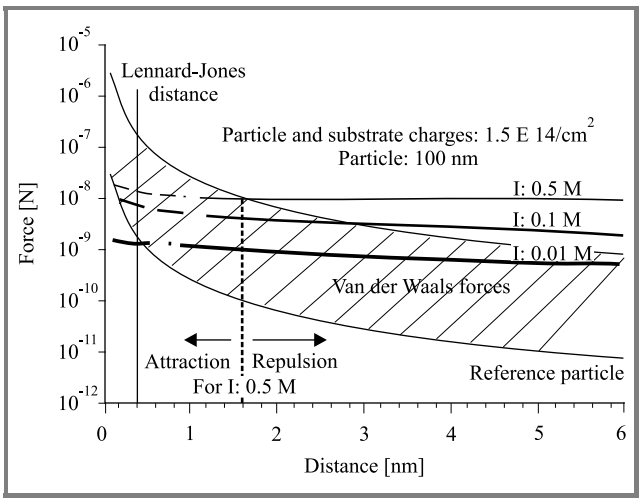


Fig. 6. Comparison between the order of magnitude of van der Waals forces (hatched area) and forces of electrostatic origin at different ionic strengths as a function of the separation distance. Calculations of van der Waals forces for distances lower than 0.4 nm and calculations of electrostatic forces for distances lower than 2 nm are not valid.

This theoretical etching thickness is increased in practice by the dynamic behavior of the removal process. Indeed, at the beginning of the separation, a competition occurs between the etching speed and the re-attraction speed of the particle due to van der Waals interactions (see Fig. 7).

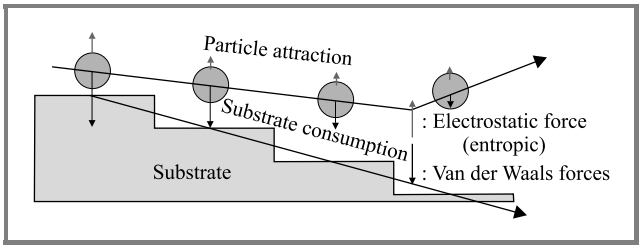


Fig. 7. Illustration of the dynamic behavior of the particle removal process by etching and electrostatic repulsion.

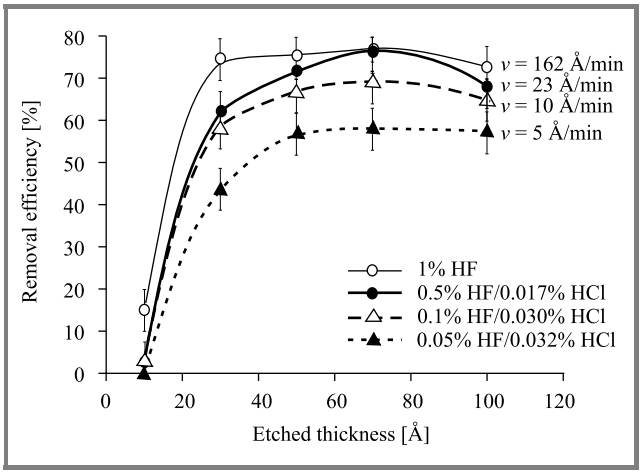


Fig. 8. Cleaning of 3000 SiO_2 particles deposited on a 200 mm SiO_2 substrate in HF solutions exhibiting different etching speeds (ionic strength and pH constant: 0.024 M and 1.4, respectively).

This phenomenon is verified in Fig. 8. When the etching speed is low, a higher material consumption is necessary to reach the nominal removal efficiency that is also lower. Particle removal processes, such as SC1 [2], IMEC clean [10], DDC [11, 12], etc., therefore have to present very fast etching kinetics. This requirement is compatible with the fast processes necessary for the new single wafer cleaning tools.

Particle removal by etching and electrostatic repulsion does not seem to present any limitation in terms of particle size since the necessary consumption of the material is acceptable. In practice, this amount can be limited by increasing the etching speed, the ionic force of the solution and absolute values of Zeta potentials.

3.2. Cleaning by drag forces

As shown in Fig. 9, drag forces induced by a continuous liquid jet are able to sweep along even very small particles. Nevertheless, very high pressures (50 bars) have to be used to have a chance of removing all particle types with a micro spray. In this calculation, we assume as the initial approximation that the thickness of the laminar boundary layer is zero just under the jet impact and that the order of magnitude of the drag force is approximated by (2). The jet therefore has to scan the whole wafer surface.

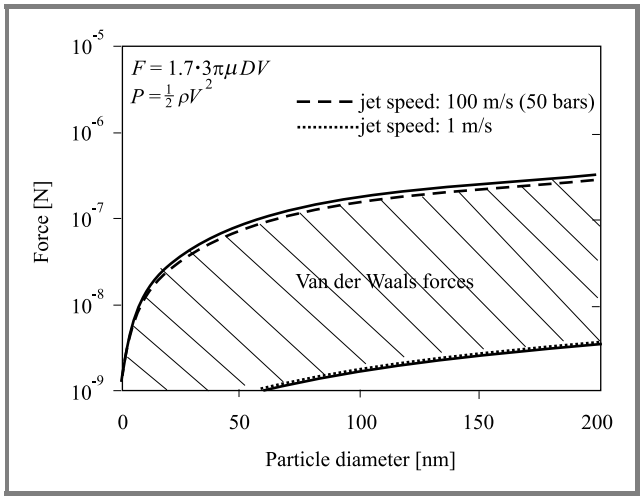


Fig. 9. Van der Waals and drag forces.

This cleaning method does not seem to be suitable for patterned wafers where particles are not accessible to the jet.

3.3. Cleaning by shock waves

The instantaneous overpressure induced by a shock wave (water hammer) on the cross section area of a particle generates a force likely to overcome the attractive van der Waals forces according to:

$$F_{shock} = \frac{\pi D^2}{4} \rho\, c V\,, \tag{8}$$

where: D – particle diameter, ρ – mass density of the media, c – wave velocity in the media ($c = 1500$ m/s in water), V – speed of the media versus the particle.

Droplet jet. The particle removal process using a jet of droplets called "Soft Spray" has recently been proposed. It consists in spraying a mixture of liquid and gas onto

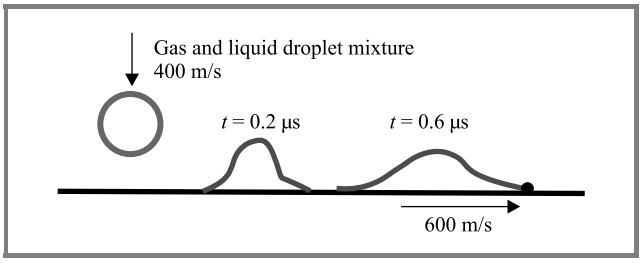


Fig. 10. Evolution of high velocity liquid droplet crashing at a perpendicular surface. Calculations taken from [13].

the wafer leading to a very high velocity heterogeneous jet. The generated droplet is projected onto the substrate with speeds in the range of 400 m/s. As represented in Fig. 10 calculations show that, when crashing, the droplet front even accelerates and strikes the particle at a speed of about 600 m/s [13]. Unlike a continuous jet, the exerted force is generated here by the shock wave from the droplet front applied to the particle surface.

Resonant acoustic cavitation. Cavitation in a liquid is due to the implosion of μ-bubbles after the loss of the equilibrium pressure conditions between the inside and outside the bubble:

$$P_{in} - P_{out} = \frac{4\gamma_{lg}}{D}\,, \tag{9}$$

where: γ_{lg} – liquid/gas surface tension, D – bubble diameter.

The density of the energy liberated during bubble collapse is considerable as temperatures of 3000 K and pressures of 1000 atm are reached very locally. The bigger the bubble the higher the potential liberated energy. When a bubble collapses close to the surface, it can induce a microjet of liquid toward the surface that can reach a very high velocity of many hundreds of meters per second. This jet produces a very intense local shock wave. This phenomenon has been observed with a high-speed camera on bubbles in the mm range [14] and by nanosecond electrochemistry [15] with bubbles generated in an ultrasonic bath at 20 kHz. Acoustic cavitation has been observed in the megasonic range – up to 850 kHz – by sonoluminescence [16]. Nevertheless, it is not possible to conclude that the shock waves produced with megasonic frequencies that generate smaller bubbles are also induced by the same jet phenomenon.

As shown in Fig. 11, the periodic pressure wave variations generated in a sonic bath tend alternatively to increase and

decrease the bubble diameter to satisfy (9). Bubbles are initially present as germs in the media. When a bubble grows, its exchange surface increases, making the desorption of the dissolved gas present in the supersaturated liquid easier. In this way, during the successive pressure cycles, the bubbles grow until they reach the resonant size that depends on the acoustic wave frequency.

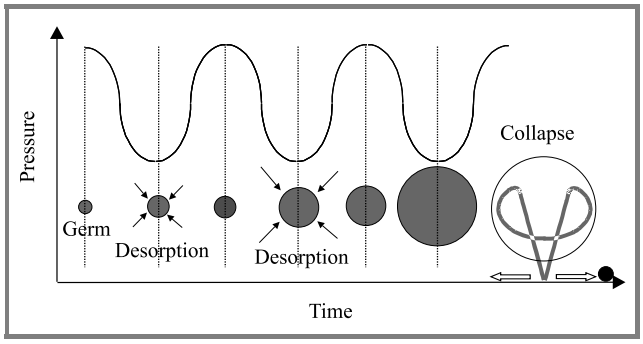


Fig. 11. Evolution of bubble sizes in an ultrasonic excitation until collapse (experimentally observed at the resonant size of the bubbles).

In water, neglecting the effects due to surface tensions and considering the transformations as adiabatic, the resonant frequency f_0 of bubbles is proportional to their radius R with a good approximation:

$$R_{[m]} f_{0[Hz]} \approx 3.26\,. \tag{10}$$

Thus $R = 75$ μm at 40 kHz and 3 μm at 1 MHz.

At this frequency, the oscillation speed of the bubbles is maximum leading to the well known resonant cavitation phenomena observed in ultrasonic baths [17].
The megasonic efficiency strongly depends on the concentration and nature of the dissolved gas. High quantities of poorly soluble gas seem to be favorable for particle removal [18]. Acoustic wave transmission in the media is limited by the presence of big bubbles. Acoustic waves generate different streamings that prevent reattachment by carrying the removed particles far from the surface [19]. Extensive efforts still have to be made to understand the actual removal mechanisms occurring in megasonic baths.

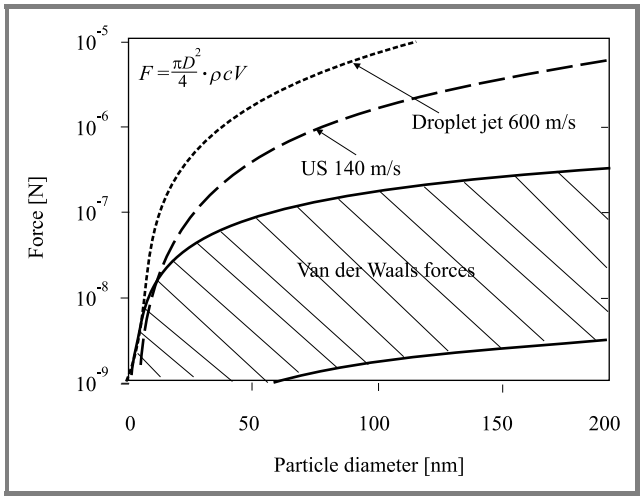


Fig. 12. Van der Waals and shock wave forces due to droplet jets and ultrasonic waves.

Cavitation is able to overcome van der Waals forces and, unfortunately, even to deteriorate the quality of the material. Using higher frequencies leads to smaller bubbles and consequently lower energies, which partly prevents material and pattern degradation but theoretically also decreases the removal forces for cleaning particles (limitation for the biggest size). Using a lower acoustic power leads to the same effect.
As shown in Fig. 12, the shock waves are theoretically able to remove all types and sizes of particles.

3.4. Cleaning using capillary forces

Capillary forces can potentially remove particles when they are located at the liquid/gas interface. This configuration is achieved, for example, during fast evaporation of a liquid phase or when wafers cross the liquid/gas interface of a bath.

Fast evaporation. This consists first in depositing a liquid medium at the wafer surface and then evaporating this liquid phase very quickly by decreasing the pressure (Fig. 13). The last fragments of the liquid can pull the particles off by the capillary force (or by simple mechanical drive). Different fluids have been envisaged, such as H_2O, CO_2, NH_3, ...

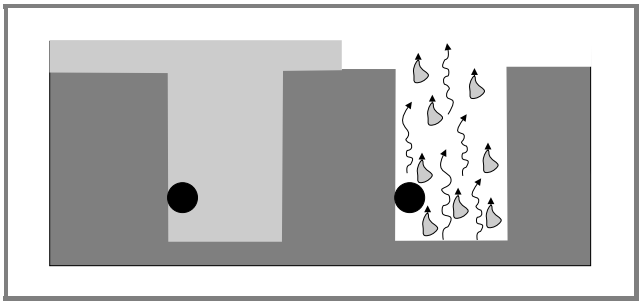


Fig. 13. Illustration of the cleaning principle by fast evaporation.

Bath interface. A.F.M. Leenaars [20] studied the capillary forces acting on a particle attached to a vertical substrate and located at the meniscus level of the air/liquid interface of a bath. In the particular case depicted in Fig. 14 (the liquid wets the substrate and not the particle), the maximum force is given by (11):

$$F_\gamma^{\max} = 2\pi R \gamma_{lg} \sin^2\left(\frac{\theta}{2}\right) \cos\alpha\,. \tag{11}$$

As seen in Fig. 15, theoretically, the capillary forces are not able to remove all types of particles, even in the favorable case of particles perfectly hydrophilic in water (a case of high surface tension).

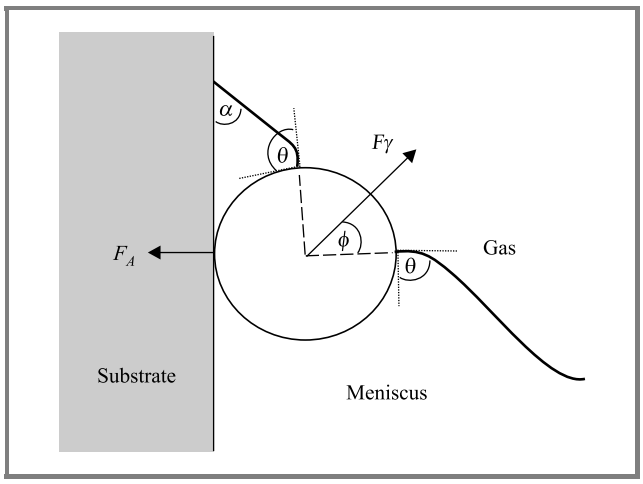


Fig. 14. Forces exerted on a particle attached to a substrate located at a gas/liquid interface.

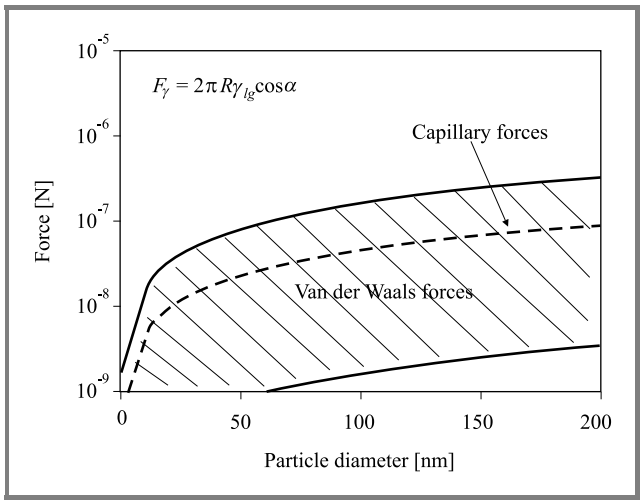


Fig. 15. Evolution of the van der Waals and maximum capillary forces ($\cos\alpha = 1$, $\gamma_{lg} = 0.072$ N/m).

3.5. Cleaning using high acceleration

A very high acceleration due to the thermal expansion of the substrate and/or the particles heated up with a laser beam is likely to remove particles [21]. In this case, the force exerted on a particle of mass m is, in the first approach, given simply by:

$$F = m\gamma. \tag{12}$$

In practice, the acceleration is limited by the acceptable laser fluence leading to the melting of silicon. This threshold corresponds experimentally to the removal of the first alumina particles of about 100 nm (optimistic scenario). The corresponding acceleration calculated using Eq. (12) is in the 10^6 g range and thus higher than the one measured experimentally by Dobler *et al*. [22]. Anyway, as shown in Fig. 16, this method is not suitable for removing fine particles. To improve the removal capability of laser cleaning, a thin layer of liquid is first condensed from steam onto the substrate. In this case, the cleaning

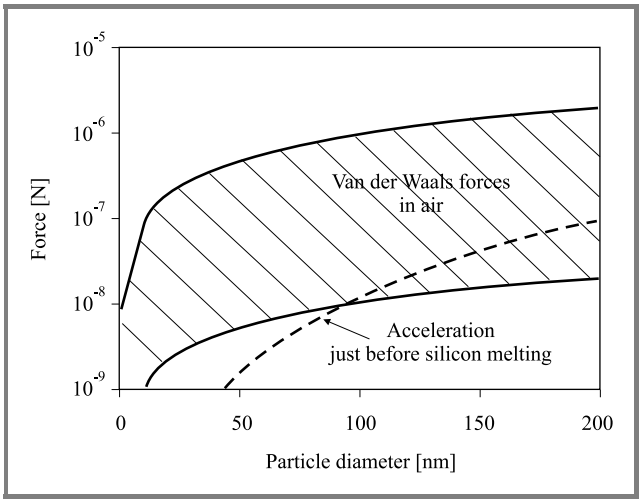


Fig. 16. Evolution of the van der Waals forces and the force generated by an acceleration able to remove the first 100 nm alumina particles, in air.

mechanism proposed by [21] would be close to the phenomenon of the cavitation process depicted above and consequently to its performance.

3.6. Cleaning by kinetic energy

In 1988, researchers at IBM Watson Research Center began to study cryogenic particle removal. This process uses the kinetic energy of a distribution of solid aerosol particles obtained for example by expansion cooling of gas such as Ar, N2, etc., [23]. This aerosol is then eliminated by sublimation. The particles just liberated are evacuated far from the substrate by thermophoresis or by a gaseous flow. If we consider that only one aerosol particle of mass m and velocity V reaches the particle at any one time, the removal condition is given by:

$$\int_{h_0}^{\infty} \frac{AR}{6h^2}\,\mathrm{d}h = \frac{AR}{12h_0} = \frac{1}{2}mV^2. \tag{13}$$

Collective effects may however occur.

In order to be able to remove particles located in lines and vias, the aerosol particles have to be smaller than the pattern dimensions.

4. Conclusion

In this work a range of van der Waals interactions covering 2 orders of magnitude is considered to take into account the majority of the variety of shapes and materials of the actual particles. This range has been determined by considering

the possible variations of the Hamaker constant and the difference between the ideal rigid sphere and real particles presenting finite contact areas, flattening effects, etc.
The different particle removal processes can be classified according to the physical effects used, such as electrostatic, drag and capillary forces, shock waves, acceleration, or kinetic energy. By comparing the attractive van der Waals forces and those generated by these effects it is then possible to predict the intrinsic capabilities and limitations of the different cleaning processes, particularly for the fine particles that have to be considered for the next IC generations. Three kinds of particle removal processes have been identified, namely universal processes able to remove all particle sizes and types even from patterned wafers, processes that present the same theoretical ability but are actually limited by the accessibility of the particles, and finally cleanings that are not able to remove all particle sizes.

- Particle removal by etching and electrostatic repulsion is the process used the most often through the SC1 cleaning step. This method does not seem to present any limitation in terms of particle size since the necessary material consumption is acceptable. In practice, this amount can be limited by increasing the etching speed, the ionic force of the solution and absolute values of Zeta potentials. It is also possible to remove all particle types by shock waves generated for example by megasonics in aqueous media. Nevertheless this method is limited by the erosion of the materials and by the mechanical resistance of the microstructures.

- High-speed aqueous jets, droplet jets, and aerosol sprays are theoretically able to remove all particle types accessible to the jet. Limitations arise from the mechanical resistance of the patterns and for particles hidden in the microstructures.

- Methods using capillary forces and high accelerations are not able to remove all particle types.

References

[1] International Technology Roadmap for Semiconductors, 2001, http://public.itrs.net/Files/2001ITRS/Home.htm

[2] W. Kern and D. Puotinen, "Cleaning solutions based on hydrogen peroxide for use in silicon semiconductor industry", RCA rev. 31, p. 187, 1970.

[3] F. Tardif, *Nettoyages par voie humide en micro-électronique dans l'ouvrage "Procédés de fabrication en micro-électronique"*. Hermes Science Publ., 2003.

[4] R. Comolet, *Mécanique expérimentale des fluides*. Tome II, 3rd ed. Paris: Masson, 1982.

[5] M. O'Neill, "A sphere in contact with a plane wall in a slow linear shear flow", *Chem. Eng. Sci*., vol. 23, p. 1293, 1968.

[6] G. Ziskind, M. Fishman, and C. Gutfinger, "Resuspension of particulates from surfaces to turbulent flows. Review and analysis", *J. Aerosol Sci*., vol. 26, pp. 613–644, 1995.

[7] J. Israelachvili, *Intermolecular & Surface Forces*. 2nd ed. Academic Press, 1997.

[8] P. C. Hiementz, *Principle of Colloid and Surface Chemistry*. 2nd ed. New York: Dekker Press, 1986.

[9] F. Tardif, I. Constant, R. J.-M. Pellenq, and A. Delville, "A new approach for particle removal based on a Monte Carlo simulation of electrostatic interactions", in *7th Int. Symp. Part. Surf*., New York, USA, 2000.

[10] M. Meuris, P.-W. Mertens, A. Opdebeeck, H.-F. Schmidt, M. Depas, G. Vereecke, M.-M. Heyns, and A. Philipossian, "The IMEC clean: a new concept for particle and metal removal on Si surfaces", *Solid-State-Technol*., vol. 38, no. 7, pp. 109-110, 112, 114, 1995.

[11] F. Tardif, T. Lardin, C. Paillet, J. P. Joly, A. Fleury, P. Patruno, D. Levy, and K. Barla, "Optimization of HF and oxidant wet cleanings before 7 nm gate oxide: introduction to "DDC": diluted dynamic clean", in *ECS*, Chicago, USA, 1995.

[12] F. Tardif, T. Lardin, B. Sandrier, P. Boelen, R. Matthews, I. Kashkoush, and R. Novak, "Performances of "DDC": diluted dynamic clean before 4.5 nm gate oxide", in *ECS*, Paris, France, 1997.

[13] N. Hirano, K. Takayama, J. Falcovitz, T. Katoaka, K. Shimada, and E. Ando, "Microscopic analysis of particle removal by gas/liquid mixture high-speed flow", in *UCPSS'98*, Oostende, Belgium, 1998.

[14] J. C. Isselin, A. P. Alloncle, and M. Autric, "On laser induced bubble near a solid boundary: contribution to the understanding of erosion phenomena", *J. Appl. Phys*., vol. 84, no. 10, pp. 5766–5771, 1998.

[15] E. Maisonhaute, P. C. White, and R. G. Campton, "Surface acoustic cavitation understood via nanosecond electrochemistry", *J. Phys. Chem. B*, vol. 105, pp. 12087–12091, 2001.

[16] D. Zhang, "Fundamental study of megasonic cleaning". Ph.D. thesis, University of Minnesota, 1993.

[17] L. H. Thompson and L. K. Doraiswamy, "Sonochemistry: science and engineering", *Ind. Eng. Chem. Res*., vol. 38, pp. 1215–1249, 1999.

[18] P. Besson, O. Keller, and G. Ching, "Particle removal evaluation in DI water with megasonic activation", in *Techn. Conf., Semicon. Eur.*, Munchen, Germany, 2002.

[19] A. Busnaina, J. Taylor, and I. Kashkoush, "Measurement of the adhesion and removal forces of sub micrometer particles on silicon substrates", *J. Adhes. Sci. Technol*., vol. 7, p. 441, 1993.

[20] A. F. M. Leenaars, "Particles on surfaces I: detection, adhesion and removal", Ed. K. L. Mittal. New York: Plenum Press, 1988.

[21] X. Wu, E. Sacher, and M. Meunier, "The modeling of excimer laser particle removal from hydrophilic silicon surfaces", *J. App. Phys.*, vol. 87, no. 8, pp. 3618–3627, 2000.

[22] V. Dobler, R. Oltra, J. P. Boquillon, M. Mosbascher, J. Boneberg, and P. Leiderer, "Surface acceleration during dry laser cleaning of silicon", *Appl. Phys. Mat. Sci. & Proc. A*, vol. 69, pp. 335–337, 1999.

[23] J. J. Wu, D. Syverson, T. Wagner, and J. Weygand, "Wafer cleaning with cryogenic argon aerosols", in *Semicon. Int*., Boise, USA, 1996.

François Tardif is the Head of the Ultra Clean Process Laboratory in the Microelectronics Department at CEA-Leti, Grenoble, France. He graduated from the engineering school of l'Institut National des Sciences Appliquées in Toulouse and received the Ph.D. degree in materials sciences from the University of Marseille, Marseille,

France. He is a member of the Electrochemical Society, he has co-organized congresses in the fields of silicon cleaning and ultra trace contamination measurements and has authored more than 50 papers.
e-mail: TardifFr@chartreuse.cea.fr
CEA Leti
17 rue des Martyrs
F-38054 Grenoble cedex 9, France

Adrien Danel received the diploma in engineering in 1994 and the Ph.D. degree in physics and microelectronics from the National Polytechnical Intitute of Grenoble, France, in 1999. He is a research engineer at CEA Leti, working with silicon surface characterization by non-invasive methods and focusing on advanced cleaning processes.
e-mail: DanelAd@chartreuse.cea.fr
CEA Leti
17 rue des Martyrs
F-38054 Grenoble cedex 9, France

Olivier Raccurt received the B.Sc. degree in physics from the Université Claude Bernard, Lyon, France, in 2000, and the M.Sc. degree in materials science from the Institut National des Sciences Appliquées, Lyon, France, in 2001. He received the Ph.D. degree in physics and material science from the Université Aix-Marseille, Marseille, France, in 2004. This work was conducted at CEA Leti, Grenoble, France, and was focused on the mechanisms of stiction of microsystems during releasing step, involving studies on etching processes and cleaning of silicon substrate for micro-electronics, using atomic force microscopy for topological characterisation. He is currently holding a postdoctoral position at CEA Leti, where he is working on electrowetting on dielectric to move small droplets of biological liquid within the framework of a lab-on-chip project. His current research interests include the behaviour of liquid on solid at microscale dimension and atomic force microscopy characterisation.
e-mail: RaccurtOl@chartreuse.cea.fr
CEA Leti
17 rue des Martyrs
F-38054 Grenoble cedex 9, France

Invited paper

Integrated gas chromatograph

Jan A. Dziuban, Jerzy Mróz, and Jan Koszur

Abstract—A portable gas chromatograph with integrated micromachined flushed injector and thermal mass detector (TCD) has been developed. The silicon/glass injector operates in a fixed volume (2×7 μL) or electronically operated mode. An integrated, pneumatically operated, fast cross-valve is applied in the injector. The TCD detector consists of two Pt microheaters and thermoresistors packaged in a silicon/glass micromachined chip. The temperature of two capillary molecular sieve separation columns is controlled by a thick-film heater fabricated on polyimide foil. The chromatograph is equipped with two 16-bits microprocessors communicating with the external portable PC. The instrument may operate in the on-line continuous analysis mode.

Keywords—portable gas chromatograph, integrated injectors, silicon micromechanics.

1. Introduction

In this paper the results of a long-lasting research program on portable gas chromatograph are presented. The program was carried out under the auspices of the State Committee for Scientific Research. The main goal of the research was to obtain a miniature gas chromatograph (GC) with a flushed gas injector and fast, nano-dead-volume TCD. The device is equipped with appropriate software and analog/digital circuits and is capable of ensuring continuous detection of combustible atmosphere in deep coal mining industry and/or continuous monitoring of the environment (long-term air quality, drifts, emission of polluters, etc.).

2. Experiment

The portable gas chromatograph consists of two micromachined parts: a fixed or electronically adjusted gas injector and thermal-mass-detectors. Restek (USA) ID 0.32 mm Rt-Msieve 5A capillary columns, each 10 m long, are positioned in a thermally isolated chamber. A thick-film heater fabricated on elastic polyimid foil surrounds the columns and ensures proper adjustment and stabilization of the temperature (50–120°C) (see Fig. 1).

The injector consists of two Borofloat 33 glass substrates, silicon (100) deep-micromachined substrate polished on both sides, and polyimide foil coated with Teflon®. The gas circuit of the injector includes one integrated fast cross-valve (1 ms) operated pneumatically, two fixed-volume chambers (7 μL each) and two narrow channels for active and reference gas column supplying (Fig. 2). The fabrication process of the injector includes new processes of 3D-structure substrate-by-substrate assembling: selective bonding of glass and silicon substrate using polyimide foil (Fig. 3).

The TCD sensor applied in the GC has been designed and fabricated by the Institute of Electron Technology in Warsaw. In the device a thin-film Pt heater and a thermo-resistor

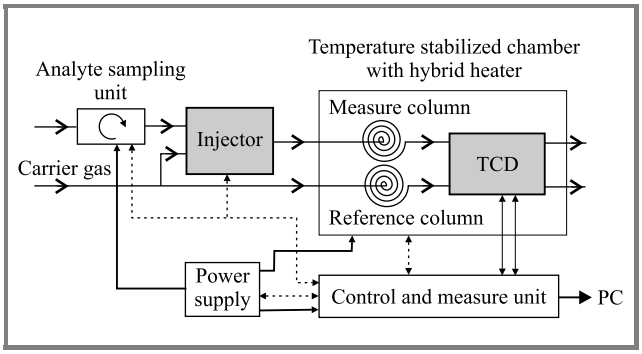


Fig. 1. Block diagram of the gas chromatograph with integrated micromachined components – gas injector and TCD sensor (grey areas).

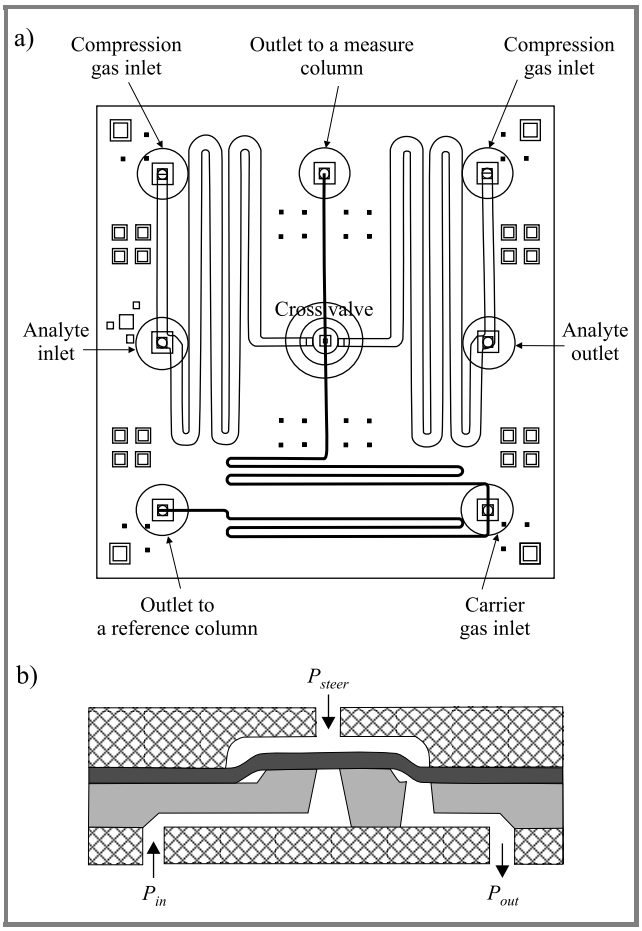


Fig. 2. The injector: (a) layout of the device; (b) cross-section of a microvalve; for $P_{steer} < P_{in}$ the microvalve is closed.

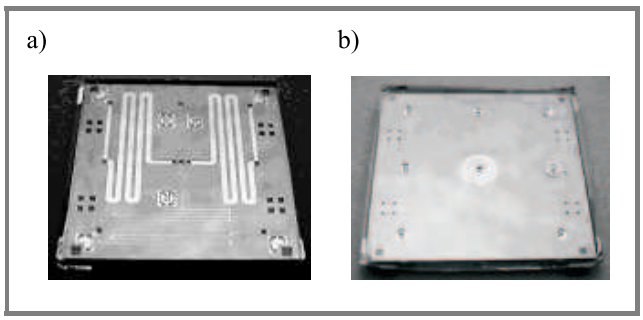


Fig. 3. The injector: (a) front-side: gas channels, fixed volume chambers and microvalve (in the middle) may be seen; (b) back-side view: analyte, carrier gas and control pressure inlets.

have been formed on a perforated membrane made of a double-layer SiO_2/Si_3N_4 system fabricated using LPCVD. The membrane extends over a gas channel formed by means of deep wet etchnig of a (100) silicon wafer. This structure is covered with a glue-bonded, 1 mm thick glass plate. The device chip is assembled with the capillaries and packaged in a metallic case, then wire-bonded connections are made (Fig. 4).

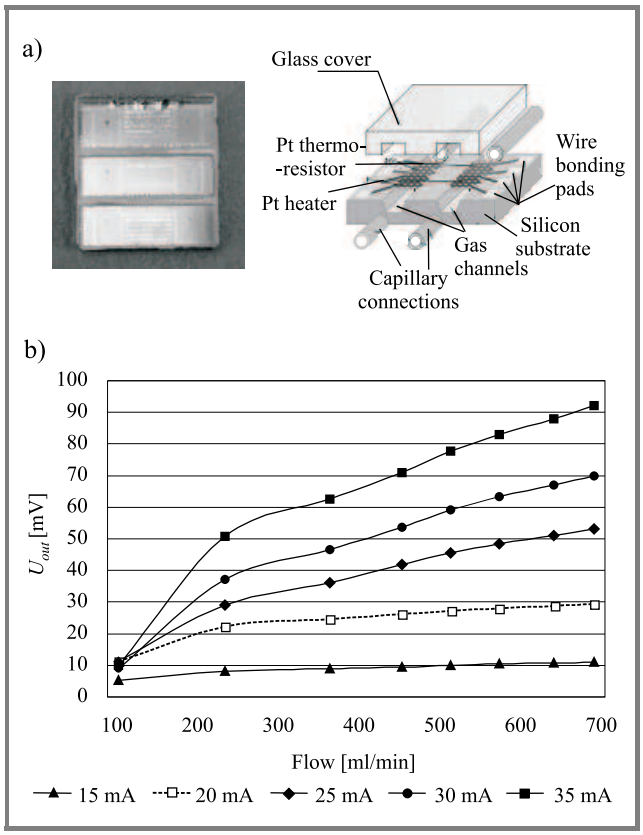


Fig. 4. TCD detector: (a) diagram (right) and view of the fabricated device (left); (b) most important characteristics – U_{out} versus N_2 flow for different levels of supply current.

In the solution reported here two independent sensors have been formed in a single chip and connected in the full Wheatstone bridge configuration in the self-calibrating mode.

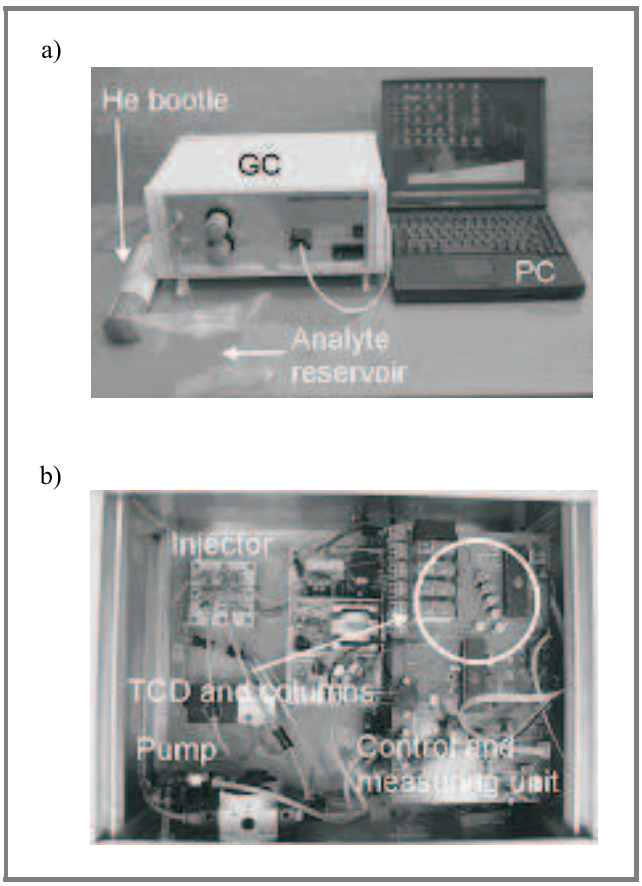


Fig. 5. Portable gas chromatograph: (a) the instrument, carrier gas bottle, and PC; (b) inside view.

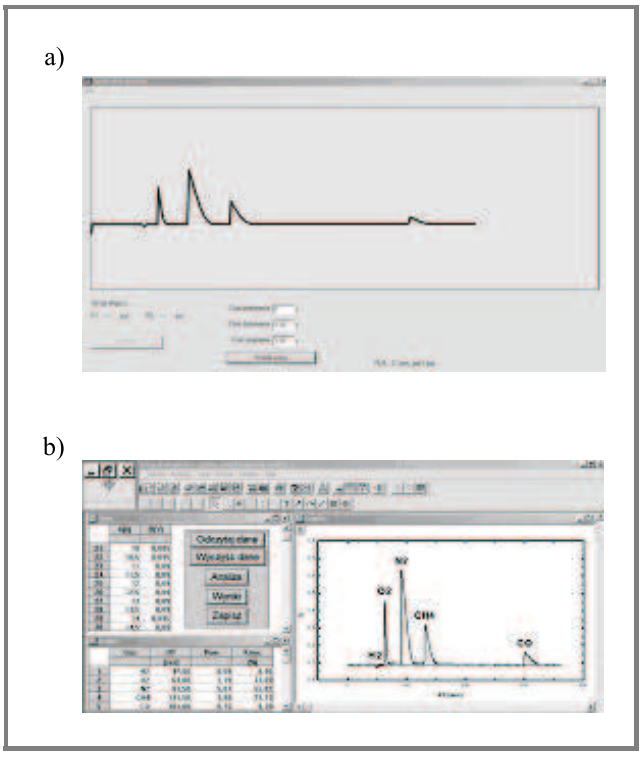


Fig. 6. Results of analysis: (a) electrical output signal and virtual control board of the GC; (b) processed chromatogram of a gas mixture and data table.

Table 1
Technical parameters of the GC reported here

Dimensions/weight:
15 cm × 45 cm × 30 cm/∼ 3.5 kg

Environmental conditions:
Indoor, automatic sampling from atmosphere (build in pump) or external source of analyte

Supply:
230 $V_{(AC)}$, 50–60 Hz, 200 VA

Carrier gas:
Helium carrier gas, external source, min. pressure 220 kPa

Analysis time:
Depends on gas mixture and applied column, 400–700 s

Injector:
Silicon/glass micromachined with polyimide membrane, the volume of injected analyte determinated by cross-valve open time or total volume of the dosage loop (14 μL $\pm$1%), possible backflush and independent temperature stabilization

Detector:
Silicon/glass micromachined two-way thermo conductive detector (TCD) with four Pt micro spirals set up in auto compensating Whetsfone's bridge circuit, death volume in nL range

Chromatographic column:
Depends on application (RT-Msieve 5 A, 10 m long, temperature 50–150°C)

Control unit/software:
External PC laptop-type with RS 232 or USB port, specialized software under Windows98/NT/2000/XP platform, all parameters of analysis electronicaly adjusted

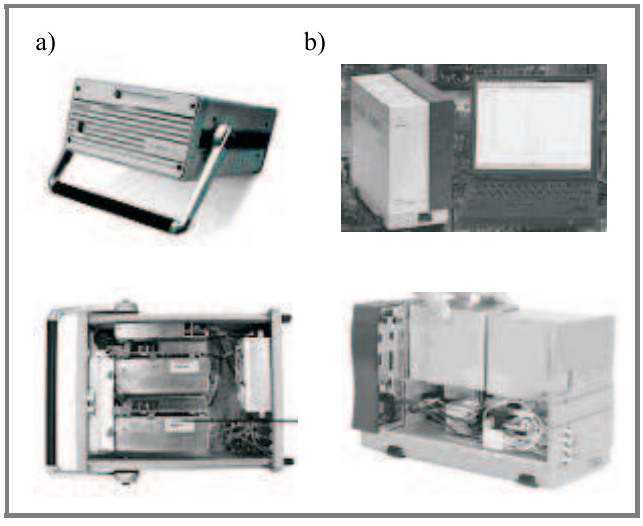


Fig. 7. The competing portable GCs: (a) Agilent 3000 Micro-GC; (b) Varian CP-4900 Micro-GC.

A printed circuit board with two 16-bit microcontrolers and specialized software designed and fabricated by the Institute of Automation and Electrification of Mining Industry – EMAG is used in the chromatograph.

The GC has been mounted in a metal box containing gas/electrical facilities (pump, pressure regulators and sensors, control valves). Both the injector and the TCD with the capillary columns enclosed in a thermally stabilized case are assembled in this box (Fig. 5).
The GC is able to operate either in the single-shot or in the continuous, automatic mode of analysis. All parameters of an analysis have to be set up by software from an external portable PC. The pressure of the carrier gas has to be adjusted manually, prior to the first analysis.
Chromatograms of gas mixtures may be observed immediately (real-time measurements), stored in memory and edited in the form table data and processed chromatograms with the description of detected gases added (Fig. 6).
Technical parameters our GC are similar to those of the well known Agillent/HP instrument (Table 1). The external view of both devices is shown in Fig. 7.

3. Conclusions

The most advantageous feature of our GC is the technical solution applied (possible only due to a complex know-how procedure). The approach resulted in technical parameters similar to those obtained by big producers of similar equipment in the world.
It is well known that the first integrated gas chromatograph was developed in the late 70's of the 20th century by Angell, Terry and Jerman [1]. After several years of redesigning, production of the portable gas chromatograph was undertaken by MTI Inc. (USA) and carried out in the late 80's and early 90's. Later the instrument became widely used under the name of Hewlett Packard (Agilent USA, http://www.chem.agilent.com).
At the moment, few projects devoted to the development of integrated gas chromatographs exist in Europe [2, 3], although many companies have been developing their own custom-designed gas meters based on the solutions used by Agilent. The lack of an independent market of key components and/or assembled portable gas chromatographs is the most important barrier blocking wider application of miniature gas chromatographs for on-line analysis in environment protection, gas systems electroengineering systems, etc.
We have built a portable (weighting aprox. 3.5 kg) gas chromatograph for continuous analysis of a gas mixture. The lowest detection level obtained is a few ppm, the average analysis time is below 6 minutes (depending on a column applied), the repetition time is 10 seconds. All parameters of the GC may be adjusted electronically, the results of the analysis may be visualized immediately. At the moment it seems that our gas chromatograph is the only European solution comparable to Agilent/HP standards.
We would like to emphasize that the GC presented here is the effect of a long-lasting cooperation between several partners from academic and research institutes in Poland (Wrocław, Warsaw, Katowice). The success of this shared efford relies heavily on programs of the State Committee for Scientific Research PBZ 2705 (basic technologies

applicable in GC fabrication) and PBZ 01915, as well as on other projects carried out by EMAG Katowice.
We hope that good scientific results reported here will be followed by technology transfer and industrial implementation of the integrated GC.

Acknowledgements

We would like to thank the graduate students of the Faculty of Electronics and Faculty of Microsystems Electronics and Photonics of Wrocław University of Technology Wojciech Bula and Dawid Zalewski for their help, to Dr. Rafał Walczak, Dr. Anna Górecka-Drzazga, Dr. Łukasz Nieradko, Dr. Krzysztof Małecki of Faculty of Microsystems Electronics and Photonics of the above mentioned University, and Dr. Małgorzata Szczygielska and Marcin Małachowski of the Center EMAG Katowice for the development and fabrication of the integrated components and GC. Last but not least, we would like to express our deepest appreciation of the technological effort of Paweł Kowalski, Dr. Jan Łysko, Jerzy Jaźwiński and Dr. Zenon Gniazdowski of the Institute of Electron Technology from Warsaw.
The research has been financed (mainly) by PBZ 01915 and (partly) 9T11A03710 programs of the State Committee for Scientific Research.

References

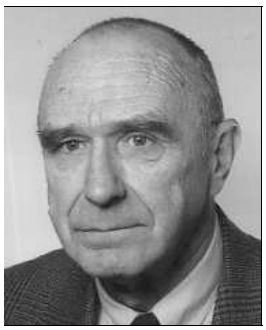
[1] S. Terry, J. Jerman, and J. Angell, "A gas chromatographic air analyzer fabricated on silicon wafer", *IEEE Trans. Electron. Dev.*, vol. ED26, no. 2, pp. 1880–1886, 1979.

[2] M. B. Stone and J. A. de Mello, "Life, the universe and microfluidics", *Lab Chip*, no. 2, pp. 58–64, 2002.

[3] Actaris Inc., "The meter of natural gas" (nonpublished technical letter).

Jan A. Dziuban was born in Sanok, Poland, in 1951. He got the M.Sc. degree in 1974, then Ph.D. in 1978 and D.Sc. in 2002. He has been Assistant Professor at Wrocław University of Technology since 1978, the Directeur de Recherche at CNRS Besancone, Institute FEMTO since 2004. He has been a member of the Board of Polish Society of Sensor Technology since 1992, a member of Steering Committee of Eurosensors Conference since 1993, a member of Steering Committee of IVNC since 1999, a member of Commission of Microengineering of PAS, Committee of Machines Building Polish Academy of Sciences since 2003. He is a founder of micromechanics of silicon in Poland. He is the author of over 110 scientific works and patents.
e-mail: Jan.Dziuban@pwr.wroc.pl
Faculty of Microsystem Electronics and Photonics
Wrocław University of Technology
Janiszewskiego st 11/17
50-372 Wrocław, Poland

Jerzy Mróz received the M.Sc. degree from the Mining Electrical Engineering Department, AGH University of Science and Technology in 1969 and Ph.D. degree in engineering in 1989. As an engineer, he worked in GIG on sensors of ventilation parameters in mines. Since 1975 he's been working in the Research and Development Center for Electrical Engineering and Automation in Mining (EMAG). At present he is involved in the Safety Systems and Ventilation Control Department.
e-mail: mroz@emag.katowice.pl
The Research and Development Center
for Electrical Engineering and Automation
in Mining (EMAG)
Leopolda st 31
40-189 Katowice, Poland



Jan Koszur was born in Warsaw, Poland, in 1943. He received the M.Sc. degree from Department of Electronics, Warsaw University of Technology, Poland, in 1968. His scientific interest is focused on design and fabrication of the silicon microsystems. He is working now as a Chief of Silicon Sensors Department of Institute of Electron Technology (IET), Warsaw.
e-mail: koszur@ite.waw.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Invited paper

Low frequency noise in advanced Si bulk and SOI MOSFETs

Jalal Jomaah, Francis Balestra, and Gérard Ghibaudo

Abstract—A review of recent results concerning the low frequency noise in modern CMOS devices is given. The approaches such as the carrier number and the Hooge mobility fluctuations used for the analysis of the noise sources are presented and illustrated through experimental data obtained on advanced CMOS SOI and Si bulk generations. Furthermore, the impact on the electrical noise of the shrinking of CMOS devices in the deep submicron range is also shown. The main physical characteristics of random telegraph signals (RTS) observed in small area MOS transistors are reviewed. Experimental results obtained on 0.35–0.12 μm CMOS technologies are used to predict the trends for the noise in future CMOS technologies, e.g., 0.1 μm and beyond. For SOI MOSFETS, the main types of layout will be considered, that is floating body, DTMOS, and body-contact. Particular attention will be paid to the floating body effect that induces a kink-related excess noise, which superimposes a Lorentzian spectrum on the flicker noise.

Keywords—CMOS, SOI, low frequency noise, fluctuations, kink-related excess noise, DTMOS.

1. Introduction

The low frequency (LF) noise and fluctuations in metal-oxide-semiconductor (MOS) devices have been the subject of intensive research during the past years. The LF noise is becoming a major concern for scaled-down devices, since the $1/f$ noise increases as the reciprocal of the device area [1–38]. Excessive low frequency noise and fluctuations could lead to serious limitation of the functionality of the analog and digital circuits. The $1/f$ noise is also of paramount importance in RF circuit applications where it gives rise to phase noise in oscillators or multiplexers [33]. The development of submicron CMOS technologies has led to the onset of new type of noise, i.e., random telegraph signals (RTS's) [8–10, 15], yielding large current fluctuations, which can jeopardize the circuit functionality. In addition, thanks to their structure, the SOI technologies present several intrinsic properties for analog and RF applications. For instance, as it is well established now, these interesting devices allow the reduction of the power consumption at a given operating frequency.

Moreover, the high insulating properties of SOI substrates, in particular with the use of high resistivity material, leads to high performance mixed-signal circuits [34–36]. However, in order to use this kind of devices in such applications, low frequency noise, which can directly impact RF or analog integrated circuits, needs to be thoroughly evaluated. Following the specifications of the ITRS roadmap, the $1/f$ noise amplitude is predicted to decrease in modern technologies. But, the maximum signal is also lowered with decreasing operation voltage leading to a deterioration of the signal to noise ratio. Therefore, accurate characterization of the noise behavior has to be established.

The aim of this paper is first to present recent issues about the low frequency and RTS noise in CMOS devices. The approaches such as the carrier number and the Hooge mobility fluctuations used for the analysis of noise sources will be presented and illustrated through experimental data obtained on advanced CMOS SOI and Si bulk generations. The application of low frequency noise measurements as a characterization tool for large area MOS devices will also be discussed. Besides, the main physical parameters characterizing the RTS's in small area MOS transistors will be briefly reviewed. Experimental data obtained on 0.35–0.12 μm CMOS technologies will be used to predict the trends for the noise in the future CMOS technologies, e.g., 0.1 μm and beyond.

2. Low frequency noise in large area devices

2.1. Carrier number fluctuations and correlated mobility fluctuations

In the classical carrier number fluctuation approach, the fluctuations in the drain current stem from the fluctuations of the inversion charge nearby the Si-SiO_2 interface, arising from the variations of the interfacial oxide charge after dynamic trapping/detrapping of free carriers into slow oxide border traps. These interface charge fluctuations δQ_{it} can be equivalently equated to a flat band voltage variation $\delta V_{fb} = -\delta Q_{it}/(WLC_{ox})$. Moreover, in a more detailed analysis one should also take into account the supplementary mobility change $\delta\mu_{eff}$ due to the modulation of the scattering rate induced by the interface charge fluctuations. The drain current fluctuations therefore read [16, 22]:

$$\delta I_d = -g_m \delta V_{fb} - \alpha I_d \mu_{eff} \delta Q_{it}, \qquad (1)$$

where g_m is the transconductance, μ_{eff} is the effective mobility, α is the Coulomb scattering coefficient ($\approx 10^4$ Vs/C for electrons and 10^5 Vs/C for holes).

This leads to a normalized drain current and input gate voltage noise $S_{Vg} = S_{I_d}/g_m^2$ for strong inversion:

$$S_{I_d}/I_d^2 = (1+\alpha\,\mu_{eff}\,C_{ox}\,I_d/g_m)^2\left(\frac{g_m}{I_d}\right)^2 S_{V_{fb}} \qquad (2a)$$

and

$$S_{V_g} = S_{V_{fb}}\left[1+\alpha\,\mu_0\,C_{ox}(V_g-V_t)\right]^2, \qquad (2b)$$

where μ_0 is the low field mobility, C_{ox} is the gate oxide capacitance, V_t is the threshold voltage, $S_{V_{fb}} = S_{Q_{it}}/(WLC_{ox}^2)$ with $S_{Q_{it}}$ (C^2/Hz/cm^2) being the interface charge spectral density per unit area, W the device width and L the device length.

The spectral density of the oxide interface charge depends essentially on the physical trapping mechanisms into the oxide. For a tunneling process, the trapping probability decreases exponentially with oxide depth x, so that the flat band voltage spectral density takes the form [2, 3]:

$$S_{V_{fb}} = \frac{q^2 \mathrm{k}T\,\lambda\,N_t}{WLC_{ox}^2 f^\gamma}, \qquad (3)$$

where f is the frequency, γ is a characteristic exponent close to 1, λ is the tunnel attenuation distance (≈ 0.1 nm), kT is the thermal energy and N_t is the volumetric oxide trap density (eV^{-1} cm^{-3}).

For a thermally activated trapping process [7], the trapping probability decreases exponentially with the cross section activation energy E_a, so that the flat band voltage spectral density reads [11]:

$$S_{V_{fb}} = \frac{q^2 \mathrm{k}^2 T^2 N_{it}}{WLC_{ox}^2 f^\gamma \Delta E_a}, \qquad (4)$$

where ΔE_a is the amplitude of the activation energy dispersion and N_{it} is the oxide trap surface state density (eV^{-1} cm^{-2}). In both trapping mechanisms, the $1/f$ nature of the spectrum stems from the uniform distribution in log scale of the involved time constants [5].

It should be mentioned that Eqs. (2a,b) also applies to the nonlinear regime of MOSFET operation [23]. Indeed, the drain voltage dependence of S_{I_d}/I_d^2 is naturally accounted for by that of the transconductance to drain current ratio squared $(g_m/I_d)^2$. This can been checked experimentally by comparing the respective variations of S_{I_d}/I_d^2 and $(g_m/I_d)^2$ with drain voltage V_d.

2.2. Hooge mobility fluctuations

In the Hooge model [4], the drain current noise results from the fluctuations of the carrier mobility through the variations in the scattering cross section entering the collision probability likely due to phonon number fluctuations [6]. This leads to a flicker noise with the amplitude inversely proportional to the total number of carriers in the device. The normalized drain current noise and input gate voltage noise in ohmic operation can then be written in the form [17, 22]:

$$\frac{S_{I_d}}{I_d^2} = \frac{q\alpha_h}{WLQ_i f} \qquad (5)$$

and

$$S_{V_g} = \frac{q\alpha_h}{WLfC_{ox}}(V_g-V_t)\left[1+\theta(V_g-V_t)\right]^2, \qquad (6)$$

where Q_i is the inversion charge, α_h is the Hooge parameter ($\approx 10^{-4}-10^{-6}$), and θ is the mobility attenuation coefficient.

Similarly, in the case of the nonlinear region of MOSFET operation, the inversion layer is no longer uniform along the channel and the normalized drain current can be expressed as [17]:

$$\frac{S_{I_d}}{I_d^2} = \frac{q\alpha_h}{fWL^2}\int_0^L \frac{\mathrm{d}y}{Q_i(y)} = \frac{q\alpha_h}{fWL^2}\int_0^{V_d}\frac{W\mu_{eff}}{I_d}\mathrm{d}\phi_c$$
$$= \frac{q\alpha_h < \mu_{eff} > V_d}{fL^2 I_d}, \qquad (7)$$

where $< \mu_{eff} >$ is the average mobility along the channel. Indeed, in the ohmic region, Eq. (7) reduces to Eq. (5). Therefore, in all the cases (ohmic and nonlinear regions), the normalized drain current noise due to Hooge mobility fluctuations varies as the reciprocal drain current.

2.3. Impact of series resistance

The impact of series resistance on the low frequency noise can simply be obtained by adding to the channel current noise the contribution of the excess noise arising from the access region. For instance, in the linear region, the total drain current noise becomes [26]:

$$\frac{S_{I_d}}{I_d^2} = \left(\frac{S_{I_d}}{I_d^2}\right)_{channel} + \left(\frac{I_d}{V_d}\right)^2 S_{R_{sd}}, \qquad (8)$$

where $S_{R_{sd}}$ is the spectral density of source-drain series resistance.

For the carrier number fluctuations, one gets:

$$\frac{S_{I_d}}{I_d^2} = \left(1+\alpha\,\mu_{eff}\,C_{ox}\,I_d/g_m\right)^2\left(\frac{g_m}{I_d}\right)^2 S_{V_{fb}} + \left(\frac{I_d}{V_d}\right)^2 S_{R_{sd}}. \qquad (9)$$

2.4. Diagnosis of the low frequency noise sources

2.4.1. Measurements in the linear regime

A generic procedure for the diagnosis of the excess LF noise sources in a MOS transistor can be drawn from the above analyses [22]. The normalized drain current noise versus drain current characteristics in a log-log scale can first be inspected for comparison with Eqs. (2) and (7). If the normalized drain current spectral density varies with the drain current as the transconductance to drain current

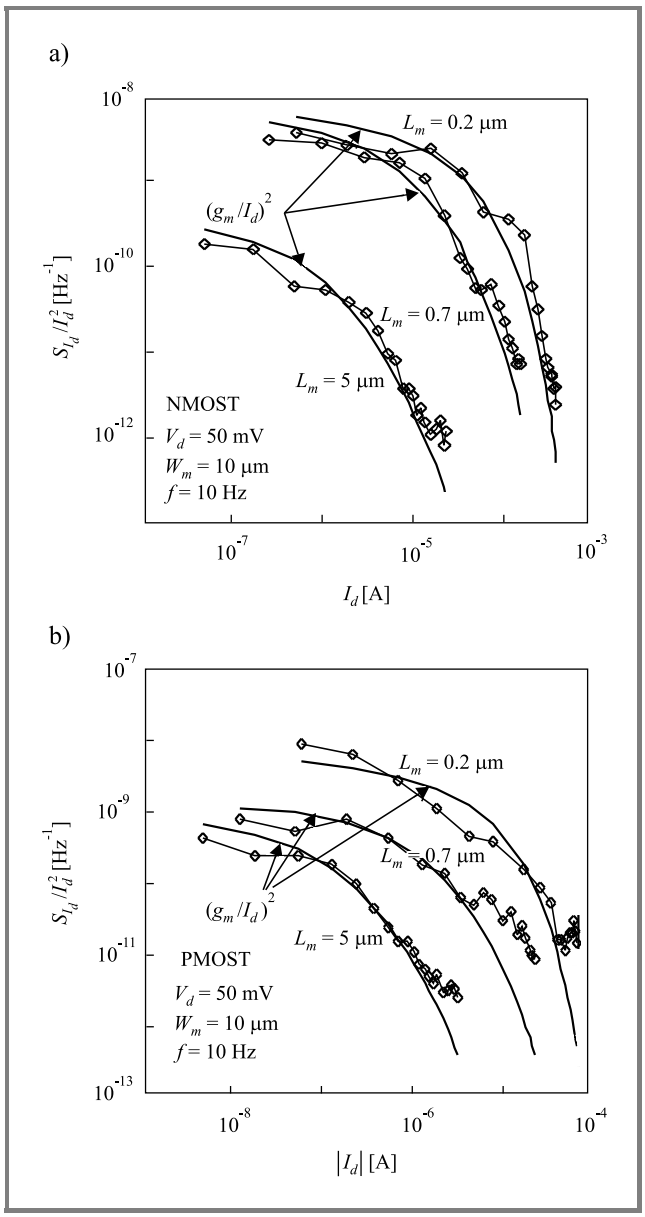


Fig. 1. Variation of the normalized drain-current noise versus drain current for (a) n-channel and (b) p-channel 0.18 μm CMOS devices.

ratio squared, one can likely conclude that carrier number fluctuations dominate. Furthermore, if the associated input gate voltage noise shows a parabolic dependence on the gate voltage at strong inversion, the correlated mobility fluctuations might be involved ($\alpha \gg 1$ in Eq. (2)).

Hooge mobility fluctuations can also be diagnosed using the $S_{I_d}/I_d^2(I_d)$ plot in a log-log scale. If the normalized current noise varies as the reciprocal of the drain current from weak to strong inversion, it can be concluded that Hooge mobility fluctuations dominate.

Moreover, if the normalized drain current noise increases at high drain current, this is indicative of an enhanced LF noise contribution of the access resistance (see Eq. (8)).

Figure 1 gives typical examples of $S_{I_d}/I_d^2(I_d)$ characteristics for 0.18 μm CMOS Si bulk devices. They illustrate

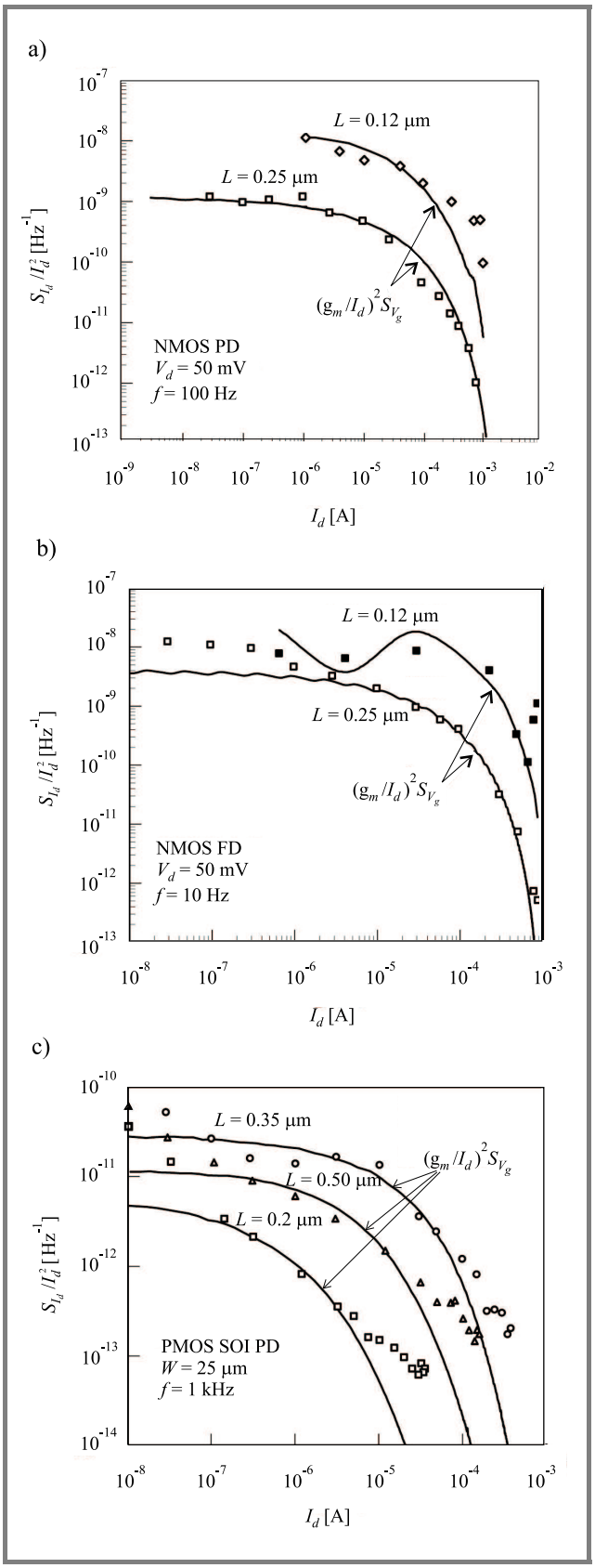


Fig. 2. Normalized drain-current power spectral density S_{I_d}/I_d^2 at $V_d = 50$ mV as a function of drain current at different channel lengths for (a) NMOS FB partially-depleted; (b) NMOS fully depleted and (c) PMOS FB partially-depleted SOI devices. Full lines: $S_{V_g} * (g_m/I_d)^2$.

a very good correlation between the normalized drain current noise and the transconductance to drain current ratio squared from weak to strong inversion. Therefore, in this case, it can be inferred that carrier number fluctuations are the main LF noise source. This behavior is representative of modern n- and p-channel devices ($L < 0.35$ μm), where both device types operate in surface mode due to the dual poly gate material. In contrast, for 0.35 μm technology, it was found that p-type transistors obey Hooge mobility model due to the buried architecture of the channel with n^+ Si poly gate [29]. Therefore, in general, for surface mode operated MOSFETs, the LF noise is found to result from carrier number fluctuations whereas, for volume mode operated devices, Hooge mobility fluctuations should contribute more.

Figure 2 shows the normalized drain current power spectral density S_{I_d}/I_d^2 in the ohmic regime ($V_d = 50$ mV), plotted as a function of the drain current for n-channel partially-depleted (PD) SOI MOSFETs (Fig. 2a) and for n-channel fully-depleted (FD) SOI MOSFETs (Fig. 2b) for two channel lengths: $L = 0.25$ and 0.12 μm. In this plot, the straight line represents the front gate power spectral density S_{V_g} multiplied by the ratio $(g_m/I_d)^2$, where g_m stands for the gate transconductance. A good correlation is obtained between these two quantities, confirming the results predicted by the McWhorter model, which associates the $1/f$ noise with the carrier number fluctuations. Moreover, some difference in strong inversion can be observed (p-channel, Fig. 2c), which is attributed to the correlated mobility fluctuations.

2.4.2. Measurements in the saturation mode

Low frequency noise characterization of SOI devices also needs to be carried out in the saturation mode, where bias conditions give birth to the kink-related excess noise observed in SOI devices. The drain current spectral density S_{I_d}, normalized to the width, is plotted (Fig. 3) versus the applied drain voltage V_d for 0.12 μm floating body (FB) PD SOI and for $L = 0.25$ μm with FB-PD and body-contacted (BC) technologies.

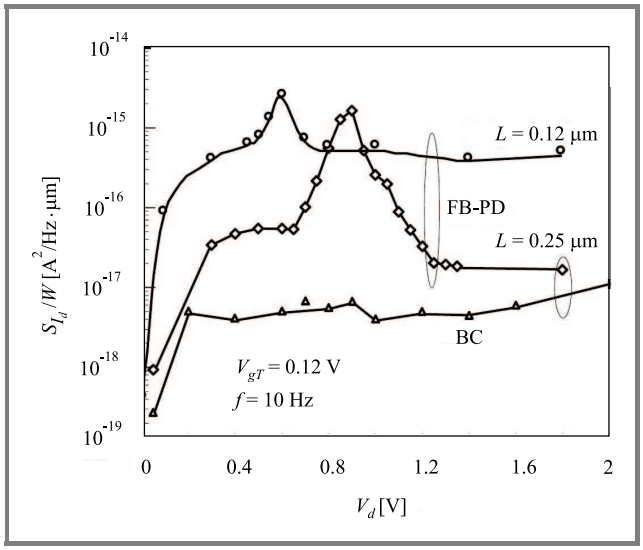


Fig. 3. Drain-current power spectral density, normalized to the width W as a function of drain voltage, at $f = 10$ Hz for FB-PD devices with two channel lengths ($L = 0.25$ and 0.12 μm) and for BC devices with $L = 0.25$ μm.

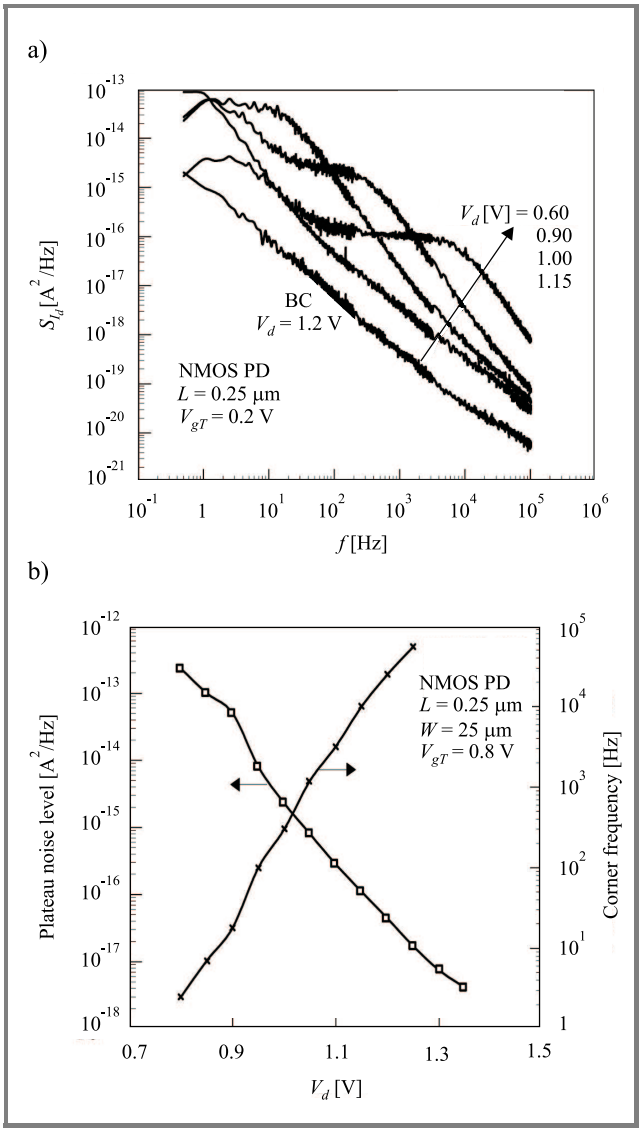


Fig. 4. (a) Drain current power spectral density versus frequency for 0.25 μm FB and BC partially-depleted devices, at different drain voltages and (b) drain bias dependence of the plateau noise level and the corner frequency for 0.25 μm PD device with a floating body.

A substantial kink effect is clearly observed in the case of FB-PD devices, and a low frequency excess noise occurs for a drain bias corresponding to the kink effect in static measurements. This noise peak shifts with the frequency, towards higher V_d. However, the kink effect disappears when the SOI film is connected to the ground. The floating body effect induces a kink-related excess noise, which superimposes a Lorentzian spectrum on the flicker noise,

characterized by a corner frequency, f_c, and a plateau noise level, S_{I_d} (0).

Several mechanisms have been proposed to explain this excess noise, such as trap-assisted generation-recombination noise [37] or shot noise amplified by floating body effect [38]. This behavior is clearly shown in Fig. 4a for a 0.25 μm PD device at different values of V_d. It is important to note that the corner frequency and the plateau noise level of the Lorentzian spectrum depend both on the drain bias (Fig. 4b).

For PD devices, f_c increases and the noise level of the plateau of the Lorentzian spectrum decreases continuously as the drain bias is enhanced. This is due to a close correlation between the substrate current, I_B, and the kink phenomenon. The noise plateau level is reversely proportional to I_B and the corner frequency is proportional to the substrate current. Moreover, a relevant point here is the difference between the kink noise overshoot for 0.25 and 0.12 μm PD technologies. The magnitude for $L = 0.12$ μm is only one decade high, contrary to the 0.25 μm SOI CMOS technology for which two orders of magnitude were obtained.

3. Random telegraph signals in small area MOS devices

The observation of RTS's in small area MOS transistors is commonly attributed to individual carrier trapping at the silicon-oxide interface [8–10, 14, 15, 18]. The drain current RTS fluctuations can be interpreted as conductance modulation originating from carrier number and/or subsequent mobility fluctuations.

Figure 5 displays a typical time domain waveform of the drain current illustrating the three main RTS parameters. As the transition times from low to high level (and *vice versa*) are Poisson distributed random variables, the RTS parameters have to be measured using statistical analysis [12, 21]. The histogram of the drain current amplitudes, which is no longer Gaussian, provides the average drain current RTS amplitude ΔI_d. The average values of the high and low level time constants represent, for an acceptor like trap, the capture time τ_c and the emission time τ_e, respectively.

The drain current spectrum of a random telegraph signal has a Lorentzian shape [1]:

$$S_{I_d} = 4A\Delta I_d^2 \frac{\tau}{1+\omega^2\tau^2}, \tag{10}$$

where $\tau = (1/\tau_c + 1/\tau_e)^{-1}$ is the effective time constant, $A = \tau/(\tau_c+\tau_e) = f_t(1-f_t)$ is the space mark ratio $\omega = 2\pi f$ is the angular frequency and f_t is the trap occupancy factor, $f_t = 1/\left\{1+\exp\left[(E_t-E_f)/\mathrm{k}T\right]\right\}$ with E_t being the trap energy and E_f the Fermi level position.

The drain current RTS amplitude can be calculated assuming that the trapping of an elementary charge q in the channel changes the local conductivity [12, 23–25]. It is easy to show that, to a first order approximation, the relative drain current RTS amplitude is given by [25, 27]:

$$\frac{\Delta I_d}{I_d} = \frac{g_m}{I_d}\frac{q}{WLC_{ox}}\left(1-\frac{x_t}{t_{ox}}\right), \tag{11}$$

where x_t is the distance of the trap to the Si-SiO_2 interface and t_{ox} is the gate oxide thickness.

This expression, which applies to the ohmic and nonlinear regions, shows that the RTS amplitude should vary with drain and gate voltages as the transconductance to drain current ratio. It is also worth noticing that it represents the amplitude version of Eq. (2a) for $\alpha = 0$.

Figure 6 displays typical variations of the relative drain current RTS amplitude with gate and drain voltages as obtained on NMOSFETs, which illustrate the validity of Eq. (11). This relation is rather well appropriate for the strong inversion region. However, it does not explain quantitatively the huge sample-to-sample variations of the RTS amplitude observed in weak inversion [27]. It has been proposed that a proportionality factor depending on the trap area could be introduced in Eq. (11) in order to extend its range of application [27].

The capture and emission times are in general governed by the Shockley-Read-Hall statistics [12]:

$$\tau_c = \frac{1}{\sigma n_s v_{\mathrm{th}}} \quad \text{and} \quad \tau_e = \frac{1}{\sigma n_1 v_{\mathrm{th}}}, \tag{12}$$

where v_{th} is the thermal velocity, n_s is the surface carrier concentration, n_1 is the surface carrier concentration when

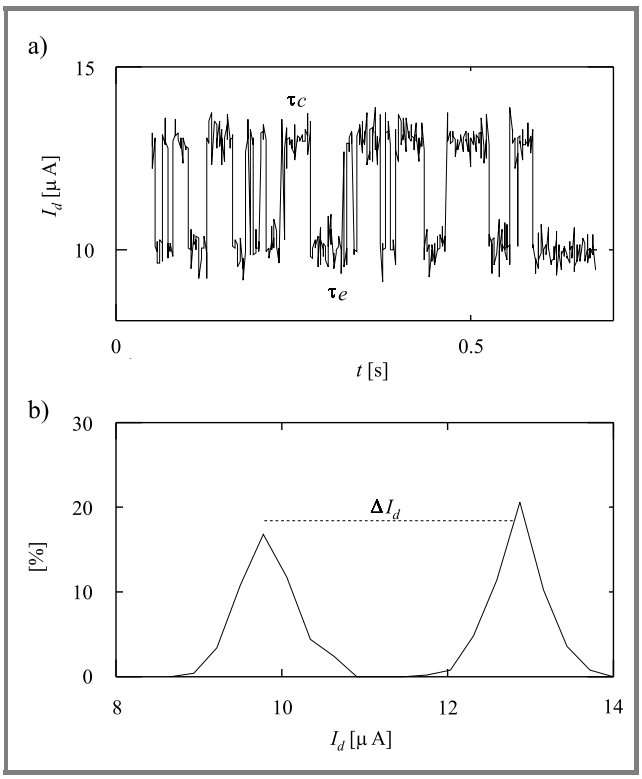


Fig. 5. Typical time-domain plot of the drain current for RTS noise (a) and corresponding amplitude histogram (b).

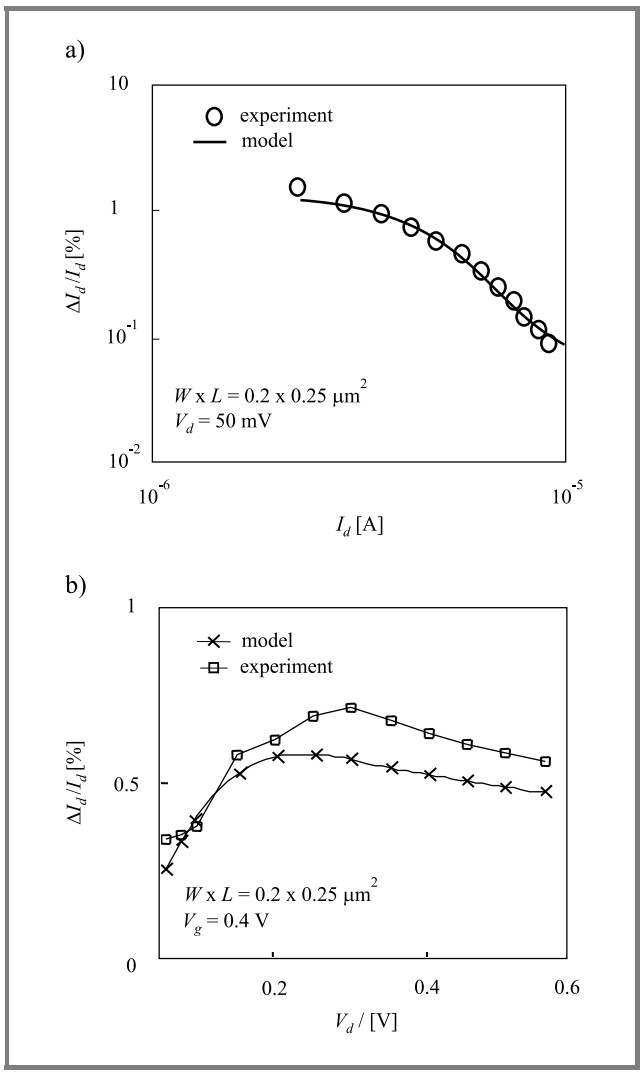


Fig. 6. Typical variations of the capture and emission times with gate (a) and drain (b) voltages. The lines show the results from the SRH model with appropriate parameters.

the Fermi level E_f equals the trap energy E_t, and σ is the trap cross section including tunneling effect and activated process [12, 20].

4. Low frequency noise in DTMOS

In this section, low frequency noise (LFN) in n- and p-channel dynamic-threshold MOSFET's on unibond substrate (SOI) is investigated. Two different types of transistors have been used: with and without current limiter. The LFN in DTMOS is analyzed in ohmic and saturation regimes. The impact of the use of a current limiter (clamping transistor) is shown. An explanation based on floating body effect inducing excess noise is proposed.

Due to a high-leakage current when the body is strongly forward biased, a small size current limiter is added between the gate and body in a DTMOS structure. Figure 7 shows the two device types. Note that the clamping transistor complicates the design but prevents the body potential (V_{BS}) from exceeding 0.65 V and allows the operation at 1 V gate bias without high gate current. Figure 8 shows drain and gate static currents versus gate voltage for 0.25 μm N-DTMOS with and without a clamping transistor. We can note the gate current limitation thanks to the clamping transistor.

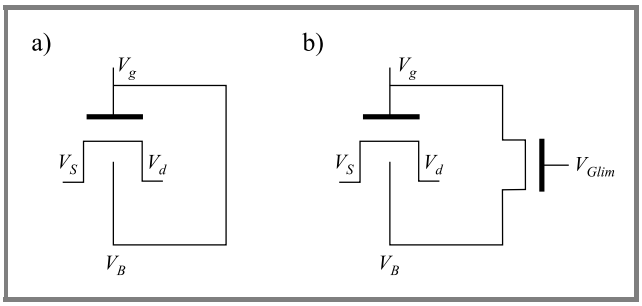


Fig. 7. Schematics of the two transistor types. DTMOS (a) without a clamping transistor and (b) with a clamping transistor.

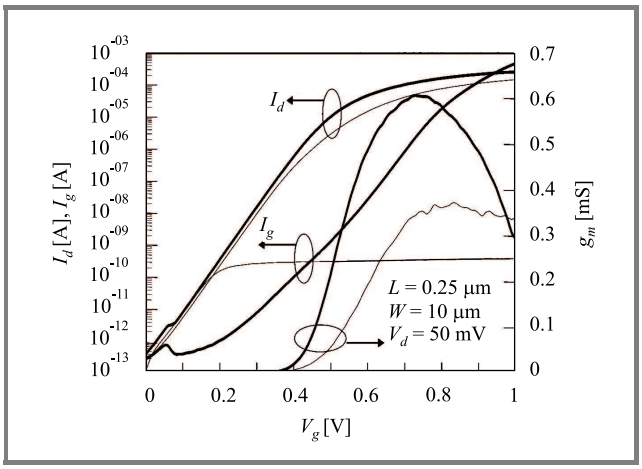


Fig. 8. Drain and gate currents, and gate transconductance versus gate voltage at $V_d = 50$ mV for 0.25 μm DTMOS with and without current limiter.

In Fig. 9a the drain current power spectral density (S_{I_d}) in the ohmic region ($V_d = 50$ mV) is plotted as a function of frequency at various levels of drain current, for 0.25 μm n-channel DTMOS without a clamping transistor. $1/f$ spectra are obtained for all drain-current levels. Similar results have also been obtained in PMOS transistors.

If a clamping transistor is used with a DTMOS, however, $S_{I_d}(f)$ shows Lorentzian spectra (Fig. 9b). These spectra are strongly dependent on the gate voltage of the current limiter (V_{Glim}). The plateau level of the Lorentzian spectrum decreases and the corner frequency increases when V_{Glim} increases (Fig. 9b). This behavior is quite similar to the noise overshoot (at high V_d) induced by the kink effect in partially depleted SOI, which is due to the impact ionization mechanism. In our case, V_d is too weak to induce the impact-ionization current, however, the clamping transistor current flows through the body inducing an excess noise

following Lorentzian spectra. When V_{Glim} is between zero and 1 V, the body potential is fixed by the clamping transistor current and, consequently, this current increases the body potential inducing a direct biasing of the source-body junction and a kink-excess-like noise is obtained, even at low drain voltage. At $V_{Glim} = 1$ V, the clamping transistor is on, therefore, the body is directly connected to the gate. In this case, a quasi $1/f$ behavior is obtained (Fig. 9b).

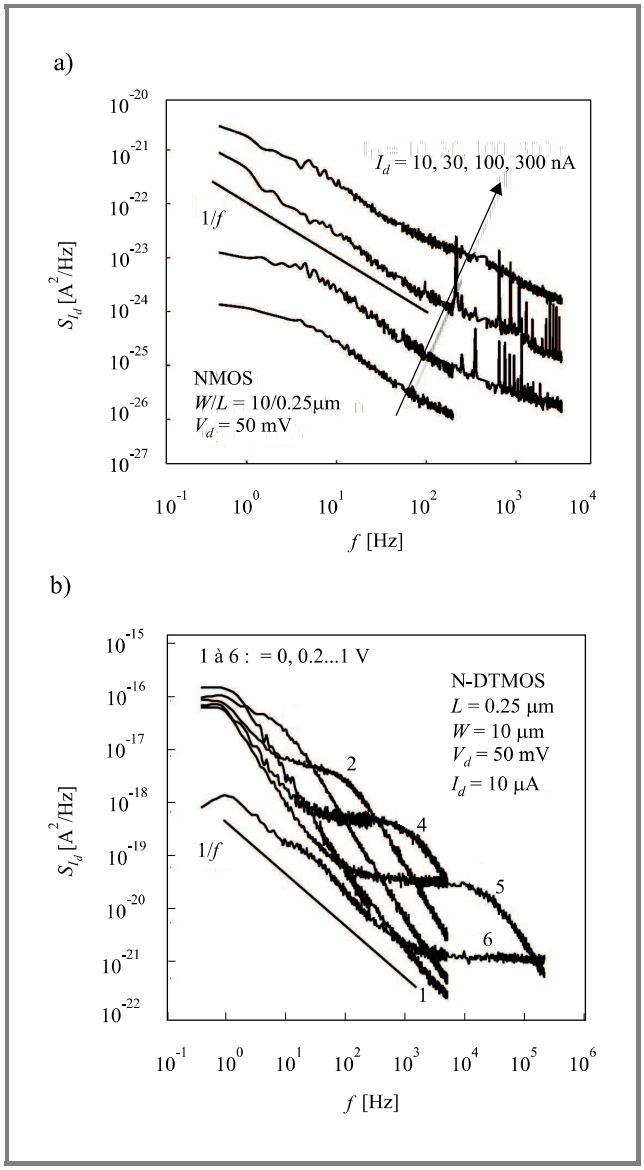


Fig. 9. Drain current power spectral density versus frequency in ohmic region for 0.25 μm n-channel DTMOS without (a) and with (b) current limiter.

Figure 10 shows the normalized drain current noise versus drain current for N- and PMOS transistors in BC and DTMOS modes. A good correlation with $(g_m/I_d)^2$ for n- and p-channel is obtained. It confirms, therefore, that the noise source in these devices is due to the carrier number fluctuations. Moreover, for all devices, the same drain-current noise is obtained, at least in weak inversion.

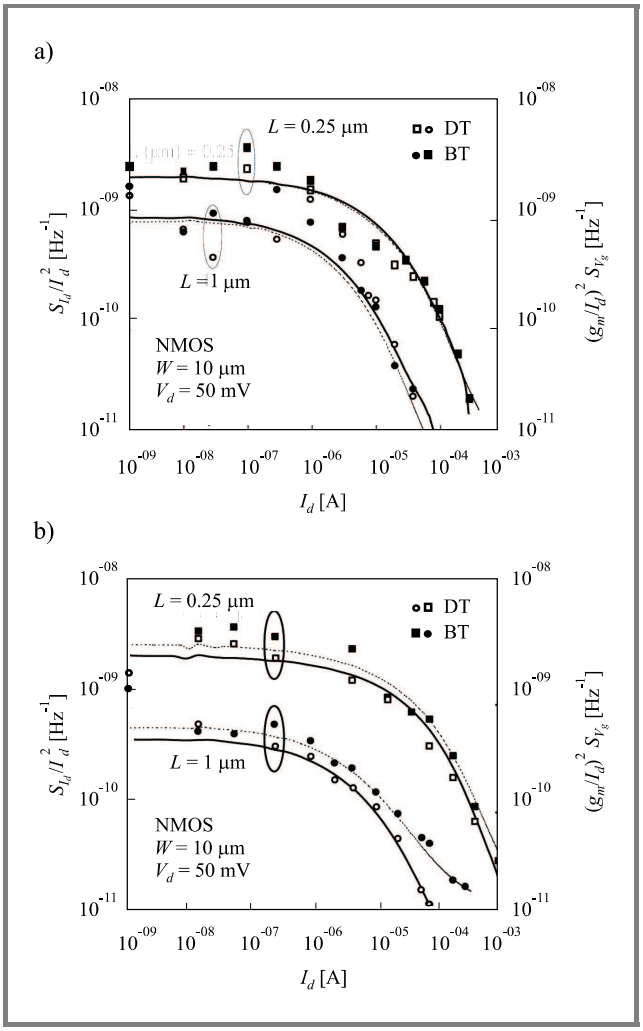


Fig. 10. Normalized drain current power spectral density $S_{I_d}/{I_d}^2$ at $V_d = 50$ mV and two channel lengths for (a) NMOS and (b) PMOS SOI devices. Full symbols for body-contacted and open symbols for DTMOS.

Some difference in strong inversion can be observed (1 μm p-channel BTMOSFET), which is attributed to the correlated mobility fluctuations.

5. Gate induced floating body effects

Low frequency excess noise associated to gate induced floating body effect is reported in partially depleted SOI MOSFETs with ultra-thin gate oxide. This is investigated with respect to floating body devices biased in linear regime. Due to a body charging from the gate, a Lorentzian-like noise component is superimposed on the conventional $1/f$ noise spectrum. This excess noise exhibits the same behaviour as the kink-related excess noise previously observed in partially depleted devices in the saturation regime.
Indeed, scaling metal-oxide-semiconductor devices to very-deep submicron dimensions has resulted in an aggressive shrinking of the gate oxide thickness. In this ultra-thin

gate oxide range, direct tunneling from the gate clearly appears, and increases exponentially with decreasing oxide thickness. Taking into consideration the case of floating body PD SOI MOSFETs with a 2 nm front gate oxide, the gate-to-body tunneling component becomes large enough to charge the body of the devices, resulting in gate induced floating body effects (GIFBE) observed even in the linear regime.

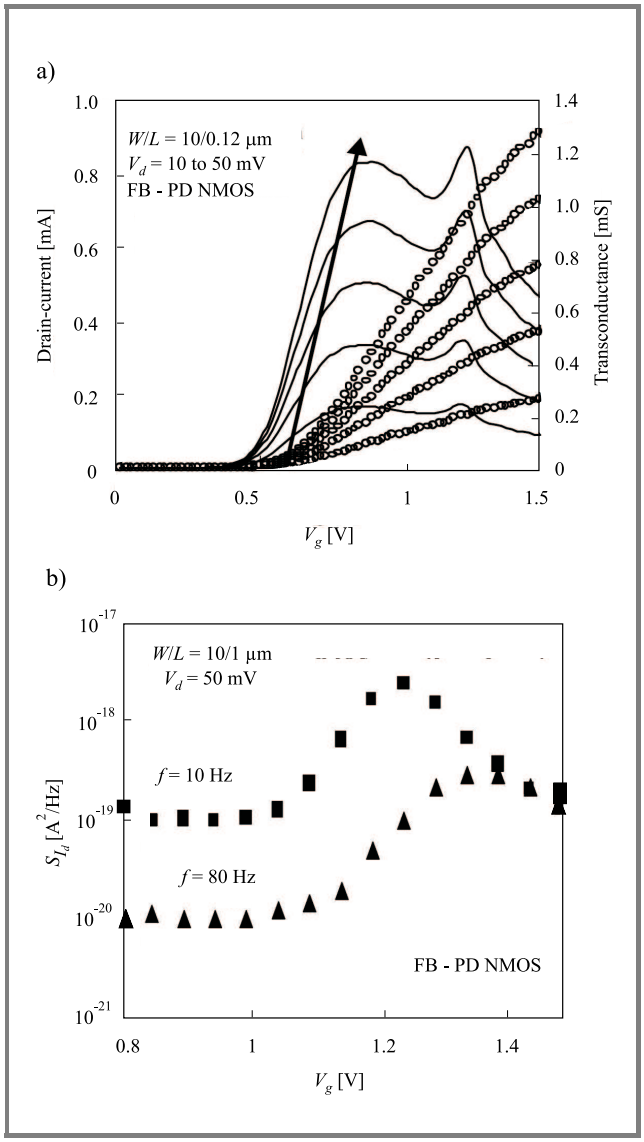


Fig. 11. (a) I_d and g_m transfer characteristics of a $W/L = 10/0.12$ μm PD NMOSFET at low drain voltage ($V_d = 10$ to 50 mV), and (b) drain current power spectral density of a $W/L = 10/1$ μm PD NMOSFET with $V_d = 50$ mV. The two different frequencies are 10 Hz and 80 Hz.

Figure 11a illustrates the drain current and transconductance measured in the linear regime for $V_d = 10$ up to 50 mV for a $W/L = 10/0.12$ μm PD NMOSFET. A sudden increase of the drain current is noticeable close to $V_g = 1.1 - 1.2$ V regardless of the drain bias. This results in an unforeseen second hump in the transconductance characteristics, the value of which exceeds that of the normal peak. The same feature is observed in PMOSFETs. Then, we considered the drain current power spectral density versus the applied front-gate bias for two different frequencies ($f = 10$ and 80 Hz). The results are plotted in Fig. 11b for a $W/L = 10/1$ μm FB-PD SOI MOSFET biased with a drain voltage $V_d = 50$ mV. The level of noise overshoot attributed to the GIFBE is almost two decades for this device. The shift of the noise peak towards higher gate biases with increasing measurement frequency is clearly shown in Fig. 10b.

6. Summary and conclusion

A brief review of recent results concerning the low frequency noise in modern CMOS devices has been given. The approaches such as the carrier number and the Hooge mobility fluctuations used for the analysis of the noise sources have be presented and illustrated through experimental data taken on advanced CMOS generations. The use of low frequency noise measurements as a characterization tool for large area MOS devices has also be discussed. Moreover, the main physical parameters characterizing the RTS's in small-area MOS transistors have been reviewed. An overview of the low frequency noise in both partially and fully depleted SOI CMOS technologies has been given. An enhancement of the overall noise level is noticeable when reducing the channel length. As regards partially depleted devices, the kink-related excess noise magnitude is reduced with the channel length, especially in terms of Lorentzian-like spectra and corner frequency evolution. The LFN in DTMOS was studied in ohmic and saturation regimes and the impact of the use of a current limiter was thoroughly analyzed. Finally, the impact of the oxide thickness thinning on the noise was also shown.

References

[1] S. Machlup, "Noise in semiconductors: spectrum of two parameters random signals", *J. Appl. Phys*, vol. 25, p. 341, 1954.

[2] A. McWhorter, "Semiconductor Surf. Phys.", University of Pennsylvania Press, Philadelphia, 1957.

[3] S. Christensson and I. Lundstroem, "Low frequency noise in MOS transistors_I", *Solid State Electron.*, vol. 11, p. 797, 1968.

[4] F. Hooge, "1/f noise", *Physica*, vol. 83B, p. 14, 1976.

[5] A. van der Ziel, "Flicker noise in electronic devices", *Adv. Electron. & Electron Phys.*, vol. 49, p. 225, 1979.

[6] R. Jindal and A. van der Ziel, "Phonon fluctuation model for flicker noise in elemental semiconductors", *J. Appl. Phys.*, vol. 52, p. 2884, 1981.

[7] P. Dutta and P. Horn, "Low frequency fluctuations in solids: 1/f noise", *Rev. Mod. Phys.*, vol. 53, p. 497, 1981.

[8] K. Ralls, W. Skocpol, L. Jackel, R. Howard, L. Fetter, R. Epworth, and D. Tennant "Discrete resistance switching in submicrometer silicon inversion layers: individual interface traps and low frequency (1/f) noise", *Phys. Rev. Lett.*, vol. 52, p. 28, 1983.

[9] M. Uren, D. Day, and M. Kirton, "Random telegraph noise in silicon metal-oxide-semiconductor field effect transistors", *Appl. Phys. Lett.*, vol. 47, p. 1195, 1984.

[10] P. Restle, "Individual oxide traps as probes into submicron devices", *Appl. Phys. Lett.*, vol. 53, p. 1862, 1988.

[11] C. Surya and T. Hsiang, "A thermal activation model for $1/f$ noise in Si-MOSFETs", *Solid State Electron.*, vol. 31, p. 959, 1988.

[12] M. Kirton and M. Uren, "Noise in solid-state microstructures: a new perspective on individual defects, interface states and low frequency $1/f$ noise", *Adv. Phys.*, vol. 38, p. 367, 1989.

[13] R. Jayaraman and C. Sodini, "A $1/f$ noise technique to extract the oxide trap density near the conduction band edge of silicon", *IEEE TED*, vol. 36, p. 1773, 1989.

[14] K. Kandiah, M. Deighton, and F. Whiting, "A physical model for random telegraph signal currents in semiconductors devices", *J. Appl. Phys.*, vol. 66, p. 937, 1989.

[15] H. Nakamura, N. Yasuda, K. Taniguchi, and A. Toriumi, "Existence of double-charged oxide traps in submicron MOSFET's", *Jap. J. Appl. Phys.*, vol. 28, p. L2057, 1989.

[16] K. Hung, P. Ko, C. Hu, and Y. Cheng, "A unified model for the flicker noise in metal-oxide-semiconductor field-effect transistors", *IEEE Trans. Electron Dev.*, vol. 37, p. 654, 1990.

[17] I. M. Hafez, G. Ghibaudo, and F. Balestra, "A study of flicker noise in MOS transistors operated at room and liquid helium temperature", *Solid State Electron.*, vol. 33, p. 1525, 1990.

[18] A. Ohata, A. Toriumi, M. Iwase, and K. Natori, "Observation of random telegraph signals: anomalous nature of defects at the Si/SiO_2 interface", *J. Appl. Phys.*, vol. 68, p. 200, 1990.

[19] R. Jayaraman and C. Sodini, "$1/f$ noise interpretation of the effect of gate oxide nitridation and reoxidation in dielectric traps", *IEEE TED*, vol. 37, p. 305, 1990.

[20] D. H. Cobden, M. Uren, and M. Kirton, "Entropy measurements on slow Si/SiO_2 interface states", *Appl. Phys. Lett.*, vol. 56, p. 1245, 1990.

[21] M. Schultz and A. Karmann, "Individual, attractive defect centers in the SiO_2-Si interface of μm-sized MOSFETs", *Appl. Phys. A*, vol. 52, p. 104, 1991.

[22] G. Ghibaudo, O. Roux, C. Nguyen-Duc, F. Balestra, and J. Brini, "Improved analysis of low frequency noise in field-effect MOS transistors", *Phys. Stat. Sol. A*, vol. 124, p. 571, 1991.

[23] G. Ghibaudo, O. Roux, and J. Brini, "Modeling of conductance fluctuations in small area metal oxide semiconductor transistors", *Phys. Stat. Sol. A*, vol. 127, p. 281, 1991.

[24] E. Simoen, B. Dierickx, C. Claeys, and G. Declerck, "Explaining the amplitude of RTS noise in submicrometer MOSFETs", *IEEE Trans. Electron Dev.*, vol. 39, p. 422, 1992.

[25] O. Roux dit Buisson, G. Ghibaudo, and J. Brini, "Model for drain current RTS amplitude in small-area MOS transistors", *Solid State Electron.*, vol. 35, p. 1273, 1992.

[26] X. Li and L. Vandamme, "$1/f$ noise in series resistance of LDD MOSTs", *Solid State Electron.*, vol. 35, p. 1471, 1992.

[27] E. Simoen, B. Dierickx, and C. Claeys, "Hot-carrier degradation of the random telegraph signal amplitude in submicrometer Si MOSTs", *Appl. Phys. A*, vol. 57, p. 283, 1993.

[28] M. H. Tsai and T. P. Ma, "$1/f$ noise in hot-carrier damaged MOSFET's: effect of oxide charge and interface traps", *IEEE Electron Dev. Lett.*, vol. 14, p. 256, 1993.

[29] O. Roux dit Buisson, "Etude de bruit en $1/f$ et des fluctuations RTS aux basses fréquences dans le transistor MOS submicronique". Ph.D. thesis, INP Grenoble, 1993 (in French).

[30] M. H. Tsai and T. P. Ma, "The impact of device scaling on the current fluctuations in MOSFET's", *IEEE Trans. Electron Dev.*, vol. 41, p. 2061, 1994.

[31] C. Jakobson, I. Bloom, and Y. Nemirovsky, "$1/f$ noise in CMOS transistors for analog applications from subthreshold to saturation", *Solid State Electron.*, vol. 42, p. 1807, 1998.

[32] E. Simoen and C. Claeys, "On the flicker noise in submicron silicon MOSFET", *Solid State Electron.*, vol. 43, p. 865, 1999.

[33] Y. Nemirovsky, I. Brouk, and C. G. Jakobson, "$1/f$ noise in CMOS transistors for analog applications", *IEEE Trans. Electron Dev.*, vol. 48, p. 921, 2001.

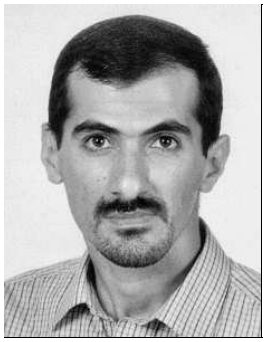
[34] A. K. Agarwal, M. C. Driver, M. H. Hanes, H. M. Hodgood, P. G. McMullin, H. C. Nathanson, T. W. O'Keeffe, T. J. Smith, J. R. Szedon, and R. N. Thomas, "MICROX^{TM}—an advanced silicon technology for microwave circuits up to X-band", *IEDM Tech. Dig.*, pp. 687–690, 1991.

[35] D. Eggert, P. Huebler, A. Huerrich, H. Kuerck, W. Budde, and M. Vorwerk, "A SOI RF-CMOS technology on high resistivity SIMOX substrates for microwave applications to 5 GHz", *IEEE Trans. Electron Dev.*, vol. 44, no. 11, pp. 1981–1989, 1997.

[36] O. Rozeau, J. Jomaah, S. Haendler, J. Boussey, and F. Balestra, "SOI technologies overview for low-power low-voltage radio-frequency applications", in *Analog Integrated Circuits and Signal Processing.* Kluwer, 2000, vol. 25.

[37] E. Simeon, U. Magnusson, A. L. P. Rotondaro, and C. Claeys, "The kink-related excess low frequency noise in silicon-on-insulator MOST's", *IEEE Trans. Electron Dev.*, vol. 41, pp. 330–339, 1994.

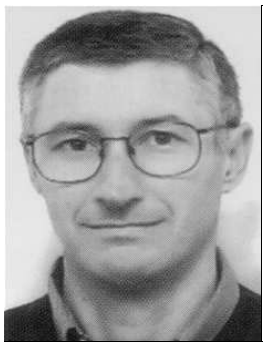
[38] Y.-C. Tseng, W. M. Huang, P. J. Welch, J. M. Ford, and J. C. S. Woo, "Empirical correlation between AC kink and low-frequency noise overshoot in SOI MOSFETs", *IEEE EDL*, vol. 19, no. 5, pp. 157–159, 1998.



Jalal Jomaah was born in Lebanon, in 1967. He graduated from the Institut National Polytechnique de Grenoble (INPG), France, in 1992. He received the M.Sc. and Ph.D. degrees in electronics from the same university, in 1992 and 1995, respectively. In 2002, he obtained the habilitation diploma from the INPG authorizing him to supervise Ph.D. dissertations. He joined the Laboratoire de Physique des Composants à Semiconducteurs (LPCS), INP Grenoble, in 1992, where he has been involved in research on the characterization, modeling, and simulation of fully- and partially-depleted silicon-on-insulator MOS transistors. He became Maître de Conférences (Associate Professor) at INPG in 1996, where he continues now his research activities at IMEP (Institute of Microelectronics, Electromagnetism and Photonics, INP Grenoble/CNRS/UJF). His main research activities were and are in the field of MOS/SOI device physics, fluctuations and low and high frequencies noise and radio-frequencies applications. He has supervised 6 Ph.D. thesis and was involved in several national and international research projects on the low and high frequency noise, reliability, modeling and characterization of SOI devices for RF applications and low temperature physics. Dr. Jomaah has coauthored over 35 publications in international scientific journals and 60 papers at international conferences (6 invited papers and review articles).
e-mail: jomaah@enserg.fr
IMEP-ENSERG (CNRS-INPG-UJF)
23 rue des Martyrs
BP. 257, 38016 Grenoble cedex 1, France

Francis Balestra was born in Digne, France, in 1960. He received the B.Sc. degree from the University of Provence, Marseille, France, in 1981 and the M.Sc. and Ph.D. degrees in electronics from the Institut National Polytechnique, Grenoble, France, in 1982 and 1985, respectively. In 1989, he obtained the habilitation diploma from the INPG authorizing him to supervise Ph.D. dissertations. He joined the Laboratoire de Physique des Composants à Semiconducteurs (LPCS), INP Grenoble, in 1982, where he has been involved in research on the characterization, modeling, and simulation of fully depleted silicon-on-sapphire MOS transistors. He became Chargé de Recherche C.N.R.S. (Centre National de la Recherche Scientifique) in 1985 and Directeur de Recherche CNRS in 2000. In 1993–94, he joined the Research Center for Integrated Systems at Hiroshima University as a visiting researcher, and worked on sub-0.1 μm MOSFETs and thin film SIMOX devices. Between 1996 and 1998 he has been Deputy Director of the LPCS, Director of the LPCS between 1999 and 2000, and he is now Director of IMEP (Institute of Microelectronics, Electromagnetism and Photonics, INP Grenoble/CNRS/UJF). He led several research teams on deep submicron CMOS, silicon on insulator devices, low temperature electronics and polysilicon emitter bipolar transistors for BiCMOS technology. He has supervised 15 research projects and 20 Ph.D. He has coordinated European, national and regional projects. Dr. Balestra has coauthored over 100 publications in international scientific journals, 150 communications at international conferences (20 invited papers and review articles), and 15 books or chapters.

e-mail: balestra@enserg.fr
IMEP-ENSERG (CNRS-INPG-UJF)
23 rue des Martyrs
BP. 257, 38016 Grenoble cedex 1, France



Gérard Ghibaudo was born in France, in 1954. He graduated from the Polytechnics Institute of Grenoble in 1979, obtained the Ph.D. degree in electronics in 1981 and the state thesis degree in physics from the same university in 1984. He became associate researcher at CNRS in 1981, where he is now Director of Research at Laboratories of Semiconductor Devices (LPCS/ENSERG now IMEP/ENSERG). During the academic year 1987–1988 he spent a sabbatical year at Naval Research Laboratory in Washington, DC, USA, where he worked on the characterization of MOSFETs. His main research activities were and are in the field of electronics transport, oxidation of silicon, MOS device physics, fluctuations and low frequency noise and dielectric reliability. Dr. G. Ghibaudo was or is a member of several technical/scientific committees of international conferences (ESSDERC'93, WOLTE, ICMTS, MIEL'95-2004, ESREF'96-98-00-03, SISC, MIGAS, ULIS, IEEE/IPFA, ...). He was co-founder of the First European Workshop on Low Temperature Electronics (WOLTE'94) and organizer of 8 Workshops/Summer School during the last 10 years. Dr. G. Ghibaudo is also member of the Editorial Board of *Solid State Electronics*. During his career he has been author or co-author of about 196 articles in international refereed journals, 310 communications and 35 invited presentations at international conferences, and of 12 book chapters.

e-mail: ghibaudo@enserg.fr
IMEP-ENSERG (CNRS-INPG-UJF)
23 rue des Martyrs
BP. 257, 38016 Grenoble cedex 1, France

Invited paper

Variability of the local ϕ_{MS} values over the gate area of MOS devices

Henryk M. Przewłocki, Andrzej Kudła, Danuta Brzezińska, and Hisham Z. Massoud

Abstract—The local value distributions of the effective contact potential difference (ECPD or the ϕ_{MS} factor) over the gate area of Al-SiO_2-Si structures were investigated for the first time. A modification of the photoelectric ϕ_{MS} measurement method was developed, which allows determination of local values of this parameter in different parts of metal-oxide-semiconductor (MOS) structures. It was found that the ϕ_{MS} distribution was such, that its values were highest far away from the gate edge regions (e.g., in the middle of a square gate), lower in the vicinity of gate edges and still lower in the vicinity of gate corners. These results were confirmed by several independent photoelectric and electrical measurement methods. A model is proposed of this distribution in which the experimentally determined $\phi_{MS}(x,y)$ distributions, found previously, are attributed to mechanical stress distributions in MOS structures. Model equations are derived and used to calculate $\phi_{MS}(x,y)$ distributions for various structures. Results of these calculations remain in agreement with experimentally obtained distributions. Comparison of various characteristics calculated using the model with the results of photoelectric and electrical measurements of a wide range of Al-SiO_2-Si structures support the validity of the model.

Keywords— MOS structure, photoelectric measurements, electrical parameters, mechanical stress.

1. Introduction

It has been previously shown (e.g., [1, 2]) that mechanical stress in metal-oxide-semiconductor (MOS) devices influences their electrical parameters. One of these electrical parameters is the effective contact potential difference (ECPD), called also the ϕ_{MS} factor. The dependence of ϕ_{MS} on stress was quantitatively estimated in [3]. On the other hand, it has been shown [2, 4, 5] that usually, mechanical stress is nonuniformly distributed under the gate of a MOS structure. In case of metal gate structures studied in this work, the mechanical stress distribution in the dielectric layer under the gate is characterized by a considerable compressive stress σ_0 in the central part of the gate and abrupt stress changes in the vicinity of gate edges, as shown in Fig. 1. It is therefore expected that ϕ_{MS} will have different values in the vicinity of gate edges and far away from them. This has not been observed until now, since conventional methods used to determine electrical parameters of MOS devices, including ϕ_{MS}, measure the parameter values which are averaged over the gate area. In this work, the lateral distributions of ϕ_{MS} values over the gate area were measured for the first time. A new measurement method was developed for that purpose, which is a modification of the photoelectric ϕ_{MS} measurement method [6], as described below.

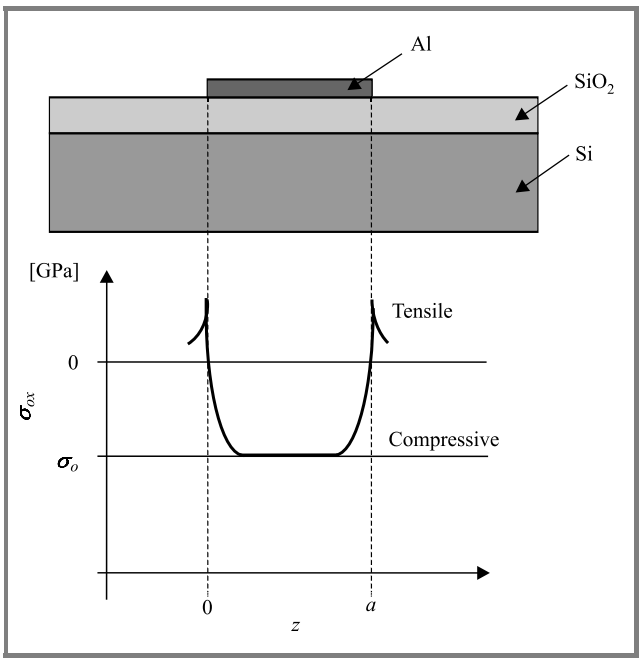


Fig. 1. The expected one-dimensional distribution of stress $\sigma_{ox}(z)$ in the oxide layer under the aluminum gate.

Photoelectric measurement results and their correlation with stress distributions are supported by electrical measurements of various MOS structures. A model of lateral distribution of ϕ_{MS} local values over the gate area has also been developed and is presented later. A range of measurement results fully supports the validity of the model.

2. Experimental

2.1. Measurement methods

Distributions of ECPD local values over the gate area were measured using a newly developed modification of the photoelectric ϕ_{MS} measurement method [6]. The principle of this modified photoelectric method is illustrated in Fig. 2. A focused beam of laser generated UV radiation illuminates a small fragment of the gate area, causing internal

photoemission to take place in this region of the MOS structure. The resulting photocurrent I versus gate voltage V_G characteristics can be taken in the external circuit. Analysis of these characteristics allows determination of the local ϕ_{MS} value in the illuminated region, as described in [6] and references therein. Hence, scanning the gate

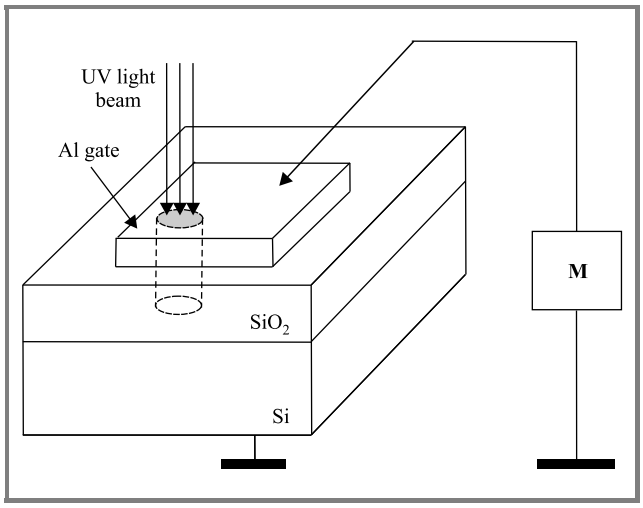


Fig. 2. Illustration of the principle of the modified photoelectric method of ϕ_{MS} determination.

with the UV light beam allows determination of the lateral ϕ_{MS} distribution over the gate area. Results obtained using this method are verified using the purely electrical $C(V)$ measurements and using the photoelectric measurements made on entire MOS capacitors of different shapes and dimensions, as described later.

2.2. Measured structures

In this work, measurements were made on Al/SiO_2/n^+-Si ($\rho \cong 0.015$ Ωcm) and on Al/SiO_2/n-Si ($\rho \cong 4$ Ωcm) capacitors of different geometries. MOS capacitors with different gate shape and area were characterized on each of the silicon wafers under investigation. The aluminum gates were either square shaped (SQ), with side $a = 0.1 \ldots 1.0$ mm, or were in the form of sets of narrow aluminum lines (STR). Although SiO_2 layers of current technological interest are thinner than $t_{ox} \cong 3$ nm, we used thicker oxides to optimize the sensitivity of the applied photoelectric methods [6]. MOS structures with three oxide thicknesses $t_{ox} = 20$, 60 and 160 nm, and two aluminum gate thicknesses $t_{\mathrm{Al}} = 35$ and 400 nm were used in this study.

2.3. Measurement results

First, the ECPD values were determined by the standard photoelectric method [6], for MOS capacitors with widely different values of R, where R is the ratio of the gate perimeter P to the gate area A of the MOS structure. Typical results of such measurements are shown in Fig. 3. It can be seen, that ϕ_{MS} values for entire MOS capacitors decrease monotonically with increasing R ratio and

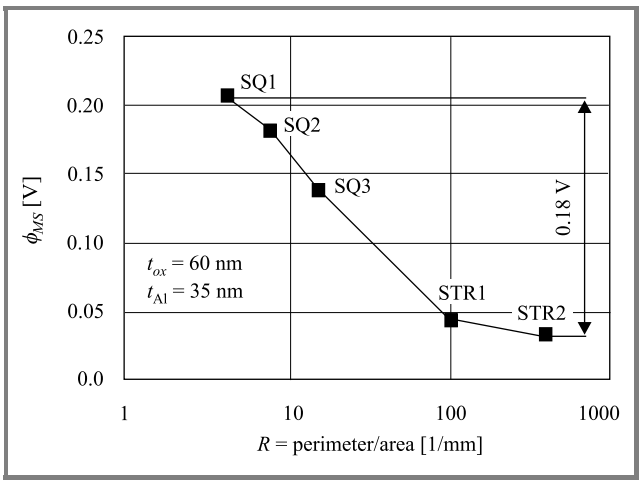


Fig. 3. The ϕ_{MS} measurement results (solid squares) for the entire SQ and STR gate Al-SiO_2-Si capacitors with different R ratios. Lines are drawn for eye guiding purposes.

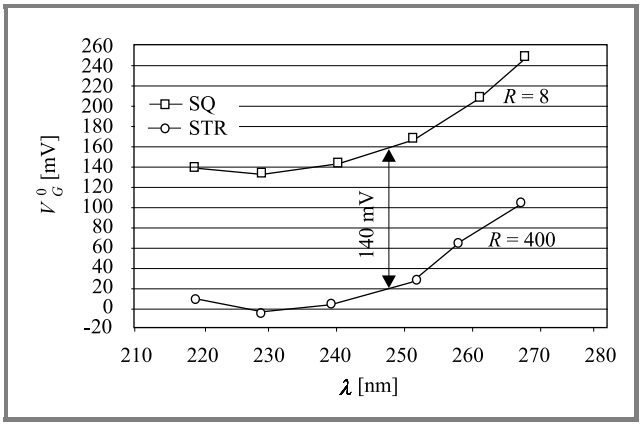


Fig. 4. Spectral characteristics of the zero-photocurrent-voltage $V_G^0(\lambda)$, taken for two Al-SiO_2-Si capacitors with different R values.

tend to saturate for large values of R. It was also found that the spectral characteristics of the zero-photocurrent-voltage $V_G^0(\lambda)$, taken for capacitors of different R values, are shifted in parallel against each other along the voltage axis, as shown in Fig. 4. Since for any MOS structure: $V_G^0(\lambda) = \phi_{MS} + \mathrm{const}(\lambda)$; [6], this result shows that for two structures with different R ratios, the difference between their ϕ_{MS} values can be determined at any wavelength λ within the used UV range. The $\phi_{MS}(R)$ dependence, illustrated in Figs. 3 and 4, suggests, although indirectly, that ϕ_{MS} values in the vicinity of gate edges are lower than far away from these edges. The direct proof of this property requires the use of measurement techniques, which allow determination of local ϕ_{MS} values in the regions, which are small in comparison with the overall dimensions of the gate. Hence, the modified photoelectric method, described above and illustrated in Fig. 2, was used for that purpose. The typical $V_G^0(z)$ characteristics (where z is a coordinate in the gate surface plane), taken using this method, are shown in Fig. 5, which proves that indeed ϕ_{MS} values at gate edges

are lower than in the middle of the gate, and that in the vicinity of gate corners the ϕ_{MS} values are still lower than near the gate edges.

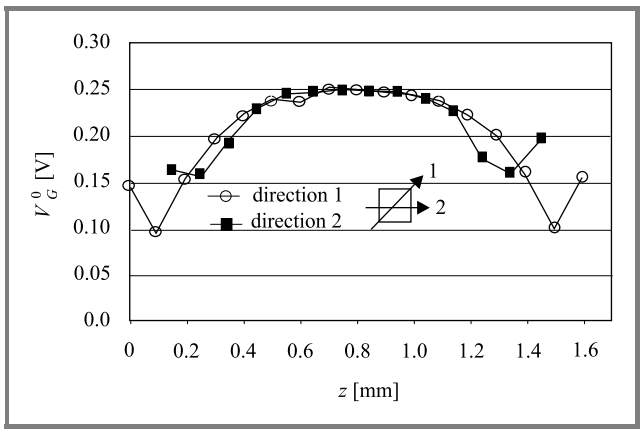


Fig. 5. Typical results of $V_G^0 = \phi_{MS} + \text{const.}$ local value measurements versus distance z, for SQ Al(35 nm)-SiO_2(60 nm)-Si(n^+) structure. Distance z is measured: 1) along the diagonal of the gate; 2) parallel to gate edges and through the middle of the gate.

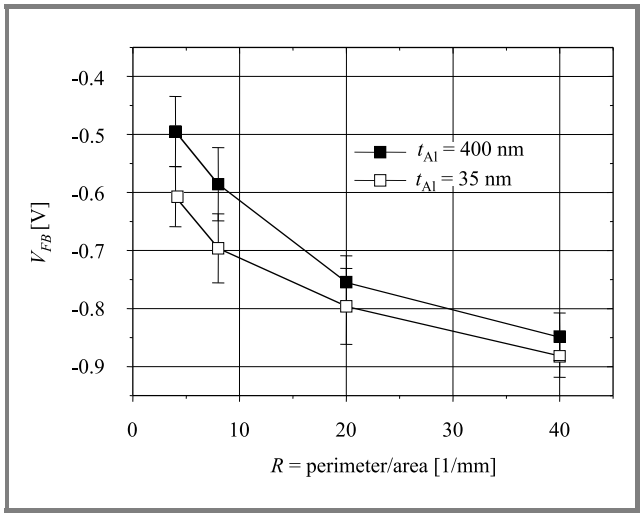


Fig. 6. Average V_{FB} values (squares) and their standard deviations (error bars) measured on SQ gate Al-SiO_2-Si capacitors with different R ratios. The solid squares are for thicker aluminum gate structures ($t_{Al} = 400$ nm), and the open squares are for thinner aluminum gate structures ($t_{Al} = 35$ nm). Lines are drawn as eye guides.

All the evidence of nonuniform ECPD distributions over the gate area, presented above, is based on photoelectric measurements. To exclude the possibility that the observed results are due to some unidentified optical effect, taking place when UV radiation illuminates the edges and corners of the gate, the results were confirmed by electrical measurements which do not involve the use of radiation. Taking into account that the ϕ_{MS} value directly influences that of the flat-band voltage V_{FB}, capacitance – voltage $C(V_G)$ characteristics were taken for MOS capacitors with different R ratio, and V_{FB} values were determined from these characteristics. This way V_{FB} values were obtained for thousands of Al/SiO_2/n-Si capacitors, with different R ratio, which were identically processed as the Al/SiO_2/n^+-Si structures measured by photoelectric methods, as described above. Results of these measurements consistently show that flat-band voltage values decrease with increasing values of R. Typical results of such measurements made on square gate (SQ) capacitors, with $R = 4$, 8, 20 and 40 mm^{-1} are shown in Fig. 6, for structures with aluminum gate thicknesses $t_{Al} = 35$ and 400 nm, fully confirming the conclusions drawn from results obtained by photoelectric methods.

3. Model

Stress distribution in the SiO_2 layer under the gate of a MOS capacitor, such as the distribution shown in Fig. 1, can be modeled in one dimension as:

$$\sigma(z) = \sigma_0 + \Delta\sigma(z) \tag{1}$$

with:

$$\Delta\sigma(z) = \Delta\sigma\{\exp(-z/L) + \exp[-(a-z)/L]\}, \tag{1a}$$

where σ_0 is the compressive stress in the central part of the gate, $\Delta\sigma = \Delta\sigma(0) = \Delta\sigma(a)$ is the maximum deviation of $\sigma(z)$ from the σ_0 value which occurs at gate edges (for $z = 0$ and $z = a$), a is the width of the gate in z-direction, and L is the characteristic length of the stress distribution. Assuming that the lateral ϕ_{MS} distribution, discussed above, is caused by lateral stress distributions, in such a way that $|\Delta\phi_{MS}(z)|$ is proportional to $\Delta\sigma(z)$, the one dimensional $\phi_{MS}(z)$ distribution can be expressed as:

$$\phi_{MS}(z) = \phi_{MS_0} + \Delta\phi\{\exp(-z/L) + \exp[-(a-z)/L]\}, \tag{2}$$

where ϕ_{MS_0} is the ϕ_{MS} value far away from gate edges, $\Delta\phi = \Delta\phi(0) = \Delta\phi(a)$ is the deviation of $\phi_{MS}(z)$ from the ϕ_{MS_0} value which occurs at gate edges (for $z = 0$ and $z = a$), and as results from experiment $\Delta\phi$ is negative. The average ϕ_{MS} value for the entire MOS capacitor, designated $\overline{\phi_{MS}}$ is given by:

$$\overline{\phi_{MS}} = \frac{1}{a}\int_0^a \phi_{MS}(z)\mathrm{d}z \tag{3}$$

and is a function of the a/L ratio.

The one dimensional analysis, presented so far, can be applied to some of the MOS capacitors of practical interest, e.g., to structures with gates in form of long and narrow stripes (STR), used in our experiments. However, in most of the cases, the problem considered is not of a one dimensional nature. For instance, for structures with square gates (SQ), of side length a, lying in the x, y plane, the $\Delta\sigma(x)$ and $\Delta\phi_{MS}(x)$ distributions are equally important as $\Delta\sigma(y)$ and $\Delta\phi_{MS}(y)$. Assuming that the properties

of the MOS system are isotropic in the x, y plane, these distributions can be expressed by Eqs. (1) and (2), with the z variable replaced by x and y. Assuming further that the principle of superposition can be applied to stress and ECPD distributions in the x, y plane, one obtains:

$$\Delta\phi_{MS}(x,y) = \Delta\phi_{MS}(x) + \Delta\phi_{MS}(y) \tag{4}$$

and

$$\phi_{MS}(x,y) = \phi_{MS_0} + \Delta\phi_{MS}(x,y)\,. \tag{5}$$

Equations (5) and (4), together with expressions for $\Delta\phi_{MS}(x)$ and $\Delta\phi_{MS}(y)$, analogous to Eq. (2), determine the $\phi_{MS}(x,y)$ distribution in the x, y plane of the gate area. Examples of $\phi_{MS}(x,y)$ distributions, calculated using the above equations, for structures with square gates (SQ) of different side lengths a are shown in Fig. 7.

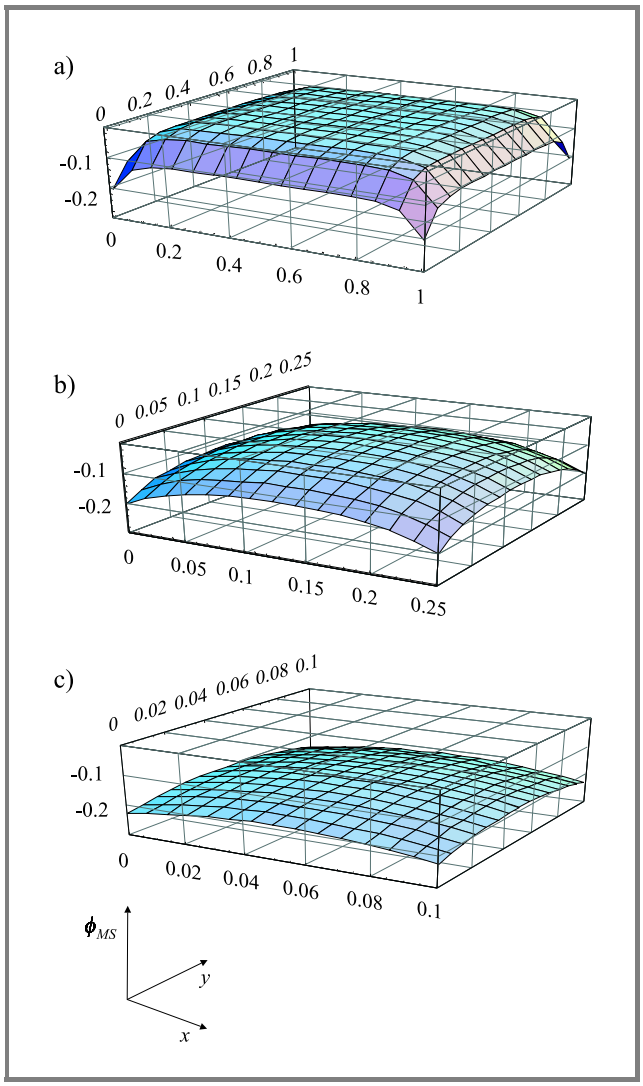


Fig. 7. The $\phi_{MS}(x,y)$ distribution calculated using the model, for SQ gate Al-SiO_2-Si capacitors of different side length a: (a) $a = 1$ mm; (b) $a = 0.25$ mm; (c) $a = 0.1$ mm, and using $\phi_{MS_0} = 0$ V, $\Delta\phi = -0.1$ V, $L = 0.05$ mm in the expressions for $\Delta\phi_{MS}(x)$ and $\Delta\phi_{MS}(y)$.

For such structures the average ϕ_{MS} value, designated $\overline{\phi_{MS}}$, is given by:

$$\overline{\phi_{MS}} = \frac{1}{a^2}\int_0^a\int_0^a \phi_{MS}(x,y)\mathrm{d}x\mathrm{d}y\,. \tag{6}$$

The integration of Eq. (6) yields:

$$\overline{\phi_{MS}} = \phi_{MS_0} + 4\,\frac{\Delta\phi\cdot L}{a}\left[1-\exp\left(-\frac{a}{L}\right)\right] \tag{7}$$

showing that $\overline{\phi_{MS}}$ is a function of a/L, as clearly seen in Fig. 7. Equation (7) can be transformed to explicitly give the dependence of $\overline{\phi_{MS}}$ on the previously introduced R ratio:

$$\overline{\phi_{MS}} = \phi_{MS_0} + \Delta\phi\cdot L\cdot R\left[1-\exp\left(-\frac{4}{L\cdot R}\right)\right]. \tag{8}$$

For capacitors with gates in form of long and narrow stripes of width a and length l, ($l \gg a$), similar reasoning leads to:

$$\overline{\phi_{MS}} = \phi_{MS_0} + \Delta\phi\cdot L\cdot R\left[1-\exp\left(-\frac{2}{L\cdot R}\right)\right]. \tag{9}$$

The dependence of $\overline{\phi_{MS}}$ on R should be reflected in the dependence of the flat-band voltage $\overline{V_{FB}}$ (as measured on the entire MOS structure) on R. Assuming that the effective charge of the MOS system does not significantly depend on R, the $\overline{V_{FB}}(R)$ dependence for square gate structures is:

$$\overline{V_{FB}} = V_{FB_0} + \Delta V\cdot L\cdot R\left[1-\exp\left(-\frac{4}{L\cdot R}\right)\right], \tag{10}$$

where V_{FB_0} is the local $V_{FB}(x,y)$ value far away from the gate edges, and $|\Delta V|$ is maximum deviation of $V_{FB}(x,y)$ from the V_{FB_0} value (ΔV is negative). Equations (8)–(10) are the equations of the model which can be used to determine $\overline{\phi_{MS}}(R)$ and $\overline{V_{FB}}(R)$ dependencies in MOS structures, as shown below.

4. Verification

To verify the model, experimental results were fit with the model equations. The best fit of model equations to the measurement results yields the model parameters for a given set of experiments, i.e., the values of ϕ_{MS_0}, $\Delta\phi$ and L, or V_{FB_0}, ΔV and L. As an example, model Eqs. (8) and (9) were fit to the ϕ_{MS} measurements illustrated in Fig. 3, and the result is shown in Fig. 8. The fit is good both for the square (SQ) gate structures (Eq. (8)) and for the (STR) structures with gates in form of narrow aluminum stripes (Eq. (9)). It should be noted that in both cases the fit is obtained for the same values of ϕ_{MS_0} and $\Delta\phi$ which strongly support the validity of the model. The (effective) L value is different for SQ and STR structures, since for STR structures L is comparable to the width a of the aluminum lines and the influences of both line edges overlap.

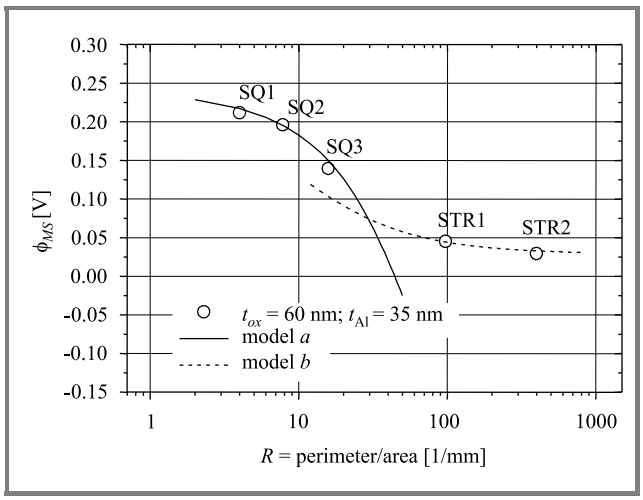


Fig. 8. Curves calculated using Eqs. (8) and (9) fit to the measurement results shown in Fig. 3, obtained for Al-SiO_2-Si capacitors with SQ and STR gates. Both curves are calculated using $\phi_{MS_0} = 0.236$ V, and $\Delta\phi = -0.106$ V in Eqs. (8) and (9). The curve calculated for SQ structures (model *a*) uses $L = 0.056$ mm in Eq. (8), and the curve calculated for STR structures (model *b*) uses $L = 0.090$ mm in Eq. (9).

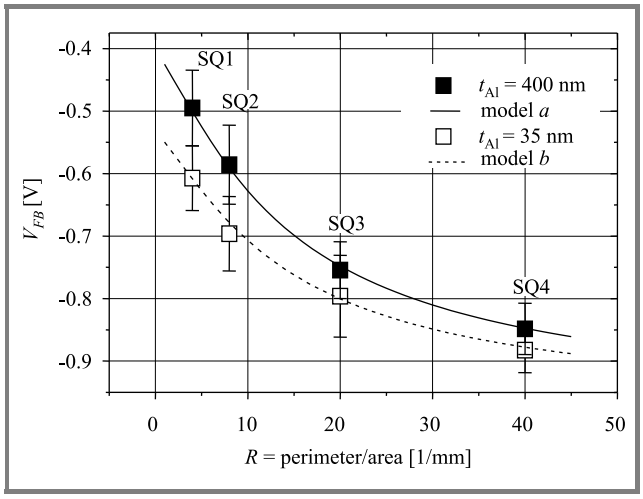


Fig. 9. Curves calculated using Eq. (10) fit to the V_{FB} measurement results shown in Fig. 6, for SQ gate Al-SiO_2-Si capacitors with different R ratio and different aluminum gate thickness. The upper curve (for $t_{Al} = 400$ nm) is calculated using $V_{FB_0} = -0.400$ V, $\Delta V = -0.148$ V, and $L = 0.170$ mm in Eq. (10). The lower curve (for $t_{Al} = 35$ nm) is calculated using $V_{FB_O} = -0.530$ V, $\Delta V = -0.115$ V, and $L = 0.170$ mm in Eq. (10).

As a result, the shape of the $\phi_{MS}(z)$ distribution is quite different in this case from the distributions obtained in case when $a \gg L$ (e.g., for large SQ structures).

The model was further verified by fitting Eq. (10) to the V_{FB} measurement results. As an example, Eq. (10) is fit to the measurement results illustrated already in Fig. 6, as shown in Fig. 9. The fit is good, both for the capacitors with thicker and the ones with thinner aluminum gates. It should be noted that the best fit gives higher values of both V_{FB_0} and $|\Delta V|$ for structures with thicker gates. This is the result that was expected since thicker aluminum layers cause higher compressive stresses σ_0 in the SiO_2 layer and larger $\Delta\sigma$ deviations at gate edges, supporting the assertion that the lateral distributions of ϕ_{MS} and V_{FB} values are caused by lateral distributions of stress σ_{ox} in the plane of the gate. Further support for this assertion was gained from an experiment based on the following reasoning. Any change in the stress distribution under the MOS system gate, should be reflected in changes of ϕ_{MS} and V_{FB} distributions. A change of stress distribution can be introduced if a gate of another MOS structure is placed close enough to the gate of investigated structure. Hence, a comparison was made between the $V_{FB}(R)$ characteristics of square gate capacitors, which were not surrounded by other capacitors in their immediate vicinity, with the characteristics of identical capacitors surrounded by protective aluminum rings of the same thickness $t_{Al} = 400$ nm, at a distance of 10 μm. The result

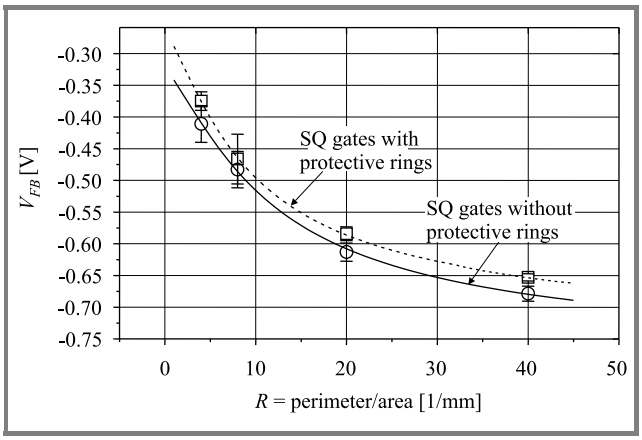


Fig. 10. The V_{FB} versus R measurement results obtained for SQ gate structures with and without the protective aluminum rings (open squares and open circles). Curves calculated using Eq. 10) are fit to measurement results for both types of structures. The upper curve (for SQ gates with protective rings) is calculated using $V_{FB_0} = -0.259$ V, $\Delta V = -0.119$ V and $L = 0.249$ mm in Eq. (10). The lower curve (for SQ gates without protective rings) is calculated using $V_{FB_0} = -0.319$ V, $\Delta V = -0.115$ V, and $L = 0.197$ mm in Eq. (10).

of this comparison is shown in Fig. 10. Similarly as in the case of capacitors with different aluminum gate thickness (see Fig. 9), a considerable shift is observed between the $V_{FB}(R)$ characteristics of structures with and without the protective aluminum rings.

5. Conclusions

The lateral distribution of local ECPD values in the plane of MOS structure's gate was studied for the first time. Photoelectric and electrical measurement methods were used in this investigation. It was found that ϕ_{MS} has a characteristic distribution over the gate area, with local ϕ_{MS} values being

highest in the middle of the gate, lower at gate edges, and lowest at gate corners. To study these ECPD distributions, the photoelectric ϕ_{MS} measurement method was modified to allow determination of local ECPD values over dimensions that are small in comparison with the dimensions of the MOS structure. The lateral distributions of ECPD are attributed to lateral distributions of mechanical stress under the gate of a MOS structure. A simple model has been proposed of the distribution of local ϕ_{MS} values over the gate area. The $\phi_{MS}(x, y)$ distributions calculated using the model remain in agreement with the ones determined by the photoelectric method. The validity of the model is further verified by electrical measurements of the flat-band voltage V_{FB} values of a range of Al-SiO_2-Si capacitors, differing in the ratio R, of the gate perimeter to gate area.

References

[1] C. H. Bjorkman, J. T. Fitch and G. Lucovsky, "Correlation between midgap interface state density and thickness-averaged oxide stress and strain at SiO_2 interfaces formed by thermal oxidation of Si", *Appl. Phys. Lett*., vol. 56, pp. 1983–1985, 1990.

[2] S. M. Hu, "Stress related problems in silicon technology", *J. Appl. Phys*., vol. 70, pp. R53–R80, 1991.

[3] H. M. Przewłocki and H. Z. Massoud, "The effects of stress annealing in nitrogen on the effective contact potential difference (ECPD), charges and traps at the Si/SiO_2 interface of MOS devices", *J. Appl. Phys*., vol. 92, pp. 2198–2201, 2002.

[4] I. De Wolf, H. E. Maes, and S. K. Jones, "Stress measurements in silicon devices through Raman spectroscopy: bridging the gap between theory and experiment", *J. Appl. Phys*., vol. 79, pp. 7148–7156, 1996.

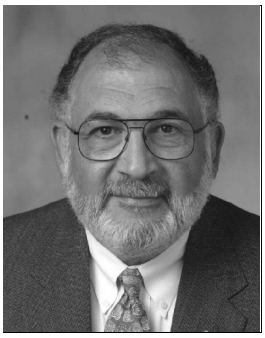
[5] K. F. Dombrowski, I. De Wolf, and B. Dietrich, "Stress measurements using ultraviolet micro-Raman spectroscopy", *Appl. Phys. Lett*., vol. 75, pp. 2450–2451, 1999.

[6] H. M. Przewłocki, "Theory and applications of internal photoemussion in the MOS system at low electric fields", *Sol. St. Electron*., vol. 45, pp. 1241–1250, 2001.

Henryk M. Przewłocki obtained the M.Sc., Ph.D., and D.Sc. degrees in 1959, 1969, and 2001, respectively, specializing in physics, technology, and measurements of MOS structures. In the period of 1965–1983, he also taught various courses to graduate and undergraduate students of the Technical University of Warsaw. At present he is the head of a research group at the Institute of Electron Technology (IET) in Warsaw, specializing in photoelectric properties and measurement methods of MOS structures. He has served as consultant and committeeman to numerous scientific, industrial and educational organizations in Poland and worldwide and has received a number of prestigious awards, including the highest Polish award, the National Award for Outstanding Achievements in Technical Sciences.
e-mail: hmp@ite.waw.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Andrzej Kudła graduated from the Faculty of Mechatronics in 1985 and joined the Department of Silicon Structure Technology of the Institute of Electron Technology (IET) in 1986. Since 1989 to 1991 he was with the ASIC Department of ATMOS ELPOL and since 1991 he is with the Department of MOS Structure Studies of IET. He received the Ph.D. degree in micreoelectronics in 1997. He is involved in optical studies with special emphasis on spectroscopic ellipsometry as well as the development of photoelectric methods of studying MOS structures. In 2001 his team received the "Master of Technology 2001" award of Polish Federation of Engineering Associations for a Multi-task System for Photoelectric Studies. He is the author or co-author of approximately 70 publications.
e-mail: kudla@ite.waw.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Danuta Brzezińska graduated from the Warsaw University of Technology, Department of Telecommunication, in 1959, receiving the M.Sc. degree in electronics. She started work in 1957 in the Industrial Institute of Electronics, where she was involved, among others, in developing thin film technology for microwave circuits. In 1968 she was granted the UN fellowship in Great Britain for six months, where she developed the sputtering techniques and equipment for thin film deposition. In 1972 she joined the Institute of Electron Technology, where she developed semiconductor devices and IC technology. She was a member of a team which received the title "Champion of Technology – Warsaw 1982". Recently she did research in physics of the MOS system, in particular, using photoelectric methods. She co-authored several papers in this field.
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland



Hisham Z. Massoud received the B.Sc. and M.Sc. degrees in electrical engineering from Cairo University, Cairo, Egypt, in 1973 and 1975, respectively, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1976 and 1983, respectively. He was at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY, in 1977 and 1980/81, where he investigated electron and hole trapping at the silicon-oxide/silicon interface at room temperature and 77 K, and the in-situ ellipsometry monitoring of the thermal growth of ultrathin silicon dioxide layers. In December 1983, he joined the Electrical and Computer Engineering Department at Duke University, Durham, NC, where he is currently Professor and Director of the Semiconductor Research Laboratory, and where he served as its Interim Chair in 2000/2001. He spent the summer of 1987 at the Microelectronics Center of North Carolina, and the summers of 1996 and 1997 at the Max-Planck Institute for Microstructure Physics, Halle, Germany. His research interests are in the areas of the thermal and rapid-thermal oxidation of silicon, silicon/silicon-dioxide interface studies, carrier tunneling studies in MOS devices with ultrathin oxides, and the effects of tunneling on MOS integrated circuits. Dr. Massoud is a Fellow of the Institute of Electrical and Electronics Engineers, and member of the American Physical Society, the Amercian Vacuum Society, the Electrochemical Society, the Materials Research Society, the American Association for the Advancement of Science, Eta Kappa Nu, Tau Beta Pi, and Sigma Xi.

e-mail: massoud@ee.duke.edu
Department of Electrical and Computer Engineering
Duke University
Durham, North Carolina 277708–0291, USA

Invited paper

Diagnostics of micro- and nanostructure using the scanning probe microscopy

Teodor Gotszalk, Paweł Janus, Andrzej Marendziak, Piotr Czarnecki, Jacek Radojewski, Roman F. Szeloch, Piotr Grabiec, and Ivo W. Rangelow

Abstract—In this paper we summarize the results of our research concerning the diagnostics of micro- and nanostructure with scanning probe microscopy (SPM). We describe the experiments performed with one of the scanning probe microscopy techniques enabling also insulating surfaces to be investigated, i.e., atomic force microscopy (AFM). We present the results of topography measurements using both contact and non-contact AFM modes, investigations of the friction forces that appear between the microtip and the surface, and experiments connected with the thermal behaviour of integrated circuits, carried out with the local resolution of 20 nm.

Keywords— scannig probe microscopy, microsystem, nanofabrication.

1. Introduction

The development of the scanning probe microscopy (SPM) investigation methods that began with the invention of the scanning tunnelling microscope in 1982 [1] made it possible to observe the structure of conductive and insulating surfaces in the air, in a nanometer scale for the first time. The so-called nearfield interactions that occur between the microtip and the surface at a distance of several nanometers (e.g., tunnelling current or heat flux flowing between the microprobe and the sample, force interaction between atoms of the tip and the surface) are observed in these measurement techniques and utilized for versatile structure characterization.

The SPM instruments have been used successfully at the universities and research institutes with for the last few years to carry out research in the nanotechnology area. Simultaneously the SPM methods were introduced in the semiconductor industry to measure nanometer-sized structures. This trend was driven, of course, by the fact that no other experimental method could provide as much information (e.g., surface roughness, elasticity, semiconductor dopant profiling, line width) on the fabricated devices.

The proper analysis of the acquired images requires however the knowledge of physical phenomena that are observed between the microtip and the sample.

2. Scanning probe microscopy experiments

For closely spaced atoms of the microtip and the surface the interaction energy is described by the Lennard-Jones potential [2]. The resulting force is illustrated in Fig. 1. There are two regions corresponding to the basic atomic force microscopy (AFM) measurement modes. In a contact AFM (C AFM) instrument the microtip touches the investigated surface. In this case the force between the microtip and the sample, which ranges typically from 1 nN up to 100 nN, is repulsive and causes the static cantilever deflection. The C AFM methods are mostly applied for the measurements of mechanical surface parameters. This technique has been developed to probe the viscoelastic and anelastic properties of submicron phases of inhomogenous materials. The measurement yields the information related to the internal friction and to the variations of the dynamic modulus of nanometer-sized volumes.

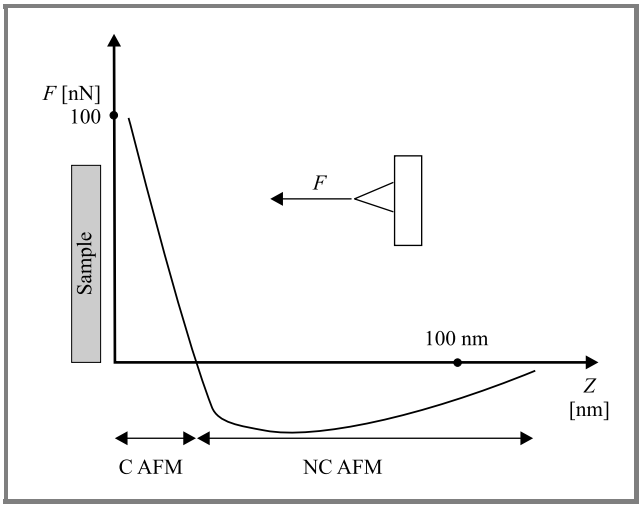


Fig. 1. Atomic force microscopy forces between the microtip and the sample.

The other C AFM method is the so-called lateral force microscopy (LFM), in which not only the vertical beam deflection but also the beam torsion are monitored. It should be noted that in classical tribology, all the investigations are carried out under heavily-loaded conditions. Therefore bulk

properties of the investigated sample dominate its tribological characteristics. In contrast to the macrotribological systems, all micro- and nanosystems operate under very light loads (few μN). In this case friction and wear are influenced only by the interactions of a few surface atom layers, therefore nanometer resolution of the investigations is required. Consequently, C AFM methods with LFM techniques are the appropriate tools for surface characterization.

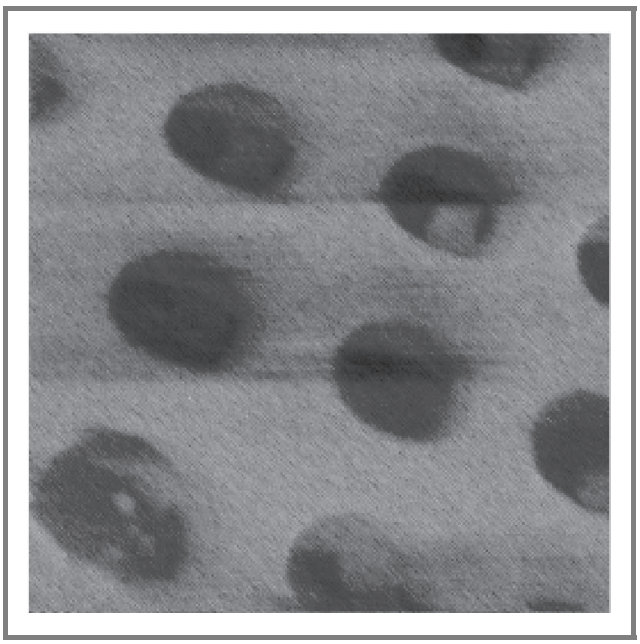


Fig. 2. Topography of a quartz/chromium point structure (scanfield 7×7 μm, structure height 120 nm, scanfield 2 lines/s).

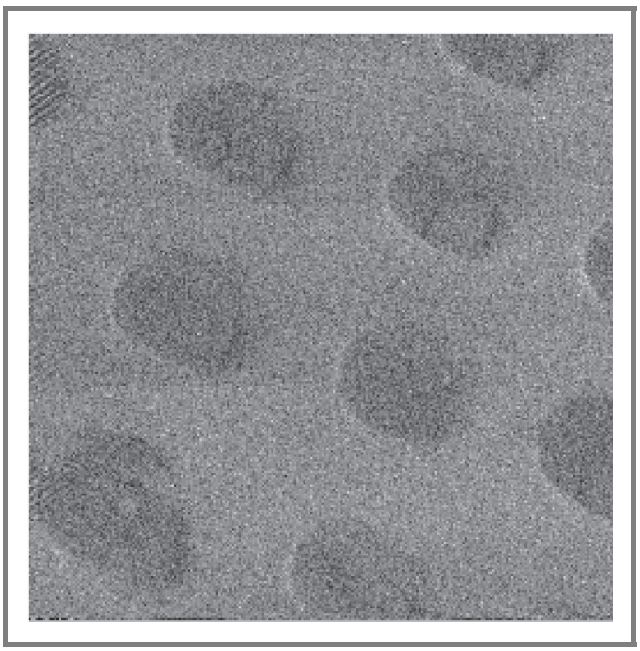


Fig. 3. Friction forces observed on a quartz/chromium point structure (load force 1 μN, maximum friction force 100 nN).

The LFM methods using the piezoresistive cantilevers were applied in the investigation of a gridlike structure of chromium dots on a quartz substrate [3]. The dots were about 120 nm high (Fig. 1). In order to perform simultaneous AFM and LFM measurements the signals corresponding to the topography and friction forces acting on the microtip were recorded. The lateral force image (Fig. 2) shows that friction forces vary on the sample surface. Higher friction was detected on the chromium dots, as shown by brighter locations in the LFM image (Fig. 3). The C AFM methods are also utilized for precise, high resolution, quantitative topography investigations. The quantitative measurement of the structure line width (the so-called critical dimension (CD) measurements) requires however the integration of the C AFM machine with the detection system of the cantilever or sample scanning movements. There are several methods to observe the deflection of the microscope piezoelectrical actuators that move the sample or cantilever in the range from 100 nm up to 100 μm with the resolution of 20 nm and the accuracy of 10 nm. In our experiments we applied the optical fiber interferometry to measure the scanning movements of the sample. Based on the observed interferometric fringes we can determine the piezoactuator deflections with the resolution of 25 nm. In Fig. 4 we show the topography of a structure consisting of chromium lines deposited on a glass substrate. The determined line width estimated based on the interferometric measurements is 1030 nm.

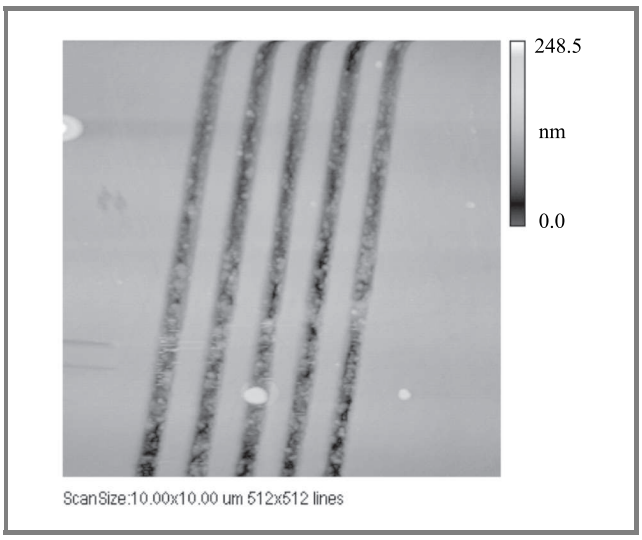


Fig. 4. Topography measurements using calibrated C AFM-quartz etched with reactive ion etching (linewidth 1 μm).

Scanning thermal microscopy (SThM), based on C AFM, is widely applied in the investigations of thermal behaviour of micro- and nanostructures. In this method a microtip containing a heat flux sensor measures locally the temperature of the sample.

The SThM system developed at the Wrocław University of Technology enables temperature changes of 20 mK to be measured with spatial resolution of 20 nm. In Fig. 5 the temperature distribution on the surface of integrated ciruit (IC) consisting of four resistors is presented. In our experiments the Wheatstone bridge formed by these resis-

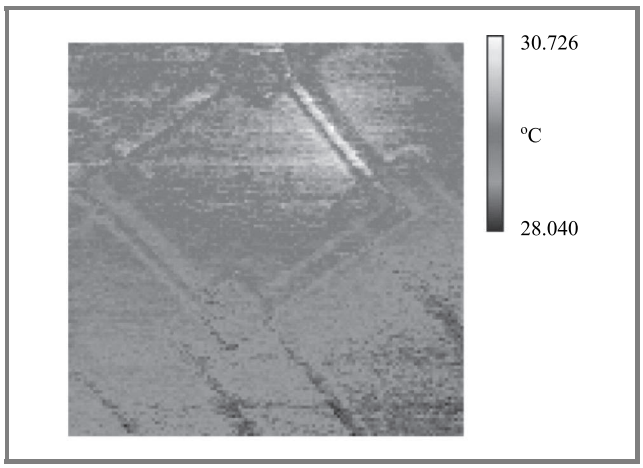


Fig. 5. Scanning thermal microscopy measurement of the temperature distribution of an integrated resistive bridge loaded asymetrically.

tors was supplied unsymetrically. The temperature in the region where most of the energy was dissipated is higher by 2 K than in the area of the other three resistors.

The experiments performed show the possibility of applying SThM methods in reliability investigations of micro- and nanostructures [5].

In the second region of the force-distance curves (Fig. 1), where the spacing between the tip and the surface is of 10–100 nm, very weak attractive interactions in the range of 0.1–0.05 nN are observed. Since the level of the interactions appearing in non-contact AFM (NC AFM) systems the force detection method is more elaborate. In this case the microscope cantilever vibrates at its mechanical resonance. The change of the beam resonance frequency or cantilever oscillation amplitude under the influence of the attractive force acting on the microtip is monitored and applied to control the probe-sample distance while sample scanning. The NC AFM techniques can be applied in the investigations of surface topography of soft materials, like, e.g., photoresists and biological samples. Because of high sensitivity of the force interaction measurements (enabling forces in the range of 0.01 nN to be detected), in NC AFM methods not only the topography investigations are performed but also electrostatic forces between the microtip and the surface can be monitored and applied for structure characterization. In this way the electrostatic force microscopy (EFM) has the potential to be a very promising analysis method for measurements of the properties of nanometer-size semiconductor devices and materials, such as dopant profiles, high-k insulator thickness. All applications indicated in literature show that EFM and SCM would be effective as a measurement tool for semiconductor devices that are continuously miniaturized. The measurement possibilities of the EFM based methods depend, however, strongly on the parameters of the applied probe. In our experiments we applied piezoresistive cantilevers [4] with conductive probes that enable voltage of 40 mV to be measured in the bandwidth of 30 Hz, and the silicon cantilevers with metallic tips for

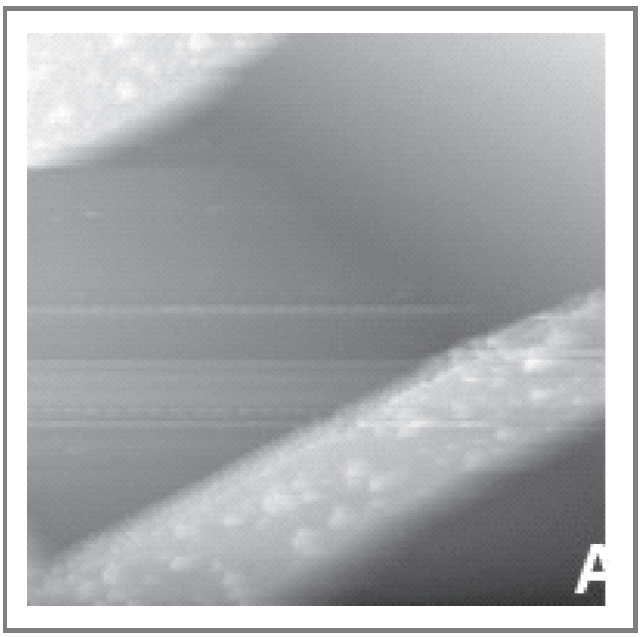


Fig. 6. Topography of an Al-SiO_2 sample observed with the EFM microscope (scanfield 45×45 μm, the structure height 300 nm).

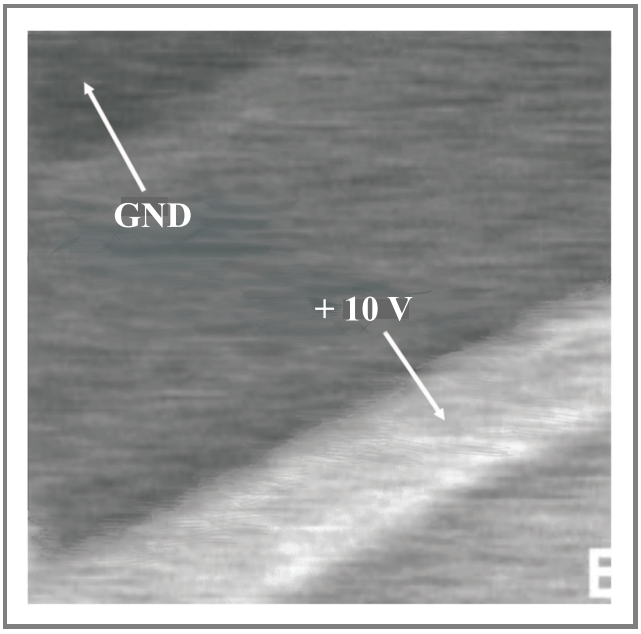


Fig. 7. Electrostatic force contrast observed on the Al-SiO_2 sample.

the measurements of voltages with the resolution of 30 mV in the bandwidth of 30 Hz [6]. In our investigations we applied a voltage $U = U_{dc} + U_{ac}\sin(\omega t)$ between the tip and the sample, which gives rise to an electrostatic force given by:

$$F = \frac{1}{2}\frac{dC}{dz}\left[(U_{dc}^2 + \frac{1}{2}U_{ac}^2) + 2U_{dc}U_{ac}\sin(\omega t) + \frac{1}{2}U_{ac}^2(1 - \sin(2\omega t)\right] = \frac{1}{2}\frac{dC}{dz}(F_{dc} + F_{\omega} + F_{2\omega}), \quad (1)$$

where C is tip-sample capacitance, ω is the frequency.

Using lock-in techniques we were able to distinguish forces corresponding to the potential at the structure surface and interactions corresponding to the variable capacitance between the microprobe and the sample. The topography of the sample formed by two Al lines deposited on the SiO_2 substrate measured with EFM microscope is presented in Fig. 6. The scan size is 45×45 μm, the structure height is of 300 nm. In our experiments we applied the voltage of 10 V to the metal strips. The image of electrostatic forces measured simultaneously with the topography is shown in Fig. 7. The dark region of the image corresponds to lower electrostatical force and grounded Al line, the brighter shade shows where the voltage of 10 V applied to the metal line was recorded.

3. Conclusions

In this paper we have presented the results of the experiments connected with the application of SPM methods in versatile diagnostics of micro- and nanostructure. The presented experiments shown that with the increasing innovations in SPM technology, the SPM instruments will be applied not only in university research techniques but also in industrial quality control and application measurements.

Acknowledgments

This work was partly funded by the Wrocław University of Technology Center of Advanced Materials and Nanotechnology grant no. 331939 and State Committee for Scientific Research grant no. 4T11B0563. We would like to thank A. Sankowska, J. Radojewski, and A. Sikora for fruitful discussion and support.

References

[1] G. Binnig, H. Rohrer, and E. Weibel, "Surface studies by scanning tunneling microscopy", *Phys. Rev. Lett*., vol. 49, p. 726, 1982.

[2] G. Binnig, C. F. Quate, and Ch. Gerber, "Atomic force microscopy", *Phys. Rev. Lett.*, vol. 56, p. 930, 1986.

[3] T. Gotszalk, P. Grabiec, and I. W. Rangelow, "Piezoresistive sensors for scanning probe microscopy", *Ultramicroscopy*, vol. 82, p. 39, 2000.

[4] T. Gotszalk, F. Shi, P. Grabiec, P. Dumania, P. Hudek, and I. W. Rangelow, "Fabrication of the multipurpose piezoresistive Wheatstone bridge cantilever with conductive microtip for electrostatic force and scanning capacitance microscopy", *J. Vac. Sci. Technol.*, vol. B16, no. 6, p. 3948, 1998.

[5] R. F. Szeloch, T. Gotszalk, and P. Janus, "Scanning thermal microscopy in microsystem reliability analysis", *Microelectron. Reliab*., vol. 42, p. 1719, 2002.

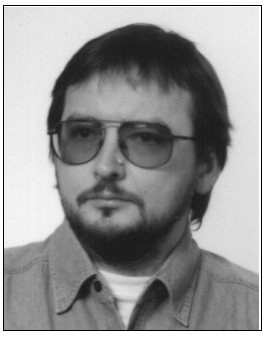
[6] T. Gotszalk, P. Czarnecki, P. Grabiec, K. Domański, M. Zaborowski, and I. W. Rangelow, "Microfabricated cantilever with metallic tip for electrostatic and capacitance microscopy and its application to investigation of semiconductor devices", *J. Vac. Sci. Technol.*, vol. B1622, p. 506, 2004.

Teodor Paweł Gotszalk was born in Wrocław, Poland, in 1966. He received the M.Sc. degrees from the Faculty of Electronics and from the Faculty of Electrical Engineering of Wrocław University of Technology in 1989 and 1991, respectively. In 1996 he received the Ph.D. degree from the Institute of Electronic Technology of the Wrocław University of Technology. Dr. Gotszalk has been honoured with Siemens research award (2000) and the prize of Polish Science Foundation FNP (1997) for his scientific work. He is the leader of the Scanning Probe Microscopy, Nanostructure and Nanometrology Group at the Faculty of Microsystem Electronics and Photonics of Wrocław University of Technology. He is the author and co-author of 80 scientific publications.
e-mail: teodor.gotszalk@pwr.wroc.pl
Faculty of Microsystem Electronics and Photonics
Wrocław University of Technology
Janiszewskiego st 11/17
50-372 Wrocław, Poland

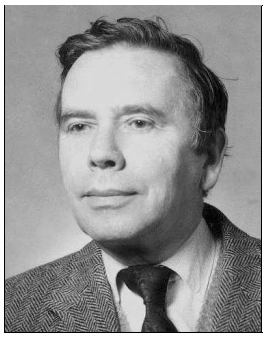
Paweł Janus was born in Żary, Poland, in 1973. He received the B.Sc. and M.Sc. degrees from the Technical University in Wrocław, Poland, in 1996 and 1998, respectively. In 2003 he received the Ph.D. degree in electrical engineering from the Faculty of Microsystem Electronics and Fotonics, Wrocław University of Technology, Poland. The thesis concerned the near- and far-field thermography in microsystems diagnostics. Since 2003, he has been with the Institute of Electron Technology (IET), Warsaw. His research interests include studies of materials for MEMS applications, micromachining of silicon microstructures, microsensors, and microactuators. His current research is silicon force sensors development for nano-mechanical applications. Dr. Janus has been active in the field of solid-state sensors and actuators and has participated in several conferences and workshops dealing with solid-state sensors and actuators. He is an author and co-author of more than 20 technical papers presented in journals and at conferences.
e-mail: janus@ite.waw.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Andrzej Marek Marendziak was born in Lubin, Poland, in 1975. He received the M.Sc. degree from the Faculty of Optoelectronics of Wrocław University of Technology in 2000. He is a member of the Scanning Probe Microscopy, Nanostructure and Nanometrology Group at the Faculty of Microsystem Electronics and Photonics of Wrocław University of Technology. He is the author and co-author of more then 10 scientific publications.
e-mail: andrzej.marendziak@pwr.wroc.pl
Faculty of Microsystem Electronics and Photonics
Wrocław University of Technology
Janiszewskiego st 11/17
50-372 Wrocław, Poland

Piotr Czarnecki was born in Szczecin, Poland, in 1977. In 2002 he graduated from Wrocław University of Technology, Faculty of Electronics, with the degree of M.Sc. in microsystem technology. In the same year he began Ph.D. studies at Wrocław University of Technology, Faculty of Microsystem Electronics and Photonics. His research interests: field electron emission from carbon nanotubes, atomic force microscopy, scanning electron microscopy.
e-mail: piotr.czarnecki@pwr.wroc.pl
Faculty of Microsystem Electronics and Photonics
Wrocław University of Technology
Janiszewskiego st 11/17
50-372 Wrocław, Poland



Jacek Mikołaj Radojewski was born in Wrocław, Poland, in 1956. In 1980 he obtained the M.Sc. degree from the Institute of Electron Technology, Electronic Faculty of Wrocław University of Technology. In 1992 he received the Ph.D. degree at the same faculty. Dr. Jacek Radojewski received awards of the Ministry of Higher Education (in 1987 and 1990) for series of scientific papers on optoelectronics and integrated optics. Currently he is the Head of Optoelectronics and Fiber Optics Laboratory. His research interests include optoelectronics and fiber optic techniques, near field optical microscopy and photon scanning tunneling microscopy.
e-mail: jacek.radojewski@pwr.wroc.pl
Faculty of Microsystem Electronics and Photonics
Wrocław University of Technology
Janiszewskiego st 11/17
50-372 Wrocław, Poland



Roman F. Szeloch has been working in the field of electronics and microsystem technology for more then 20 years and has been especially interested in the reliability of electronic components, and reliability physics and analysis. Recently his chief interest has been in studying thermal effects in micro and nanosystems and application of atomic force microscopy (AFM) and scanning thermal microscopy (SThM) for nanoscale characterisation of microsystems and micromaterials. His efforts are concentrated on linking thermal data with local distributions of defects and, in the next step, processes of degradation in microsystems.
e-mail: roman.szeloch@pwr.wroc.pl
Faculty of Microsystem Electronics and Photonics
Wrocław University of Technology
Janiszewskiego st 11/17
50-372 Wrocław, Poland

Piotr B. Grabiec received the M.Sc. and Ph.D. degrees in chemistry from Warsaw University of Technology in 1973 and 1985, respectively. In 1974 he joined the Institute of Electron Technology (IET), Warsaw, Poland, where he was involved in research on CVD and diffusion technology. His current research interests are silicon microfabrication, photodetectors, MEMS and MOEMS with special emphasis on integration of different technologies. He is a member of IEEE Electron Devices Society and a member of Electrochemical Society. He is an author of 150 papers and more than 100 conference presentations, and holds 10 patents.
e-mail: grabiec@ite.waw.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Ivo W. Rangelow has a long experience in micro- and nanoscience and technology and is actually working at the Institute for Microsystem Technology and Analysis, University Kassel, Germany. His current research interest includes management of micro- and nanostructuring branch, principal scientist on projects in the area of nanostructuring, mamolithography, pattern transfer in nanoscale, MEMS, proximal nanoprobes, carged particle projection lithography, semiconductor analysis, vacuum microelectronics, solid-gas phase reactions, semiconductor and nanodevices, quantum devices, sensors/actuators and embedded systems. Responsibilities include acquisition and management of new projects, supervision of master and Ph.D. thesis projects and lecturing. He acts as organizer and chairperson of numerous international conferences and as reviewer in scientific journals and for funding institutions. He has more than 300 publications, 42 invited conference presentations, 97 contributed conference reports, and 34 patents.

e-mail: rangelow@uni-kassel.de
Institute of Technical Physics
University of Kassel
Heinrich-Plett Str. 32
34-109 Kassel, Germany

Invited paper

An accurate prediction of high-frequency circuit behaviour

Sadayuki Yoshitomi, Hideki Kimijima, Kenji Kojima, and Hideyuki Kokatsu

Abstract—An accurate way to predict the behaviour of an RF analogue circuit is presented. A lot of effort is required to eliminate the inaccuracies that may generate the deviation between simulation and measurement. Efficient use of computer-aided design and incorporation of as many physical effects as possible overcomes this problem. Improvement of transistor modelling is essential, but there are many other unsolved problems affecting the accuracy of RF analogue circuit modelling. In this paper, the way of selection of accurate transistor model and the extraction of parasitic elements from the physical layout, as well as implementation to the circuit simulation will be presented using two CMOS circuit examples: an amplifier and a voltage controlled oscillator (VCO). New simulation technique, electro-magnetic (EM)-co-simulation is introduced.

Keywords—electro-magnetic simulation, SPICE, circuit test structure, RF CMOS, EKV2.6-MOS model, spiral inductor, CMOS VCO.

1. Introduction

The goal of RF analogue circuit design with ultimate accuracy has been sought by both modelling and circuit engineers for a long time. The necessity to consider the electro-magnetic (EM) effects has been recognized because the electro-magnetic behaviour of the signal line and the passive elements can have considerable effect on the circuit performance.

The following phenomena become prominent: the self-inductance, skin effect and mutual electrical coupling between signal lines and the electro-magnetic loss of silicon substrate. These factors are as important for the design as for the accuracy of the transistor SPICE model. Moreover, the issue becomes even more critical as the operating frequency of RF circuits is increased.

Introduction of the above phenomena into the model is not easy because of their dependence on the type of the layout, location, size of the device, the number and the structural configuration of the transistor, etc. Thus, for the circuit design of chips intended for mass production, the electro-magnetic behaviour has still not been introduced until today.

In this paper, ways to overcome this problem for the semiconductor industry have been proposed. This is a new simulation technique, EM-co-simulation [1].

The main idea is to share the results of EM simulation with the circuit simulation. Each step, as well as the results of the verification will be explained in this paper.

All the test structures presented in this paper have been fabricated in TOSHIBA's 3.3 V, 0.40 μm rule SiGe BiCMOS technology. The maximum values of f_T and $f_{\max}$ of n-p-n transistors are 30 GHz and 50 GHz, respectively, and the maximum f_T of NMOS is about 20 GHz. Three metal layers have been implemented with the uppermost layer of 3 μm thickness dedicated for the fabrication of inductors.

This report consists of the following parts:

- selection of CMOS SPICE model,
- investigation of the applicability of EM simulation,
- verification of the circuit performance by using EM-co-simulation technique.

2. Selection of CMOS SPICE model

An accurate transistor model is the most essential matter for the real circuit design. Lately, surface potential (SP) based MOS models, such as EKV [3–5], HiSIM [6] and SP [7] have become well recognized.

Among these three models, only EKV Version 2.6 (EKV2.6) has already been implemented into several commercial-based simulators, therefore, a comparison in terms of DC and small signal output characteristics between EKV 2.6 and BSIM3 Version 3.2 (BSIM3V3.2) [8, 9] was made.

2.1. Device measurement and stability test

Before starting the discussion on the accuracy of the above two models, one should rather clear the measurement stability issue first to make the discussion trustworthy. Its solution is to analyze the robustness of the measurement data by using statistical approach. The detailed procedure will be explained in the following section.

The measurement of MOS transistors with three geometries: large ($L_g = 10$ μm, $W_g = 10$ μm), short ($L_g = 0.4$ μm, $W_g = 10$ μm), and narrow ($L_g = 10$ μm, $W_g = 0.6$ μm), has been performed by two persons for 4 chips belonging to one wafer. This 8 (= 4 times 2) sets of measurements have been repeated three times. A total of 24 measurement data for each size have been collected. Agilent's HP4156 with cascade probe station has been used as a measurement tool.

To evaluate the model's accuracy two quantities Dev_g_{ms} and Dev_n_{fact} have been introduced defined by the following formulae:

$$\text{Dev\_}g_{ms} = \sum_{large,short,narrow} \sum_{\frac{I_D}{I_{spec}}=0.1}^{\frac{I_D}{I_{spec}}=10} \frac{|\text{meas}(g_{ms}) - \text{sim}(g_{ms})|}{\text{sim}(g_{ms})}, \tag{1}$$

$$\text{Dev\_}n_{fact} = \sum_{large,short,narrow} \sum_{\frac{I_D}{I_{spec}}=0.1}^{\frac{I_D}{I_{spec}}=10} \frac{|\text{meas}(n) - \text{sim}(n)|}{\text{sim}(n)}. \tag{2}$$

The measurement and calculation of g_{ms} (normalized gate-to-source conductance) and n (slope factor) have been done using formulae (3)–(9), where I_D is the drain current, U_T is the thermal voltage (= kT/q), q is the electron charge, ε_{si} is the silicon permittivity, N_{sub} is the doping concentration in silicon, and V_{TO} is the threshold voltage.

Using (4), the universal function G_s expressed by (3) [5] can be obtained from the g_{ms}, that can easily be obtained from the simulation and measurement data. In this sense, G_s is a useful figure, because both simulated and measured behaviour of a MOSFET should follow this function:

$$G_S = \frac{1}{\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{I_D}{I_{spec}}}}, \tag{3}$$

$$g_{ms} \equiv -\left.\frac{\partial I_D}{\partial V_S}\right|_{V_G,V_D} = \frac{G_S \cdot I_D}{U_T}, \tag{4}$$

where:

$$I_{spec} = 2 \cdot n \cdot U_T^2 \cdot \mu \cdot C_{ox} \cdot \frac{W}{L}, \tag{5}$$

$$n \equiv \left[\frac{\partial V_P}{\partial V_G}\right]^{-1} = 1 + \frac{\text{GAMMA}}{2\sqrt{\Psi_o + V_P}}, \tag{6}$$

where:

$$V_P \cong \frac{V_G - V_{TO}}{n}, \tag{7}$$

$$\text{GAMMA} = \sqrt{2q\varepsilon_{\text{si}}N_{sub}}/C'_{ox}, \tag{8}$$

$$\Psi_0 = 2 \cdot U_T \cdot \ln\left(N_{sub}/n_i\right). \tag{9}$$

The Dev_g_{ms} and Dev_n_{fact} have been calculated and statistical analysis has been performed with Minitab [10] software using the obtained data. Figures 1 and 2 show the statistical distribution of Dev_g_{ms} and Dev_n_{fact} and

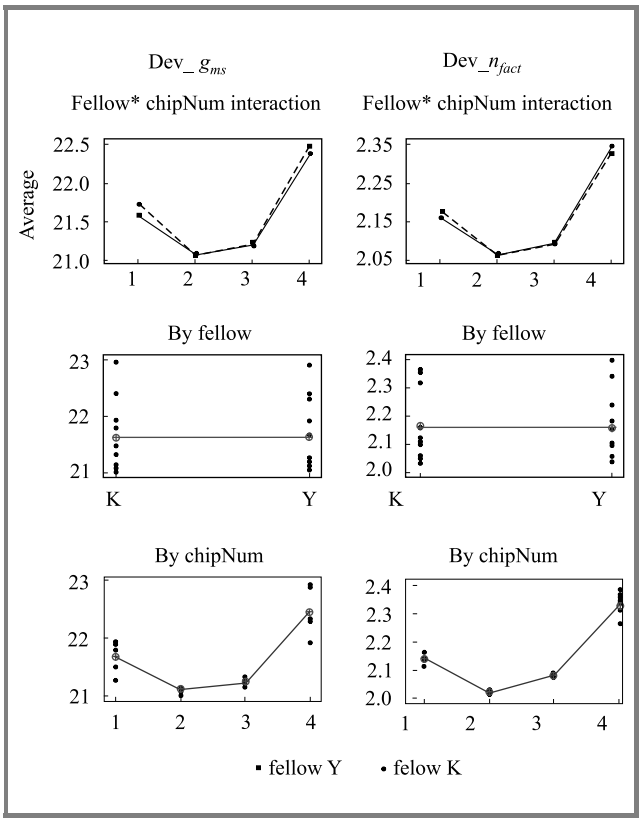


Fig. 1. Run chart of Dev_g_{ms} and Dev_n_{fact} in the case of EKV2.6 model.

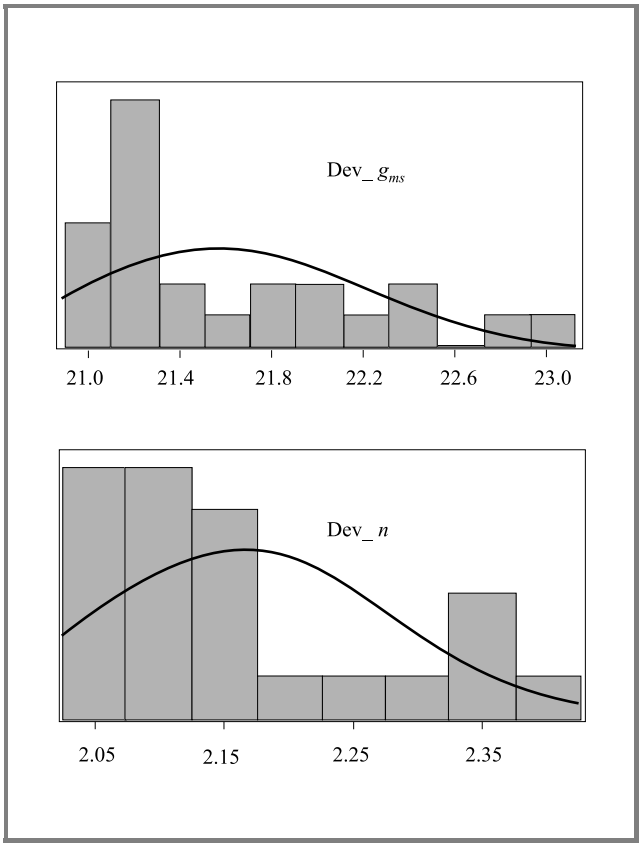


Fig. 2. Statistical distribution of Dev_g_{ms} and Dev_n_{fact} values in the case of EKV2.6 model.

Table 1 summarizes the analytical result. Gauge R&R in Table 1 is the measure of the stability of the measurement system, which is the total sum of % contribution of three factors: (1) repeatability, (2) reproductivity and (3) measurement operator. The theory requires that Gauge R&R should be less than 20% [10].

Table 1
Summary of the stability test of Dev_g_{ms} and Dev_n_{fact} in case of EKV2.6 model

Factor for robustness	Contribution [%]	
	Dev_g_{ms}	Dev_n_{fact}
Total Gauge R&R	5.49	13.12
– repeatability	5.49	13.12
– reproductivity	0	0
– fellow	0	0
Part-to-part	94.51	86.88
Total variation	100	100

The resultant Dev_n_{fact} and Dev_g_{ms} value in Table 1 was 5.49% and 13.2% respectively, which is less than 20%. It has been concluded that the way of measurement used in this study is stable and independent of the measurement operator so that the obtained measurement data is trustworthy.

2.2. Results and discussion of the benchmark test

The SPICE parameter extraction for EKV2.6 and BSIM3V3.2 models was done. The number of model parameters was 26 in the case of extraction based on EKV2.6 (including 22 original parameters and 4 parameters for extrinsic elements) and 81 in the case of BSIM3V3.2 (all 81 original parameters). Simulation data has been generated by Synopsys's HSPICE2002.2. Agilent's ICCAP has been used for verification and parameter extraction.

It should be noted that the parameters of both models include the temperature effects in the range from -40°C to 150°C. The number of the variations of geometrical test patterns used for the extraction is 19 (gate length ranges from 0.35 μm to 10 μm, and gate width from 0.6 μm to 10 μm).

After parameter extraction, simulation data was generated to calculate Dev_g_{ms} and Dev_n_{fact}, and the median value of 24 samples for each model was obtained using the Minitab software. Discussion on the modelling accuracy has been done based on this median value.

Figure 3 shows the comparison of G_s between EKV2.6 and BSIM3V3.2 for short and large devices. One should focus

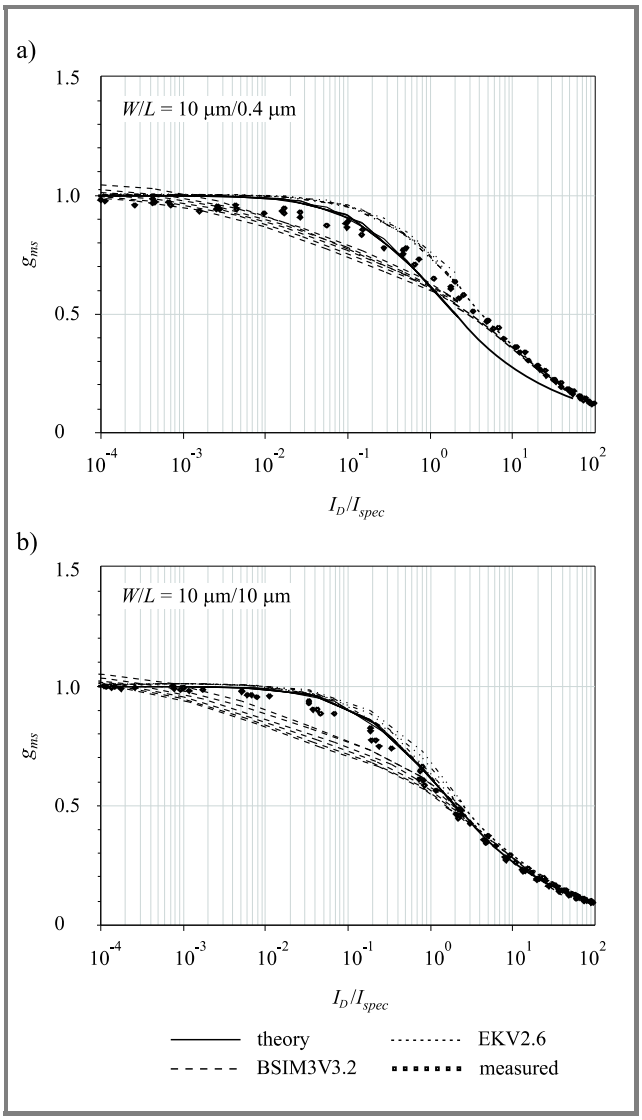


Fig. 3. The benchmark result of $G_s(= g_{ms}U_T/I_D)$ according to EKV2.6 and BSIM3V3.2 for: (a) short device ($I_{spec} = 3.0$ μA); (b) large device ($I_{spec} = 280$ nA).

on the G_s value in the range $0.1 < I_D/I_{spec} < 10$, which corresponds to the operation region of MOSFET between moderate and strong inversion, which is typically used for analogue circuits. This figure shows that both theoretical and measured G_s values match well. This means that the formula (3) is a valid expression for MOSFETs output characteristics.

As for the simulated results between two models, the EKV2.6 model is a better fit with G_s than BSIM3V3.2 in the whole range of I_D/I_{spec}. This fact indicates that EKV2.6 can describe the real behaviour of MOSFETs much better than BSIM3V3.2.

Comparison of the slope factor n has been performed and the results are given in Fig. 4. Strange behaviour has been observed in the case of BSIM3V3.2 results. The BSIM3V3.2 curve started decreasing for voltages below $V_{GB} = 0.7$ V and then increasing sharply for V_{GB} more

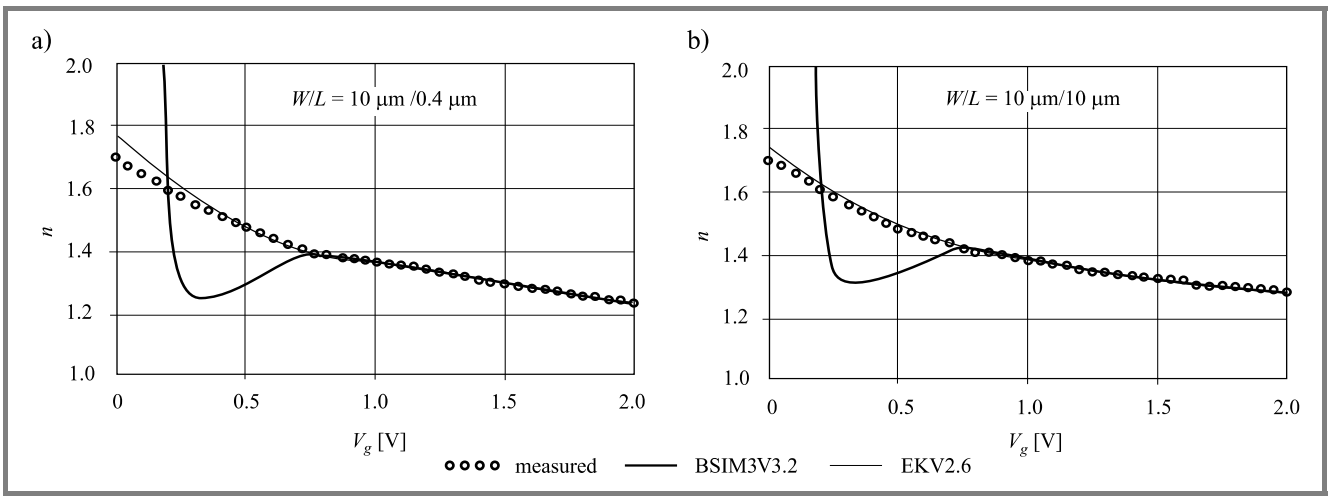


Fig. 4. The benchmark result of n factor according to EKV2.6 and BSIM3V3.2 for: (a) short device; (b) large device.

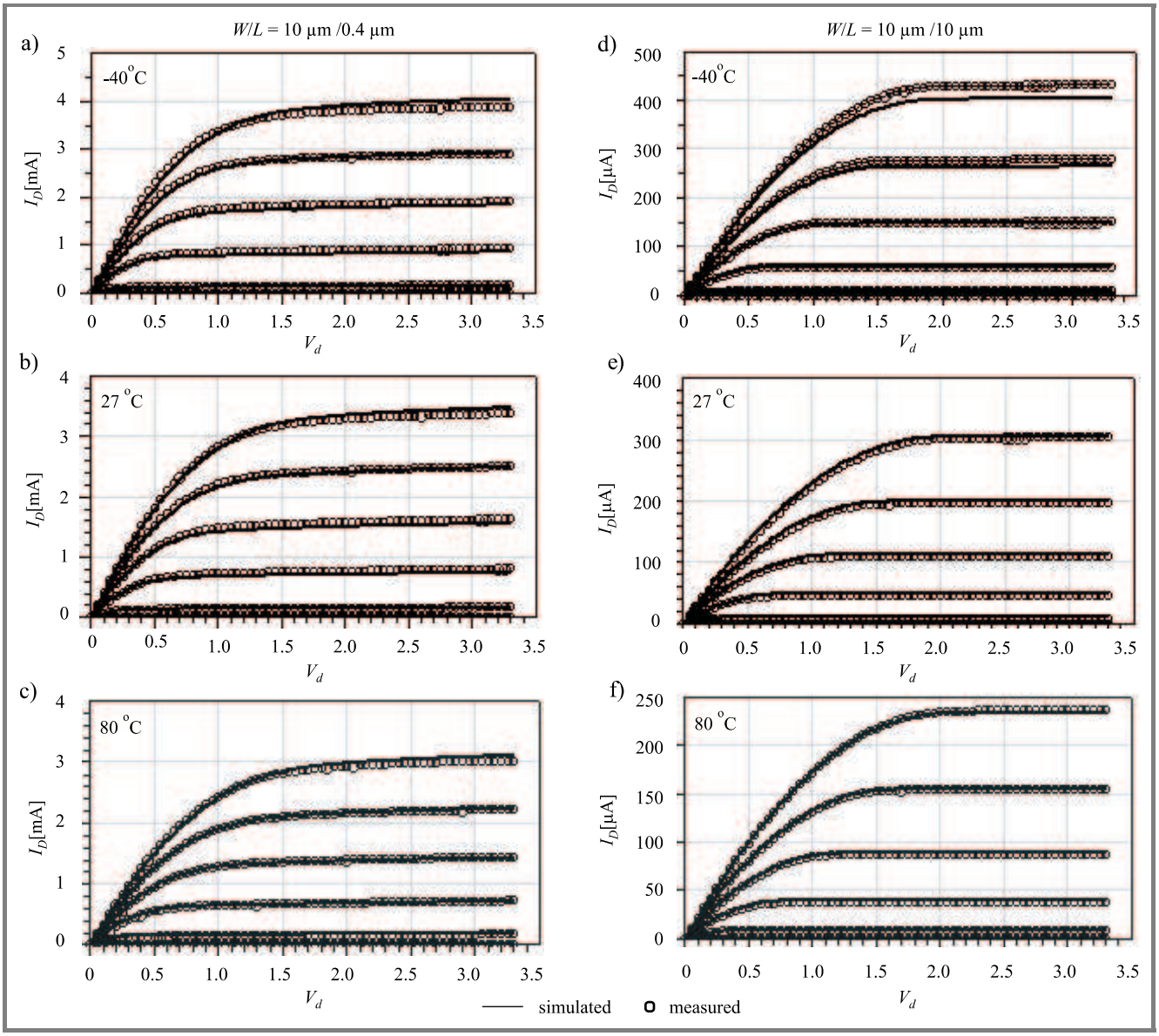


Fig. 5. Fitting result of EKV2.6 model of short (a, b, c) and large (d, e, f) device for three ambient temperatures.

closer to 0, while the EKV2.6 curve is in good agreement with the measured data.

The results are summarized in Table 2. Dev_g_{ms} for EKV2.6 was 21.0, which is approximately 1/5 of that of BSIM3V3.2 (99.0). The same holds for Dev_n_{fact} which was only 2.10 for EKV2.6, while it rose to 28.3 for BSIM3V3.2, which is 13.5 times higher. This confirms that EKV2.6 is better to use than BSIM3V3.2. These results coincide completely with those presented by Dr. Matthias Bucher [5].

Table 2

Comparison of Dev_g_{ms} and Dev_n_{fact} values calculated according to EKV2.6 and BSIM3V3.2

Contents	EKV2.6	BSIM3V3.2
Dev_g_{ms}	21.0	99.0
Dev_n_{fact}	2.10	28.3

Figure 5 shows the results of the verification of EKV2.6 models over three ambient temperatures (−40°C, 27°C and 80°C). It is amazing that this superb fit could be obtained using only 26 model parameters.

As a result of the above investigation, EKV2.6 was chosen for the entire study in this research.

2.3. Modification of the MOSFET SPICE model for RF application

In the previous section, the modelling accuracy of the intrinsic part of MOSFET has been discussed. However, the following extrinsic elements should be taken in account in the case of high frequency circuit operation:

1. Ohmic resistance of the gate material.
2. Resistance between source or drain and the substrate.
3. Coupling capacitances between gate, drain and source, respectively.

Figure 6 shows the EKV2.6 model modified for RF application [11–14]. Two added resistors R_G and R_B describe the gate and substrate resistance, respectively. Two capacitances (C_{GDFI} and C_{GSFI}) placed between the gate and drain, and gate and source are the sum of the overlap capacitances for the intrinsic part (C_{GD} and C_{GS}) and the extrinsic part describing the coupling between gate and drain, and gate and source electrodes (C_{GD_FI} and C_{GS_FI}) .

These values have been related to the physical layout by the following formulae [13, 14]. In the following section, extraction procedure of these values will be explained later.

- - - Device configuration - - -

L_g	[m]	: gate length
W_f	[m]	: finger length
Multi		: numbers of parallel devices
R_{PSH}	[Ω/square]	: sheet resistance of gate material
R_{GCT}	[Ω]	: gate contact resistance

- - - Gate resistance reduction factor - - -

$$FRG = \frac{1}{3} \quad \text{(one sided)} \tag{10}$$

$$FRG = \frac{1}{12} \quad \text{(double sided)} \tag{11}$$

- - - Formula for the extraction of R_G - - -

$$R_{GSH} = R_{PSH} \cdot FRG \tag{12}$$

$$R_G = \frac{R_{GCT} + R_{GSH} \cdot W_f / L_g}{\text{Multi}} \tag{13}$$

- - - Formula for the extraction of R_B- - -

R_{bref} : R_B value used for best fit

W_{fref} : finger length used to fit R_B

$$R_B = \frac{R_{bref} W_{fref}}{W_f \,\text{Multi}} \tag{14}$$

- - - G-D coupling capacitance - - -

$$C_{GDFI} = (C_{GD} + C_{GD\_FI})\, W_f \,\text{Multi} \tag{15}$$

- - - G-S coupling capacitance - - -

$$C_{GSFI} = (C_{GS} + C_{GS\_FI})\, W_f \,\text{Multi} \tag{16}$$

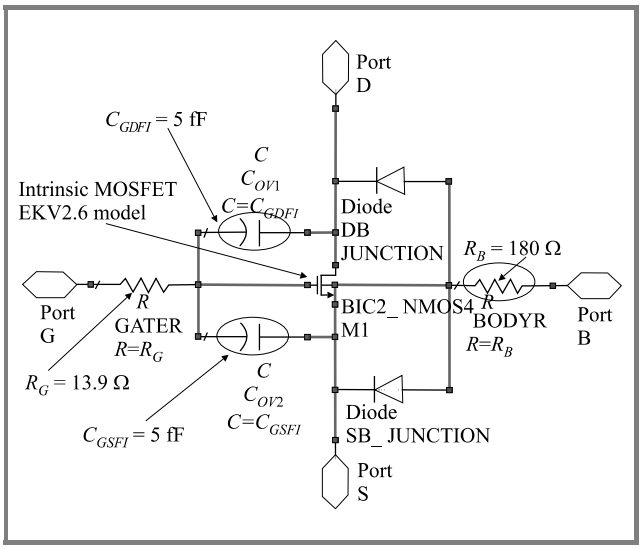


Fig. 6. Subcircuit-based EKV2.6 model with W_f (finger length) = 20 μm, L_g (gate length) = 0.4 μm, and Multi (numbers of fingers) = 5.

3. Investigation of the applicability of EM simulation

In this section, the results of the investigation of the applicability of electro-magnetic simulation will be explained using pad and inductor structures. Agilent's Momentum [1] has been used as a simulation tool.

3.1. Verification using pad structure

Figure 7 is the layout view of the pad structure used as the first case of the investigation. This is 1 port ground (G)-signal (S)-ground (G) configuration with 150 μm pitch.

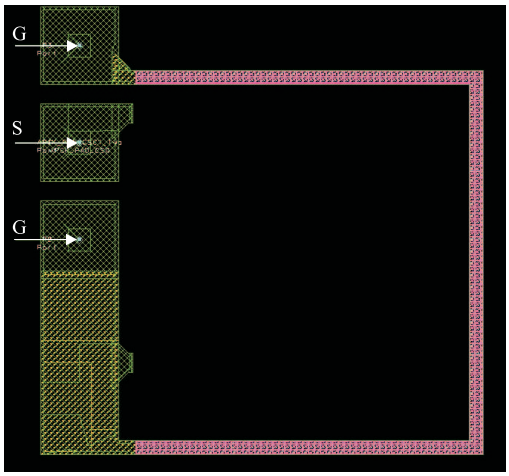


Fig. 7. Layout of the pad structure used for the verification of Momentum.

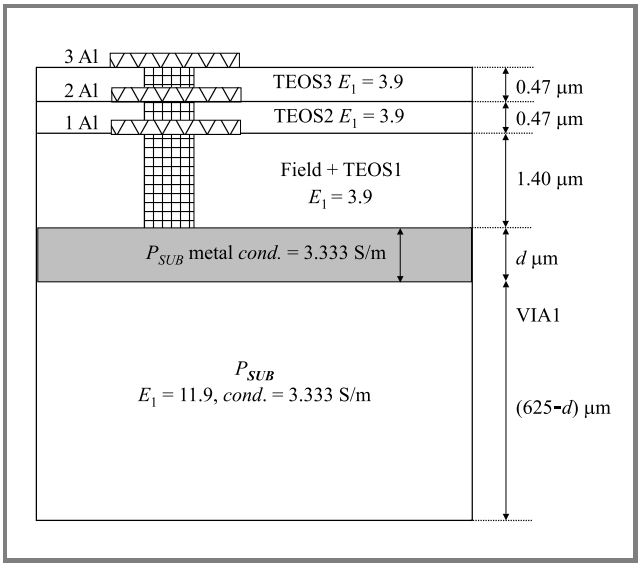


Fig. 8. Cross sectional information on TOSHIBA's 0.4 μm BiCMOS technology used for Momentum simulation.

Two G pads are connected to the silicon substrate through guard ring. Figure 8 illustrates the cross sectional information on the BiCMOS process used for the Momentum simulation. A thin conductive layer with variable thickness (d), which has the same resistivity as the bulk sub-

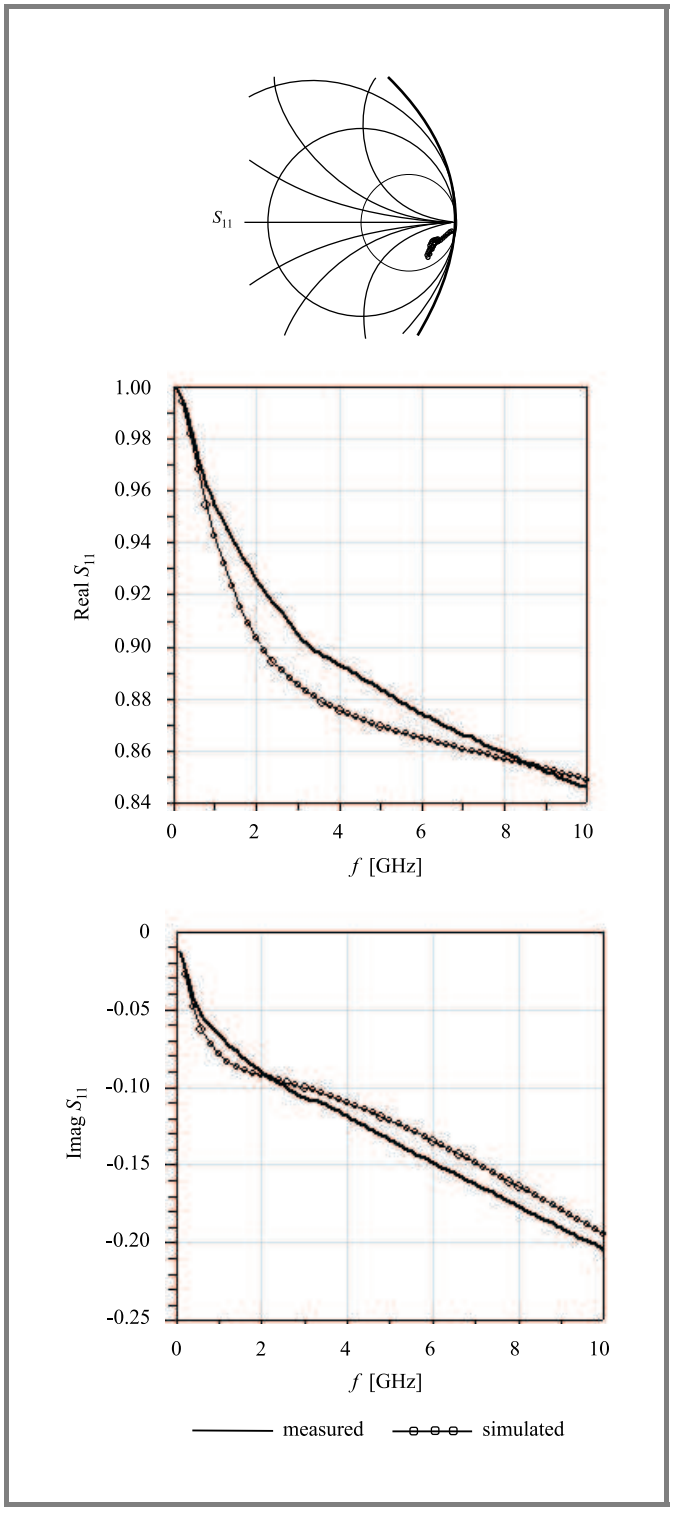


Fig. 9. Comparison of the S_{11}-parameter data obtained from Momentum simulation with measurement data of the structure shown in Fig. 8. Covered frequency range: from 100 MHz to 10 GHz.

strate, is placed on top of the substrate. Its role is to modify the amount of eddy current induced in the substrate. Because the actual value of d is unknown and typically process dependent, optimization has been used to find the best match with the measurement data.

The optimum d was obtained by fitting with S_{11} measurement data (frequency range: 0.1 GHz to 10 GHz) of the structure presented in Fig. 7. Figure 9 shows the results of S_{11} fitting with the case of $d = 60$ μm. The simulation RMS error stayed below 12.2% for the real part and 2.5% for the imaginary part, respectively.

3.2. Verification using inductor structure

The d value of 60 μm obtained previously has been verified using inductor shown in Fig. 10. The pad configuration is the same as in the previous case, and the distance between the inductor and guard-ring was 20 μm.

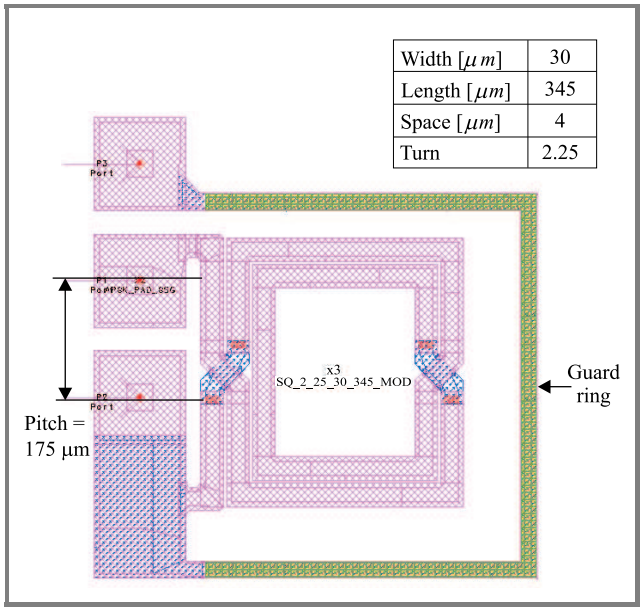


Fig. 10. Layout of the 2.25 turns of inductor used for the verification of Momentum.

Figure 11 shows a comparison of S_{11} and quality factor Q between simulation and measurement. Very good match has been obtained in this case, too. The S_{11} simulation RMS error was kept below 2.0% for the real part and 2.2% for the imaginary part, respectively.

The applicability of EM simulation has been verified using two structures, and good match between measurement and simulation has been obtained. It was achieved by adding the conductive layer with 60 μm thickness on top of the bulk silicon substrate. It is believed that the same approach may be applied to other silicon technologies as well.

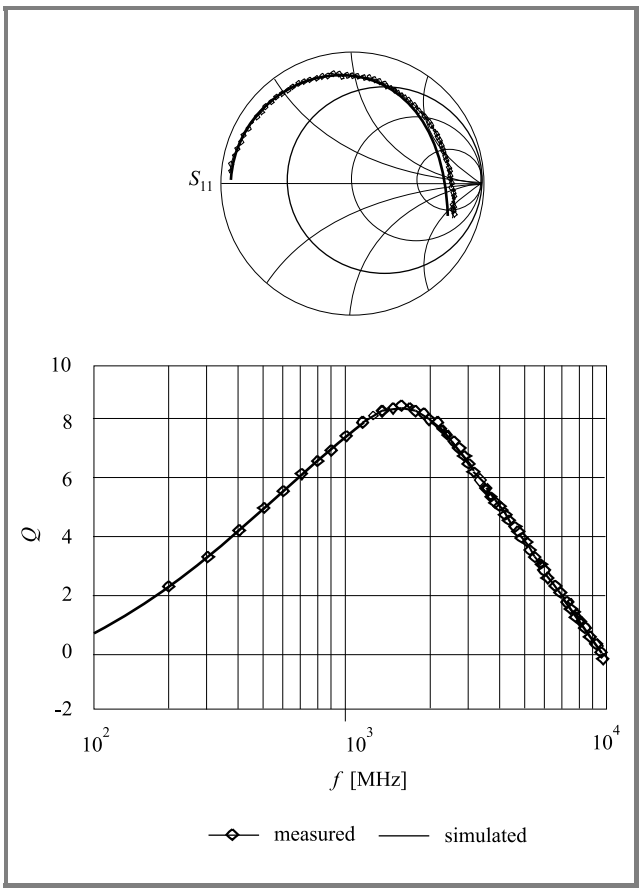


Fig. 11. Comparison of the S_{11}-parameter data obtained from Momentum simulation with measurement data of the structure shown in Fig. 8. Covered frequency range: from 100 MHz to 10 GHz.

4. Verification of the circuit performance using EM-co-simulation technique adapted for two design examples

The extended EM simulation technique applied for circuit design will be explained in this section. This is called electro-magnetic co-simulation. This technique has been applied to two cases, 1) CMOS differential amplifier and 2) 1.9 GHz CMOS VCO, and the results will be explained.

4.1. Case 1: CMOS differential amplifier

4.1.1. Description of the circuit

The layout and schematic of the CMOS differential amplifier used in this study are shown in Figs. 12 and 13, respectively. This is a single-ended configuration. Five pads,

located on the upper side of the circuit are for the DC power supply, and two sets of G-S-G pads on both sides are provided for the input/output. This circuit, which is fully dedicated for the verification of model parameters, has the following features.

1. **Unified device configuration to suppress the process fluctuation effect.**

   All transistors in the circuit are NMOSFETs with the following geometry and configuration of R_G:

   $$L_g \text{ (gate length)} = 0.4\ \mu\text{m}, \tag{17}$$

   $$W_f \text{ (finger length)} = 20\ \mu\text{m}, \tag{18}$$

   $$\text{Multi (numbers of parallel devices)} = 5, \tag{19}$$

   $$FRG = \tfrac{1}{3} \text{ (single-sided gate contact)}, \tag{20}$$

   $$R_{PSH} \text{ (sheet resistance of gate material)} = 3.5\ \Omega/\text{square}, \tag{21}$$

   $$R_{GCT} \text{ (gate contact resistance)} = 3\ \Omega, \tag{22}$$

   $$R_{GSH} = R_{PSH} \cdot FRG = 1.67\ \Omega/\text{square}, \tag{23}$$

   $$R_G = \frac{R_{GCT} + R_{GSH} \cdot W_f / L_g}{\text{Multi}} = 12.3\ \Omega. \tag{24}$$

   All resistors are used with a series or parallel connection of the 300 Ω resistor having 4 μm width. This is aimed at reducing the effect of process fluctuations, which strongly depends on the resistor width.

2. **Only 1st and 2nd metal layers have been used**.

   This is for practical reasons. A shorter turn-around time was expected.

3. **Special pattern has been fabricated to calibrate out the parasitic impedances of the pads**.

   Open calibration pattern was prepared to detect the intrinsic behaviour of the circuit.

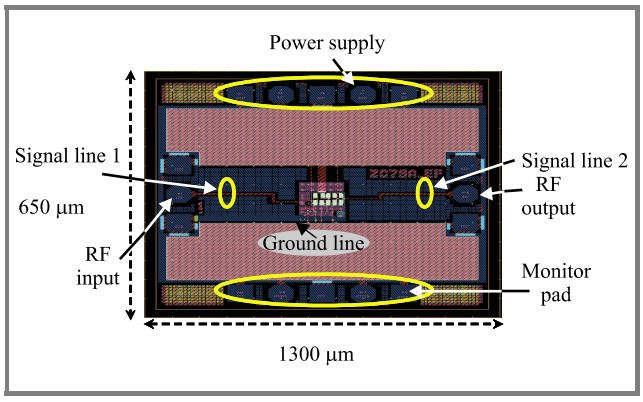


Fig. 12. Layout of CMOS differential amplifier.

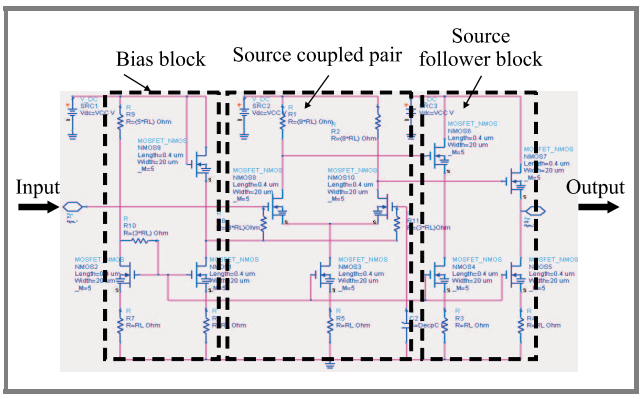


Fig. 13. Schematics of CMOS differential amplifier shown in Fig. 12.

4.1.2. Conditions of the measurement and simulation

The two port S-parameter measurement was done in the following conditions:

- frequency range: 0.1 GHz to 6 GHz with 0.1 GHz step,
- supply voltage: 2.8 V to 3.4 V with 0.2 V step,
- instrument: Agilent's HP8510B network analyzer and HP4142 DC supply.

4.1.3. Characterization of signal lines using Momentum

To predict the high frequency behaviour more accurately, the effect of series parasitic impedances of the following lines should be taken into account, because such effects cannot be removed by open de-embedding procedure only:

1) signal line between RF input pad and the core circuit,
2) signal line between RF output pad and the core circuit,
3) ground line between ground and the core circuit.

Momentum simulation has been used to characterize these three lines. The layout data of the above signal lines were picked up and used for Momentum simulation. The resulting S-parameters have been converted into the equivalent circuit as shown in Fig. 14 and the parameters of the circuit equivalent model were calculated by the formulae listed in Table 3.

In Fig. 14 L_s and R_s is the series inductance and resistance. $C_{ox1(2)}$ is the coupling capacitance between metal lines and silicon substrate, which is usually estimated by the insulator capacitance. The parallel network of $R_{sub1(2)}$ and $C_{sub(2)}$ denote the signal loss regarding eddy current generated

in the substrate. The formulae in Table 4 exhibit the case for symmetrical topology. Nevertheless, this can be expanded to the asymmetrical case by introducing y_{22} instead of y_{11}. The advantage of this methodology is that all the model parameters can be extracted directly from the measurement data without any use of optimization.

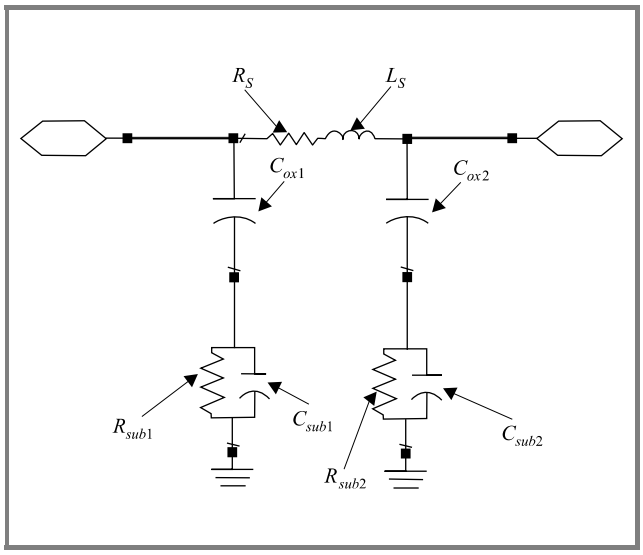


Fig. 14. Schematic description of the equivalent circuit of the signal line 1, signal line 2 and ground line shown in Fig. 12.

Table 3

Formulae for the extraction of the equivalent circuit of Fig. 14

Element	Extraction formula
L_s	$\dfrac{\mathrm{Im}\left(\dfrac{1}{y_{21}}\right)}{2\pi f}$
R_s	$\mathrm{Re}\left(\dfrac{1}{y_{21}}\right)$
$Z_{sub1(2)}$	$\dfrac{1}{y_{11(22)}+y_{21}}-\dfrac{1}{\mathrm{j}2\pi C_{ox}}$
$C_{ox1(2)}$	$\dfrac{-1}{2\pi f\,\mathrm{Im}\left[\dfrac{1}{y_{11(22)}+y_{21}}\right]}$ at low frequency end
$C_{sub1(2)}$	$\dfrac{\mathrm{Im}\left[\dfrac{1}{Z_{sub1(2)}}\right]}{2\pi f}$
$R_{sub1(2)}$	$\dfrac{1}{\mathrm{Re}\left[\dfrac{1}{Z_{sub1(2)}}\right]}$

The calculation methodology in Table 4 starts from the generation of π-structured network consisting of Y_1, Y_2 and Y_3.

Table 4

Equivalent-circuit elements of the three lines shown in Fig. 12 with the symmetrical topology assumed

Line element	L_s [nH]	R_s [Ω]	$C_{ox1(2)}$ [fF]	$C_{sub1(2)}$ [fF]	$R_{sub1(2)}$ [Ω]
Input	0.51	2.68	31.9	13.4	892
Output	0.51	2.79	54.2	13.2	1541
Ground	0.39	0.60	150.2	23.7	720

They are calculated based on the two-port y-parameters (y_{11}, y_{12}, y_{21}, y_{22}) according to the following relationship:

$$Y_1 = y_{11} + y_{12} \quad \text{or} \quad y_{11} + y_{21}, \tag{25}$$

$$Y_2 = y_{22} + y_{12} \quad \text{or} \quad y_{22} + y_{21}, \tag{26}$$

$$Y_3 = -y_{12} = -y_{21}. \tag{27}$$

Complete expressions for Y_1 and Y_2 can be written as the series impedance connection for $C_{ox1(2)}$ and $Z_{sub1(2)}$. Nevertheless, one assumption makes the extraction of $C_{ox1(2)}$, $C_{sub1(2)}$ and $R_{sub1(2)}$ simple. At the low frequency end, where the effect of eddy current is negligible, $Z_{sub1(2)}$ can be simplified to be a solo connection of $R_{sub1(2)}$. In this case, both Y_1 and Y_2 can be approximated in the following way:

$$Y_{1(2)} = \frac{1}{\mathrm{j}\omega C_{ox1(2)}} + \frac{1}{\dfrac{1}{R_{sub1(2)}} + \mathrm{j}\omega C_{sub1(2)}}$$

$$\cong \frac{1}{\mathrm{j}\omega C_{ox1(2)}} + \frac{1}{\dfrac{1}{R_{sub1(2)}}} = \frac{1}{\mathrm{j}\omega C_{ox1(2)}} + R_{sub1(2)} \tag{28}$$

Thus, $C_{ox1(2)}$ can be calculated from the imaginary part of Eq. (28). Care should be taken to obtain $C_{ox1(2)}$ value. It should be picked up from the data taken at a frequency as low as possible, where the introduced assumption is valid.

Then, $Z_{sub1(2)}$ can be defined by the remainder after subtracting the impedance of $C_{ox1(2)}$ from $Y_{1(2)}$:

$$Z_{sub1(2)} = \frac{1}{\dfrac{1}{R_{sub1(2)}} + \mathrm{j}\omega C_{sub1(2)}}$$

$$= Y_{1(2)} - \frac{1}{\mathrm{j}\omega C_{ox1(2)}}. \tag{29}$$

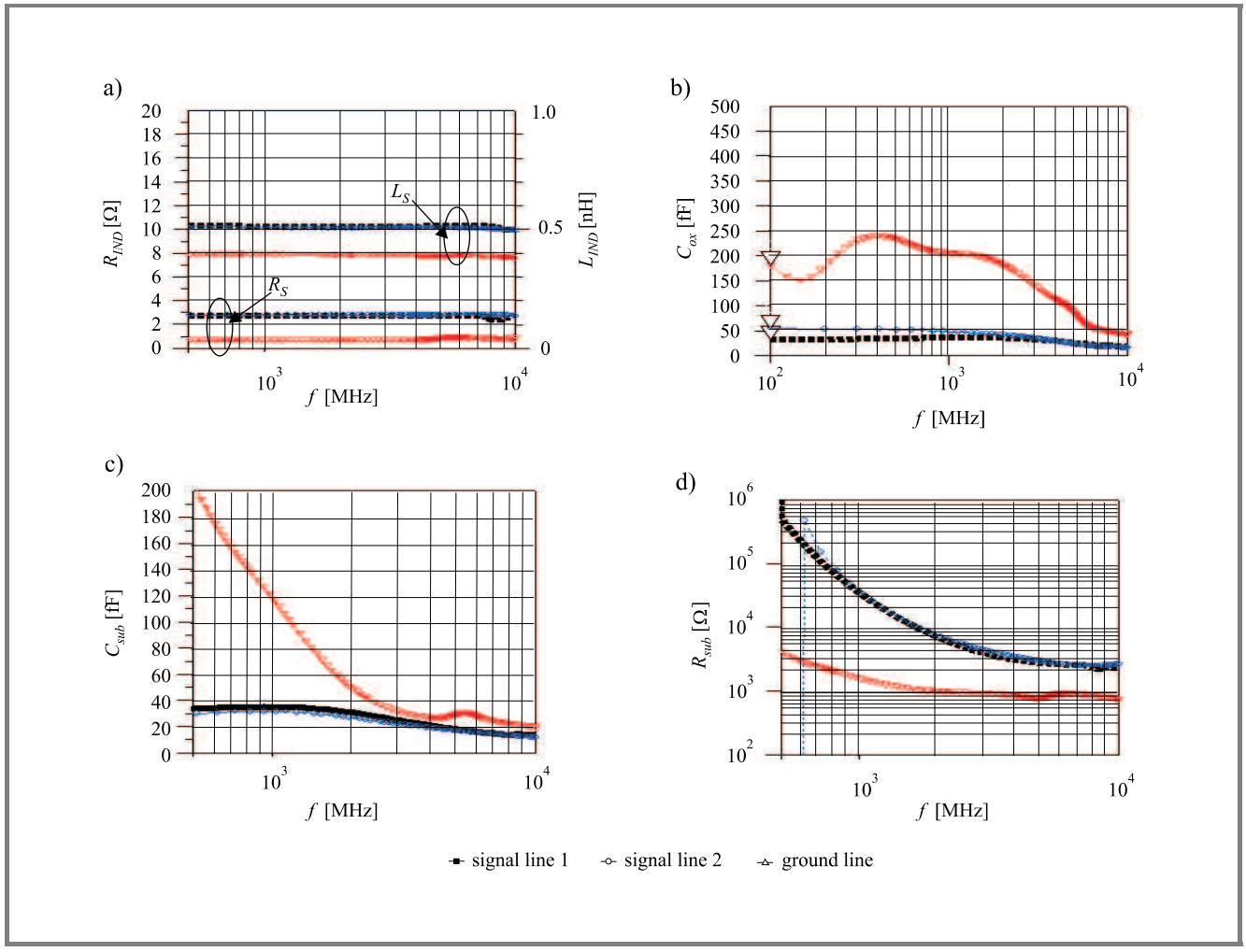


Fig. 15. Simulated frequency behaviour of L_S and R_S (a), C_{ox} (b), C_{sub} (c) and R_{sub} (d) of the three signal lines. Covered frequency range: from 100 MHz to 10 GHz.

Both $R_{sub1(2)}$ and $C_{sub1(2)}$ can be defined by the real and imaginary part of $Z_{sub1(2)}$, respectively. Figure 15 displays model parameters of the three signal lines, calculated by the formulae given in Table 3 (symmetrical topology ($Y_1 = Y_2$) is applied). Table 4 lists the obtained values.

4.1.4. Implementation of Momentum for the extrinsic parameter extraction of MOSFET models

The estimation of R_G, R_B, C_{GD_FI}, and C_{GS_FI} shown in Fig. 6 is very complex because they can hardly be recognized by the measurement data of the real device. Nevertheless, their accurate values should be extracted because they have strong influence on the high frequency characteristics.

The R_G value was extracted by the Momentum simulation using MOSFET layout data (configured by 5 fingers of 20 μm length with single-sided contact) of the gate electrode material as shown in Fig. 16. The two port y-parameters have been calculated between input port and 5 opposite ends. The R_G has been extracted by the real part of the resultant y_{21}:

$$R_G = \mathrm{Re}\left(-\frac{1}{y_{21}}\right). \tag{30}$$

Frequency dependency of R_G is plotted in Fig. 17. Extracted R_G value is approximately 13 Ω and is almost frequency independent. By comparing the estimated value as calculated in Eq. (24), it can be concluded that Eq. (13), used to calculate the initial value, can predict the R_G value with acceptable accuracy.

The determination of both C_{GD_FI} and C_{GS_FI} has been done in the same manner. The layout data shown in Fig. 18 was used for Momentum simulation. Both C_{GD_FI} and C_{GS_FI} have been calculated using the imaginary part of resultant y_{21} value:

$$C_{GS\_FI}, C_{GD\_FI} = \frac{1}{2\pi f\,\mathrm{Im}\left(\frac{1}{y_{21}}\right)}. \tag{31}$$

The result is shown in Fig. 19. Both capacitances are almost independent of the frequency and amounted to about 5 fF each.

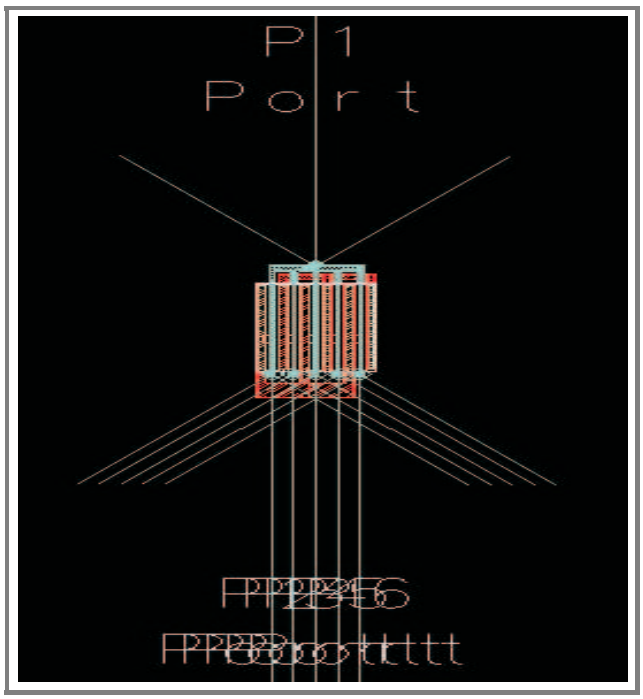


Fig. 16. Momentum setup for the extraction of R_G. Simulation target is the gate electrode material with 5 fingers of 20 μm length.

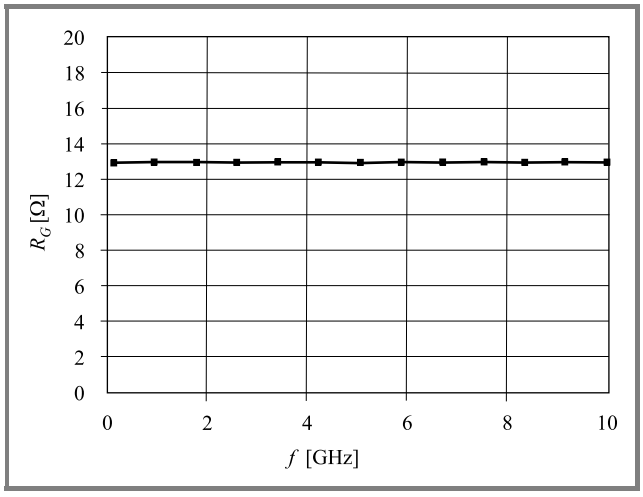


Fig. 17. Resultant R_G obtained using the setup shown in Fig. 16. Covered frequency range: from 100 MHz to 10 GHz.

The extraction of R_B was done as the last step. R_B has a strong effect on the S_{22} value of the circuit. This is because the output impedance of MOSFET, which forms the source-follower network, has a strong influence on the circuit's S_{22}. Mathematical optimization has been used to extract R_B. The effect of R_B on the circuit's S-parameters is displayed in Fig. 20. At R_B value of 180 Ω, best fit with the measurement was obtained. All parameters extracted in this section are summarized in Fig. 6.

With the lumped model of the signal lines and modified EKV2.6 model, the final simulation was performed.

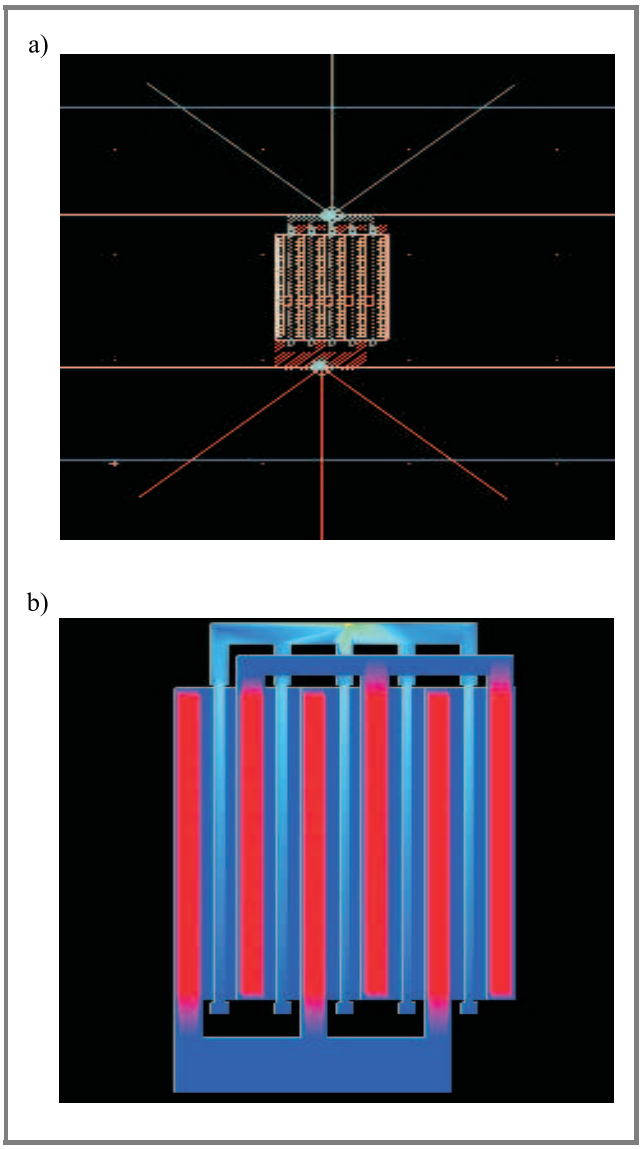


Fig. 18. (a) Setup for the extraction of C_{GS_FI} and C_{GD_FI}; (b) resultant current flow obtained from Momentum simulation. Simulation target is the gate material with the configuration as shown in Fig. 6.

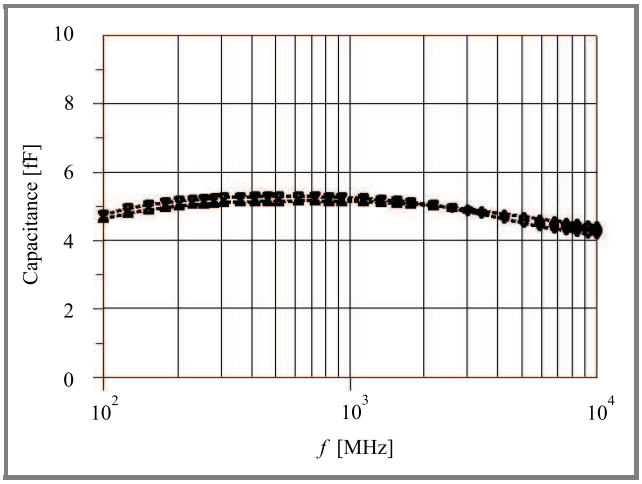


Fig. 19. Extracted C_{GD_FI} and C_{GS_FI}. Covered frequency range: from 100 MHz to 10 GHz.

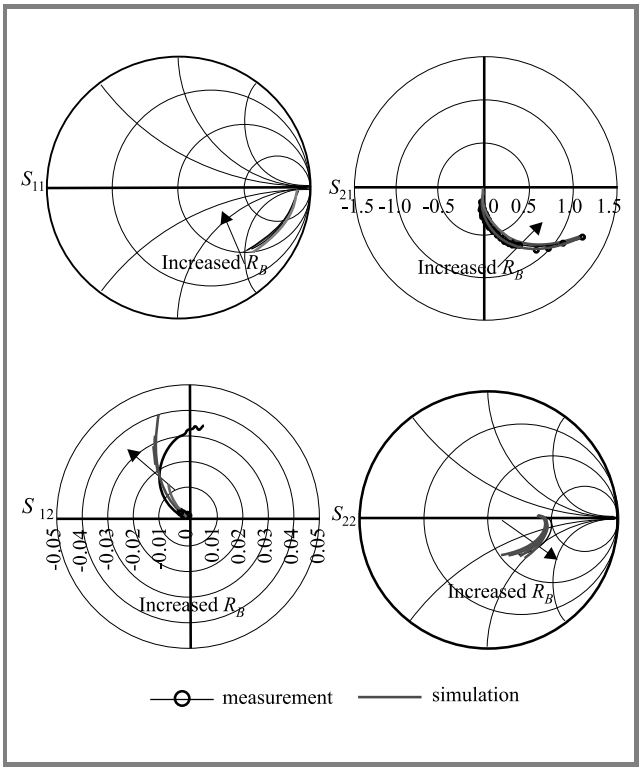


Fig. 20. Effect of R_B as shown in Fig. 6 on the *S*-parameters of the differential amplifier. Covered frequency range: from 100 MHz to 6 GHz.

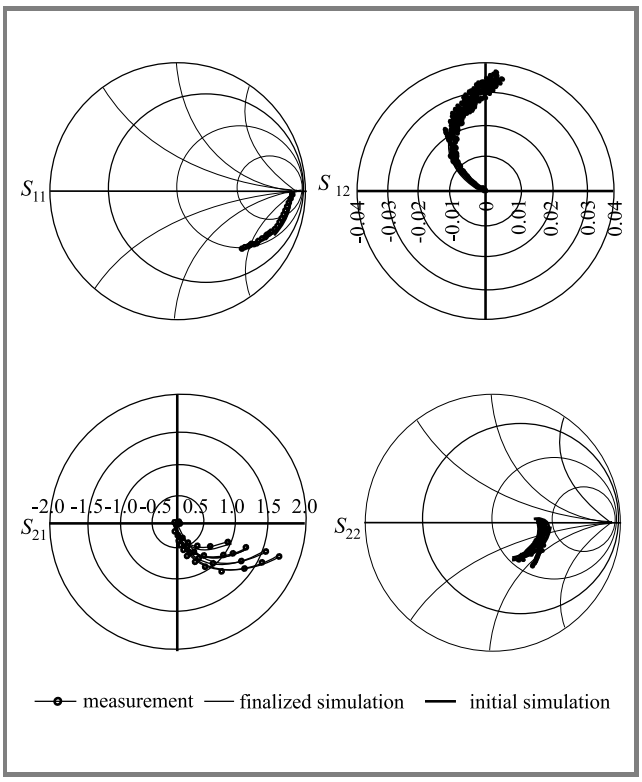


Fig. 21. Final simulation results with subcircuit-based RF-MOSFET (as shown in Fig. 6) and lumped signal line models (as shown in Fig. 13). Covered frequency range: from 100 MHz to 6 GHz.

The results are shown in Fig. 21. As seen, the simulated and measured *S*-parameter values are in very good agreement. Thus, the simulation accuracy of an RF circuit can be enhanced by the use of a proper device model and EM simulation.

4.1.5. Verification of amplifier behaviour using EM-co-simulation technique

As an extended example, EM-co-simulation technique will be presented in this section. The flow chart to execute EM-co-simulation is illustrated in Fig. 22. A layout-component, that is a symbolized layout block, can be used together with SPICE compact models as shown in Fig. 23. Agilent's ADS-2003C with Verilog-A code of EKV2.6 model was used for this simulation.

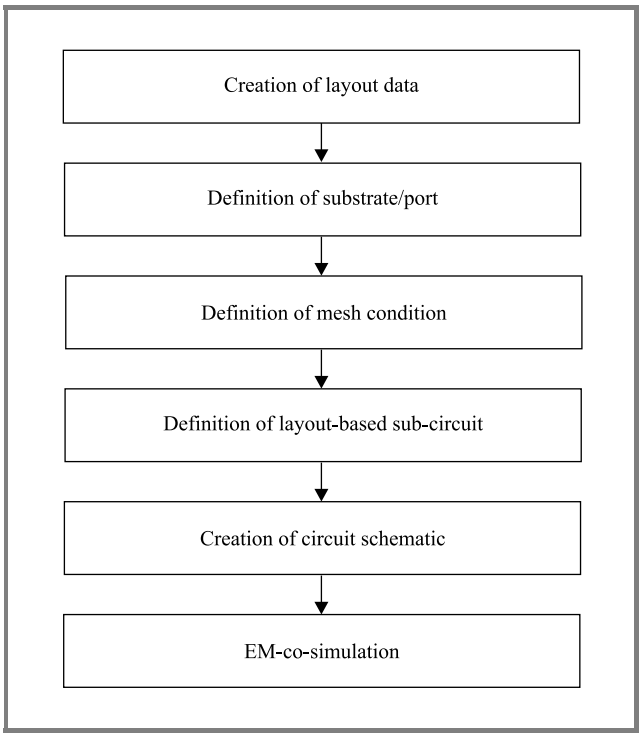


Fig. 22. Simulation flow of electro-magnetic co-simulation.

Figure 23 illustrates the view of the simulation schematic. The core part of the circuit is composed of spice components (transistors, resistors and capacitors), and other external components such as signal lines, pads are contained inside the layout-components.

At the first step of the simulation, EM simulation is invoked to analyze the *S*-parameters of the layout-component. After its completion, circuit simulation is followed by the use of the resultant *S*-parameters and compact model.

Now that both accurate transistor model and EM simulation methodology is available, simulation results shows good agreement with the measurement data as shown in Fig. 24.

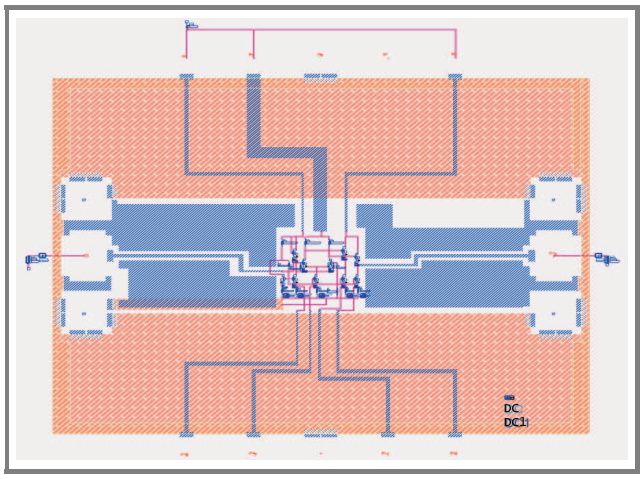


Fig. 23. The EM-co-simulation setup for the CMOS differential amplifier. Its schematic is shown in Fig. 13.

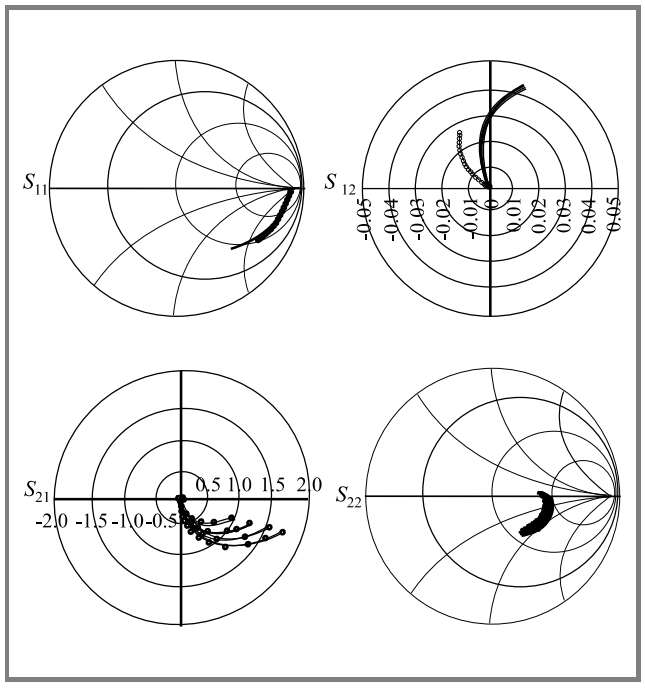


Fig. 24. Comparison between EM-co-simulation and measurement data of CMOS amplifier with the schematic shown in Fig. 13. Covered frequency range: from 100 MHz to 6 GHz.

The duration of the simulation of this example was 30 minutes on Solaris2.6 based EWS. The *S*-parameter values obtained from the Momentum simulation can be reused so that the duration of the second simulation did not take more than 30 seconds.

4.2. Case 2: 1.9 GHz CMOS VCO

4.2.1. Description of the circuit

The second adaptation example of the EM-co-simulation technique is the 1.9 GHz CMOS VCO design. Designing VCO is very difficult because many factors, such as the device model accuracy, parasitic elements and electromagnetic effect should be taken into account carefully in the course of the circuit design. The most difficult and critical issue is the characterization of tank circuit.

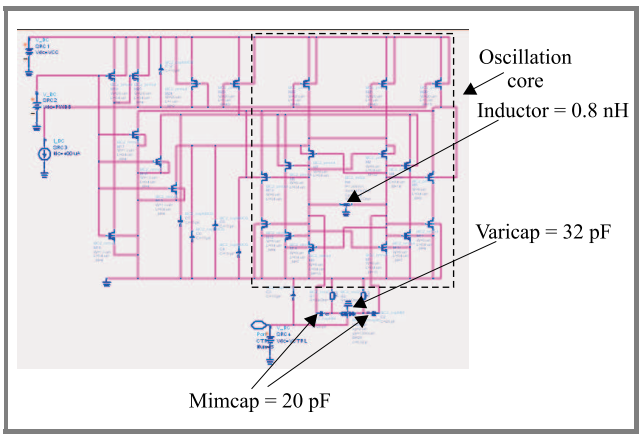


Fig. 25. Schematic of CMOS VCO for DCS1800 application.

The circuit schematic is displayed in Fig. 25. Six pairs of CMOS transistors with different numbers of fingers (unit size of 8 μm/0.4 μm) and a tank circuit (consisting of an inductor, a capacitor and a variable capacitor) constitute the main oscillator block.
The 2.75 turns of inductor used in the tank circuit have line and spacing of 40 μm/4 μm, and self inductance estimated from the Greenhouse formula [15] is 0.8 nH.

4.2.2. VCO measurement and simulation conditions

The frequency tuning characteristics and oscillation frequency have been measured with the supply voltage of 2.9 V and the control voltage (v_{ctrl}) swept from 0 V to 2.5 V with the step of 0.5 V.

4.2.3. Verification of VCO tuning behaviour using EM-co-simulation technique

The following two circuit simulations have been done for comparison:

1) circuit simulation using the schematic of Fig. 25 with the initial estimation (0.8 nH) of the inductance,

2) EM-co-simulation.

The schematic of (2) is displayed in Fig. 26. Simulation was using 7 tones' harmonic-balance simulation and Momentum simulation. The layout-component consisting of the tank circuit (inductor, leading line of varicap) and other signal lines between transistors and guard-ring was analyzed by the Momentum simulation.

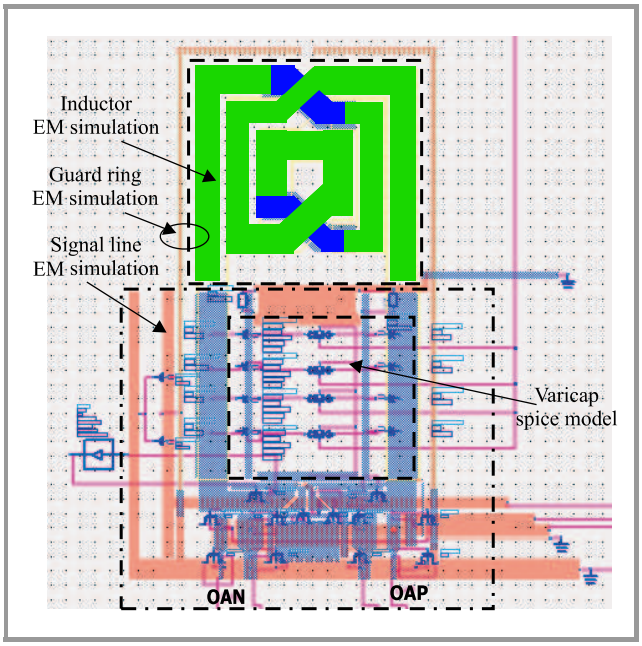


Fig. 26. Simulation setup of EM-co-simulation for CMOS VCO tuning characteristics.

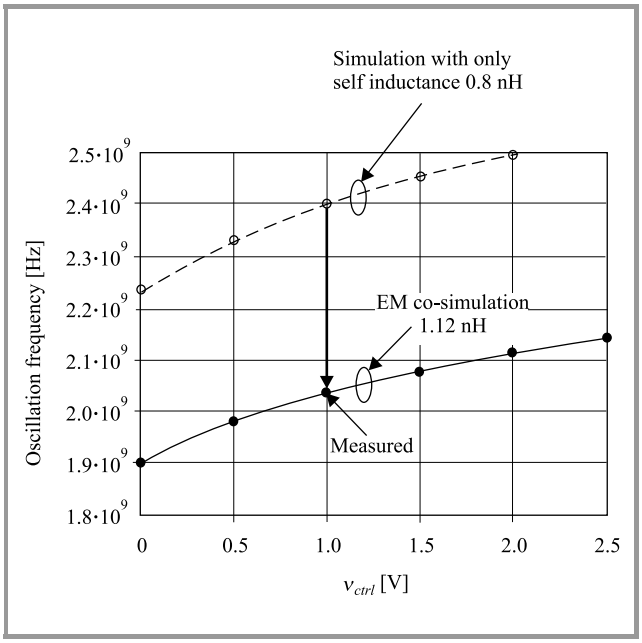


Fig. 27. Comparison of simulation results between EM-co-simulation and measurement of tuning characteristics of CMOS VCO.

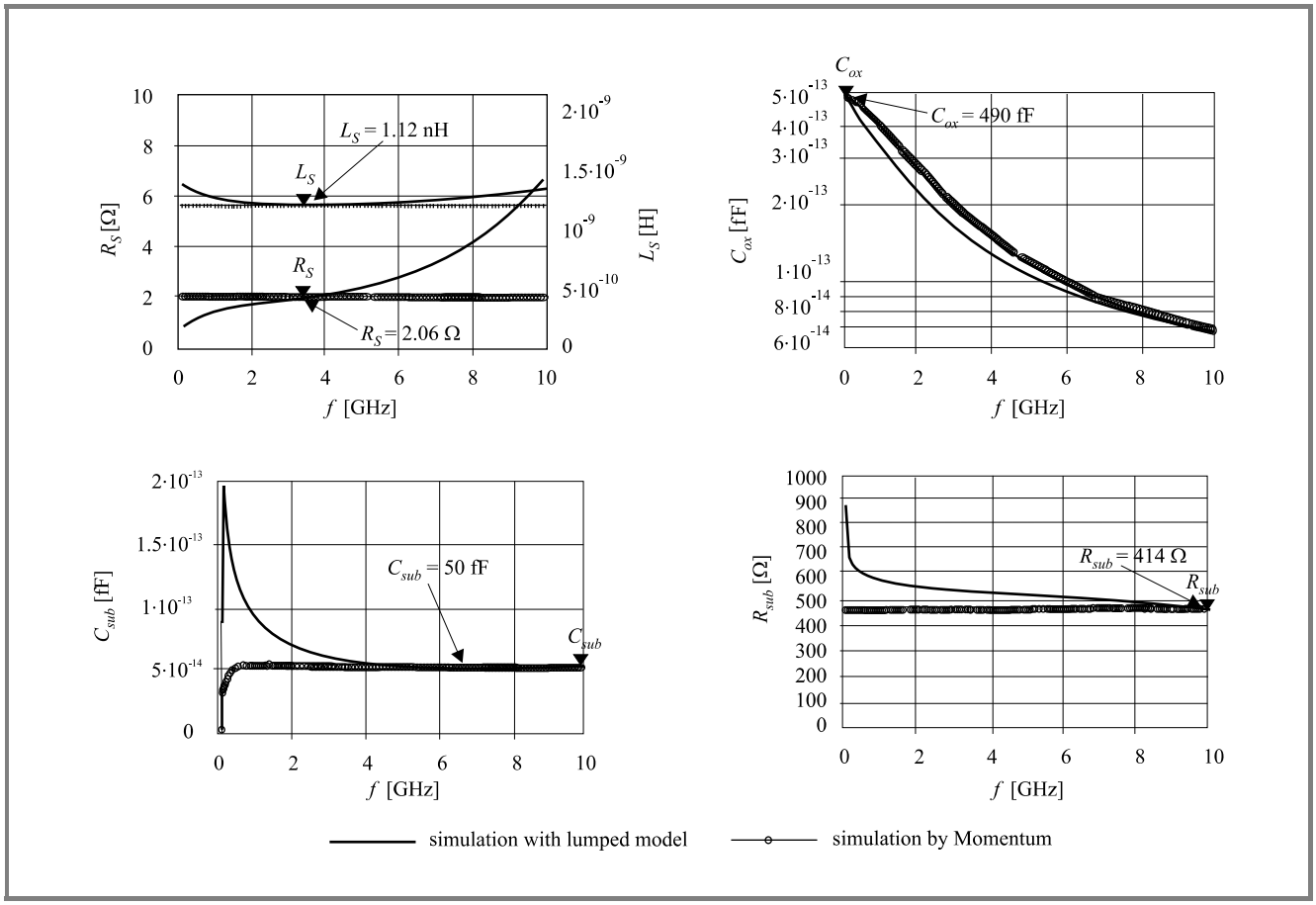


Fig. 28. Comparison of simulation characteristics of the inductor in the tank circuit between lumped model and Momentum. Model parameter values are displayed in the figure.

Simulation results for the two cases are compared in Fig. 27. The frequency deviation in the case (1) ranged between 330 MHz and 400 MHz from $v_{ctrl} = 0$ V to $v_{ctrl} = 2.0$ V, which approximately amounted to the frequency error of 16%–18%. In the case (2), the deviation has been drastically decreased to only 3 MHz difference from the measurement in the whole v_{ctrl} range. Thus, it can be concluded that EM-co-simulation is accurate and effective for the design of VCO circuits.

4.2.4. Detailed analysis of the tank circuit

The reason for the obtained accuracy in the case (2) seems to be the increase of the inductance in the tank circuit. This is confirmed by the Momentum analysis of the tank circuit layout. The values of the equivalent circuit as shown in Fig. 14 are $L_s = 1.12$ nH, $R_s = 2.06$ Ω, $C_{ox1(2)} = 490$ fF, $C_{sub1(2)} = 50$ fF and $R_{sub1(2)} = 414$ Ω. Figure 28 shows the comparison of simulation data of equivalent circuit with Momentum output.

The resultant L_s was 1.12 nH (a visible increase from the initial estimation of 0.8 nH), which means that the oscillation frequency was decreased by 15%. This increased inductance is ascribed to the parasitic inductances of the inductor's extension part, mutual coupling with the surrounding guard ring, and other electrical coupling.

Based on the above results, it can be concluded that the EM-co-simulation can incorporate the parasitic effects that cannot easily be estimated from the layout, and provides the circuit designer with the accurate prediction of high frequency circuit behaviour.

5. Conclusions

In this study, it has been proved that the proper use of good CMOS SPICE model (EKV2.6) and electro-magnetic simulation can yield a good match between measured and simulated data. Through the investigation of the applicability of EM simulation, using several test structures, it has been concluded that EM simulation is very useful for the estimation of inductor characteristics, layout parasitic components in the circuit and the extrinsic part of the transistor. EM-co-simulation is introduced as a new simulation technique. This can be used as a "post-layout tool" because of its acceptable accuracy.

To summarize, it can be concluded again that the use of both: accurate transistor model and electro-magnetic simulation is the shortest route to predict the behaviour of RF circuits in the most accurate way.

Acknowledgements

The Authors would like to express their gratitude to Dr. Matthias Bucher the Department of Electronics and Computer Engineering at Technical University of Crete for technical comments on the extraction of EKV2.6 model parameters and to Dr. Władysław Grabiński at Freescale Semiconductor Geneva Modeling and Characterization Lab for heartfelt support.

References

[1] "ADS 2003C reference manual", Agilent Technologies, 2003.

[2] "HSPICE2002.2 model manual", Synopsis, 2002.

[3] C. Enz, F. Krummenacher, and E. Vittoz, "An analytical MOS transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications", *J. Anal. Integr. Circ. Sig. Proc.*, no. 8, pp. 83–114, 1995.

[4] M. Bucher, C. Lallement, C. Enz, F. Théodoloz, and F. Krummenacher, "The EPFL-EKV MOSFET model equations for simulation, Version 2.6", Tech. Rep., Electronics Laboratory, Swiss Federal Institute of Technology Lausanne (EPFL), June 1997.

[5] M. Bucher, "Analytical MOS transistor modelling for analog circuit simulation", Ph.D. thesis, no. 2114, Lausanne, EPFL, 1999.

[6] STARC Physical Design Gr, "HiSIM1.2.0 release note", Apr. 2003, http://home.hiroshima-u.ac.jp/usdl/HiSIM.shtml

[7] G. Gildenblat and T. L. Chen, "Overview of an advanced surface-potential-based MOSFET model", in *Tech. Proc. Fifth Int. Conf. Model. Simul. Microsyst.*, Puerto Rico, 2002, pp. 657–661.

[8] "BSIM3V3 MOSFET model", http://www.device.eecs.berkeley.edu/ bsim3

[9] "BSIM4 MOSFET model", http://www.device.eecs.berkeley.edu/ bsim4

[10] Minitab Inc. "MINITAB®for Windows release 12.21 manual", 1998.

[11] C. Enz, "An MOS transistor model for RF IC design valid in all regions of operation", *IEEE Trans. Microw. Theory Tech.*, vol. 50, no. 1, 2002.

[12] M. J. Deen and T. A. Fieldy, "CMOS RF modeling characterization and applications", *World Sci.*, 2002.

[13] F. Krummenacher, M. Bucher, and W. Grabiński, "Deliverable D2.1, RF EKV MOSFET model implementation", CRAFT, Eur. Project, no. 25710, WP2, July 2000.

[14] F. Krummenacher, M. Bucher, and W. Grabiński, "Deliverable D2.2, RF EKV MOSFET model implementation", CRAFT, Eur. Project, no. 25710, WP2, July 2000.

[15] H. M. Greenhouse, "Design of planar rectangular microelectronic inductors", *IEEE Trans. Parts, Hybr., Packag.*, vol. PHP-10, no. 2, June 1974.

[16] S. S. Mohan, "Modeling, design and optimization of on-chip inductors and transformers", CIS, Stanford University, June 1999.

Sadayuki Yoshitomi was born in Sasebo, Japan, in 1965. He received the B.E. and M.E, and Ph.D. degrees from Yokohama National University, Yokohama, Japan, 1988, 1990, and 1993, respectively. The title of his doctor thesis was the "The study on the silicon dioxide film deposited by liquid phase deposition (LPD)". His background

is the physics of Si/SiO_2 interface. In 1993, he joined the Research and Development Center, Toshiba Corporation, Kawasaki, Japan, where he was engaged in the research and development of BiCMOS device technologies. His current field of interest is the application of simulation technologies for telecommunication applications such as system simulation for next generation ICs, electro-magnetic simulation and SPICE modelling.

e-mail: Sadayuki.Yoshitomi@toshiba.co.jp
TOSHIBA Corporation Semiconductor Company
Microelectronics Center
2-5-1 Kasama, Sakae-Ku
Yokohama-City, Kanagawa, 247-8585 Japan

Hideki Kimijima was born in Yokohama, Japan, in 1973. He received the B.E. and M.E. degrees from Yokohama National University, Yokohama, Japan, in 1995 and 1997, respectively. In 1997, he joined Microelectronics Engineering Laboratory, Toshiba Corporation, Kawasaki, Japan, where he was engaged in the research and development of device technology of RF-CMOS. Since 2000, he has been involved for the research and development of device technology of RF BiCMOS, especially SiGe BiCMOS process in Semiconductor Company, Toshiba Corporation, Kawasaki, Japan. His current interests are device physics of bipolar transistor and CMOS.

e-mail: Hideki.Kimijima@toshiba.co.jp
TOSHIBA Corporation Semiconductor Company
Microelectronics Center at Kitakyushu Operation
1-10-1 Shimoitozu, Kokura-Kita-Ku
Kitakyuushyu-City, Fukuoka, 803-8686 Japan

Kenji Kojima received the B.E. and M.E. degrees in electrical engineering from Kanagawa University, Yokohama, Japan, in 1990 and 1992, respectively. In 1992, he joined the Medical Equipment Division of Toshiba Corporation, Tochigi, Japan. In 1996, he joined the Semiconductor Division, Toshiba Corporation, Kanagawa, Japan. Since 2001, he has been engaged in the research and development of device technology of advanced CMOS analogue and logic in SoC research and development center.

e-mail: Kenji3.kojima@toshiba.co.jp
TOSHIBA Corporation Semiconductor Company
Advanced CMOS Technology Department
Soc Research and Development Center
8 Shinsugita-Cho, Isogo-ku
Yokohama-City, Kanagawa, 235-8522 Japan

Hideyuki Kokatsu was born in Tokyo, Japan, in 1967. He received the B.S. and M.S. degrees in theoretical physics from Shizuoka University, Shizuoka, Japan, in 1990 and 1992, respectively. In 1992, he joined the Semiconductor System Engineering Center, Toshiba Corporation, Kawasaki, Japan. He currently is engaged in the design and development of high-frequency analogue integrated.

e-mail: Hideyuki.Kokatsu@toshiba.co.jp
TOSHIBA Corporation Semiconductor Company
Microelectronics Center
2-5-1 Kasama, Sakae-Ku
Yokohama-City, Kanagawa, 247-8585 Japan

Invited paper

Trends in assembling of advanced IC packages

Ryszard Kisiel and Zbigniew Szczepański

Abstract—In the paper, an overview of the current trends in the development of advanced IC packages will be presented. It will be shown how switching from peripheral packages (DIP, QFP) to array packages (BGA, CSP) and multichip packages (SiP, MCM) affects the assembly processes of IC and performance of electronic systems. The progress in bonding technologies for semiconductor packages will be presented too. The idea of wire bonding, flip chip and TAB assembly will be shown together with the boundaries imposed by materials and technology. The construction of SiP packages will be explained in more detail. The paper addresses also the latest solutions in MCM packages.

Keywords— IC packages, SiP, wire bonding, TAB, flip chip.

1. Introduction

The development of the electronics industry is dominated by communication products, which are characterised by rapid market introduction and fast mass-manufacturing capabilities. The main drivers of this development are miniaturisation and styling or eco design and production. The industry has made great strides in reducing the package size from the dual in line package (DIP) and quad flat pack (QFP) to ball grid array (BGA) and chip scale package (CSP). Table 1 illustrates the current trends in IC packaging. Standard leadframe packages still have the largest market share. Packages like BGA and CSP have gained increasing importance and popularity. CSPs are used more often than BGAs, in packaging, e.g., of memory products, controllers and digital signal processors, due to their smaller size.

Table 1
Trends in IC packages, 2001–2006 [1]

Package type	Standard (QFP)	BGA	CSP
Max body size [mm]	30 → 32	40 → 50	27 → 20
Max number of I/Os	300 → 320	900 → 3500	400 → 650
Min lead/ball pitch [mm]	0.4 → 0.3	0.75 → 0.5	0.5 → 0.3
Max package thickness [mm]	1.4 → 1.0	2.0 → 1.0	1.2 → 0.7
Market share [%]	88 → 78	4 → 7	4 → 15
US$ ct/I/O	0.8 → 0.4	0.9 → 0.6	0.9 → 0.4

Table 2 compares the main features of QFP, BGA and CSP at a comparable I/O number in the range of 200–300. It is obvious that switching to CSP packages reduces the physical area occupancy on the PCB, about three times compared to QFP for this I/O range. Shorter connections offered by CSP reduce parasitic inductance and capacitance.

Table 2
Comparison of key features of various packages [2]

Feature	QFP	BGA	CSP
I/O	208	225	313
Pitch [mm]	0.5	1.27	0.5
Footprint [mm^2]	785	670	252
Height [mm]	3.37	2.3	0.8
Package to die ratio	8	7	1
Inductance [nH]	6.7	1.3 – 5.5	0.5 – 2.1
Capacitance [pF]	0.5 – 1	0.4 – 2.4	0.05 – 0.2

Compared to leaded packages such as QFPs, BGA and CSP packages are expensive. The cost is about twice as high, but when cost per I/O is taken into account the prices become almost the same. In addition, the substrates on which chips have to be used are generally more expensive because more layers and possibly microvias are needed.

2. Packaging technology requirements

System integration is a key element for manufacturing future products. The idea behind system integration is to combine individual components and subsystems into a functional electronic system. Assembly and packaging of the semiconductor products are an essential part of this process. The most difficult challenges facing the assembly and packaging industry, according to the ITRS roadmap, are as follows [1, 3]:

- Improved organic substrates. The substrates must be compatible with Pb-free solder processing. Improved impedance control and lower dielectric losses are needed to support higher frequency applications. The substrates must be characterised by low moisture absorption, improved planarity and low wrapage at higher process temperatures as well as low-cost embedded passives.
- Improved underfillers for flip chip on organic substrates. Underfiller must have improved flow, fast dispense/cure properties, better interface adhesion, and lower moisture absorption, as well as higher operating temperature range for automotive applications and Pb-free soldering in liquid dispense underfiller.

- Impact of Cu/low κ on packaging. Direct wire-bond and bump to Cu must be possible. Mechanical strength of dielectrics must be improved and bump as well as underfill technology must assure low κ dielectric integrity. New tests to measure the critical properties need to be developed.
- New system level technologies must be capable to integrate chips, passives and substrates. Embeded passives may be integrated into the "bumps" as well as the substrates. Bumpless area array technologies must be developed, for example for face-to-face connection.
- Pb, Sb and Br free packaging materials. New materials and processes must meet new requirements, including higher reflow temperature and reliability under thermal cycling.

To fulfil these requirements special attention must be paid to [8]:

– area array packaging (flip chip, CSP, SiP),
– cost efficient and flexible substrate materials,
– thin semiconductor chips,
– cost efficient bumping process,
– fine pitch and multilayer technology,
– integration of passive components into the substrate,
– integration of optical and electrical signal transmission,
– material characterization (model and measurement),
– environmentally compatible choice of materials and processes.

Such requirements are more prominent in cost-performance applications, such as notebooks, personal computers and telecommunication equipment [3, 4]. The key single-chip-package technology requirements are shown in Table 3. Economic aspects are especially important here. Assembly and packaging costs are expected to decrease over time on a cost-per-pin basis, but the chip and package pin-count is increasing more rapidly than the cost-per-pin is decreasing.

Table 3
Development of selected packaging parameters in the next decade [3, 4]

Parameter	Year		
	2002	2006	2010
Chip size [mm^2]	178	206	268
Cost [cents/pin]	0.75–1.44	0.56–1.03	0.48–0.98
Core voltage [V]	1.5	0.9	0.6
Package pincount	480–1320	550–1936	780–2700
Package thickness [mm]	1.0–1.2	0.8–1.2	0.65–0.80
Performance on chip [MHz]	2320	5630	12000
Performance: peripheral buses [MHz]	200/660	300/966/1062	300/1415
Junction temp. max [°C]	85	85	85
Chip interconnect pitch [μm]	35	20	20
Flip chip pitch [μm]	160	130	90

3. Bonding technology for chip packaging level

There are three chip-level interconnection technologies currently in use:

– wire bonding (WB),
– flip chip (FC),
– tape automated bonding (TAB).

Wire bonding is a well established bonding technology. This technique is used for over 90% of all interconnections. Flip chip is another technique that has been a steadly gain in importance, particularly in new applications where system performance and miniaturisation is important. It also offers special advantages in terms of high accuracy in chip placement by self-adjustment when solder bumps are on the die. This is of special interest for the integration of optical interfaces in electro-optical devices.
A concept of TAB process was initiated in the 1960s by General Electric. In the mid-90s TAB process was widely used in high volume products, such as Pentium laptops and high resolution printers [7]. Now it is not use so often. The comparison of the three bonding techniques, mentioned above, in aspect of signal propagation delay is shown in Table 4.

Table 4
Typical values for lead capacitance and inductance [7]

Bonding technique	Capacitance [pF]	Inductance min [nH]
Wire bond	0.5	6
TAB	0.6	2
Flip chip	0.1	0.2

Propagation time is directly proportional to the length of interconnection, the longer the lead the greater the propagation delay and the slower the off-the-chip system speed.

3.1. Wire bonding assembly

In "standard" packaging , the interconnection between the die and the carrier is made using the wire. Different types of wires are used for the wire bond process: gold, aluminium and copper. The die is attached to the carrier face up, then a wire is bonded first to the die, then looped and bonded to the carrier (Fig. 1) [9]. Wires are

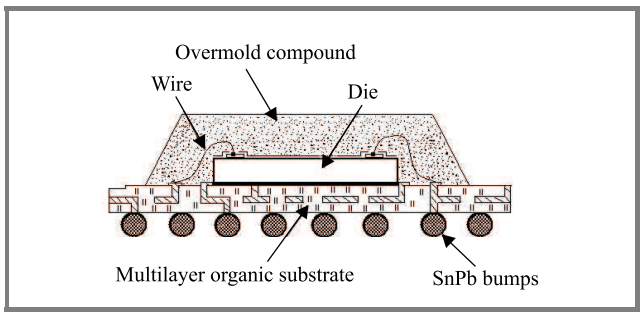


Fig. 1. Wire bonding technology applied in BGA package.

typically 1–5 mm in length and 25–35 μm in diameter. The process used most often to realize these interconnections is thermosonic bonding with gold wire.

3.2. Flip chip assembly

The term flip chip describes a method of establishing electrical connectors between the die and the package carrier. The interconnection between the die and the carrier in flip chip packaging is made through a conductive "bump" placed directly on the die surface. The bumped die is then "flipped over" and placed face down, with bumps connecting directly to carrier (Fig. 2). A bump is typically 70–100 μm high and 100–125 μm in diameter.

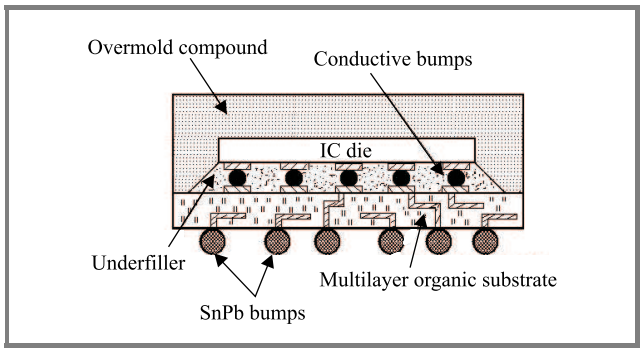


Fig. 2. Flip chip technology for inner assembly in BGA package.

The flip chip connection is generally formed in one of two ways: using solder or a conductive adhesive. The most common packaging interconnect is the eutectic SnPb solder. The solder bumped die is attached to a substrate by means of a solder reflow process. After the die is soldered, the underfiller is added between the die and the substrate. Underfiller is designed to control the stresses in the solder joints caused by the difference in thermal expansion between the silicon die and the carrier. Once cured, the underfiller adsorbs the stresses, reducing the strain in the solder bumps, greatly increasing the life of the finished package and reliability of solder joints.
The other method of establishing the flip chip connection is adhesive bonding. Various methods of joining can be used: connecting with isotropically or anisotropically conductive adhesives and using non-conducting materials. The bump material is predominantly gold. Without the danger of solder-bridge formation, smaller pitches can be implemented with adhesive techniques. Furthermore, substrates can be used which cannot withstand the melting temperature of the eutectic lead/tin solder or lead-free alloys. Adhesive flip chip bonding, based mainly on the anisotropically conductive adhesive, is the dominant technology for LC-displays assembly.
The main benefits and features that flip chip can offer are:

- High package density without the need for lead frame.
- Enhanced electrical performance as a result of short distance between chip and substrate, including minimal propagation delay, minimal transmission losses at high frequency, and low parasitic capacitance as well as inductance values.
- Low thermal resistance solder bonds serve primarily as heat-dissipating paths, thus suitable for high-speed electronic devices.
- Power can be spread directly from the core of the die to laminate (substrate). This greatly decreases the noise of the core power and improves performance of the silicon.
- Better suited to making small electronic modules, due to the low profile, reduced weight and space requirements.
- The entire surface of the die can be used for interconnecting. So, it can support vastly larger numbers of interconnects on the same die size. Flip chip technique is mainly used for assembling ball grid arrays. Such packages use PCB with, pitch above 500 μm and require large substrate areas. In the case of smaller pitch the CSP packages are used.

3.3. Tape automated bonding interconnection technology

The TAB process involves bonding Si chips to patterned metal on polymer tape, e.g., copper on polyimide, using thermocompression or solder bonding. Subsequent processing is carried out in strip form through operations such as testing, encapsulation, and burn-in, followed by excising of the individual packages from the tape and attachment to the substrate or board by outer lead bonding.
The copper leads on a thin polyimide film are made using subtractive photo-chemical process. Finally the leads have to be plated to retain solderability. Gang bonding of the bumped chip to copper tin plated leads on polyimide tape is performed using thermode heater. After chip protection TAB device is excised from polyimide tape, see Fig. 3 and next attached to the board surface bonding pads. Attachment technique includes pulse heated hot bar. This hot bar or thermode heater mechanically presses the leads onto bonding pads, which eliminates any coplanarity errors arising from lead forming. The main disadvantage of hot

bar soldering is the fact that a thermode has to be mechanically rigid, to exert the force on the component leads. Unfortunately, this force can be high enough to damage the substrate. To solve this problem, spring loaded thermode or thermode which can be held on a spherically mounted suspension unit are used. The last one allows the blades of thermode to be planarised with respect to the surface of the substrate. This yields uniform forces across all leads around the bonding site. Placement accuracy for TAB devices is tighter than for conventional SMD components, since the lead pitch is smaller in comparison to SMD. So, the machine positioning system must include reliable machined hardware, precision encoders and stable base.

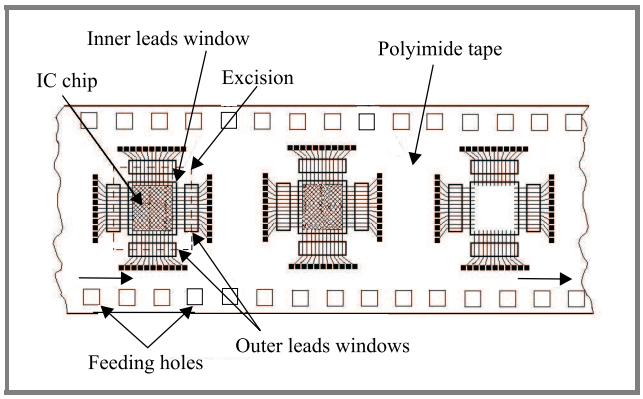


Fig. 3. The idea of TAB connection.

The width and spacing of the inner lead bonding pads of the TAB process used currently is 50 μm. The corresponding dimensions of the outer lead bonding pads are 100 μm if the substrate technology allows it. TAB offers advantages in the area of improved electrical and thermal performance. TAB process is also a very good choice for building reliable MCMs.

4. Wafer level packaging

The wafer level CSP is a packaging technology where the majority of the packaging process steps are carried out at the wafer level [3, 8]. It is the most interesting technology from the point of view of miniaturisation. These packages are made before the wafer is sawed, therefore their size cannot be larger than that of the die. Most packages of this type are made by re-routing the bond pads on the wafer, followed by bumping.

The limitation on wafer level packaging (WLP) is how many I/O can be placed under the chip. Typical application market for WLP is projected for ICs with total I/O below 100 and adequate silicon area. A key enabling technology to take full advantage of WLP is the development of wafer level test and burn-in. Most WLPs with I/O pitch equal or greater than 0.5 mm do not require the use of underfiller and can be directly implemented into a standard surface mount technology (SMT) process flow.

The reliability of WLP that use bond pad redistribution is a problem, especially when they are to be mounted on FR-4 boards. Due to different thermal expansion coefficients of silicon and printed board, the maximum die size is limited. So, the use of wafer level packages focuses on ICs with relatively low pin count (small size). Many portable equipment manufactures use underfilling. It improves resistance to solder fatigue and improves the resistance to mechanical stress, such as, shock, board wrapage and vibration. The reliability of WLP can also be improved when the stand-off between the component and the printed board is enlarged [11].

WLP technology is ideal for portable communications and related applications that require a low cost packaging solution with small form factor and improved signal propagation characteristic.

A disadvantage of WLP and flip chip is that their size is not standardized. The package has the same dimensions as the die, which can have any dimension. To avoid the need for customised test sockets, trays, reels, etc., it is best to perform as many operations as possible at the wafer level.

5. Types of IC array packages

5.1. BGA packages

The BGA package is based on a PCB substrate. The standard core thickness of substrate is typically from 0.2 to 0.4 mm with 18 μm copper on each side. The silicon chip is die bonded to the top side of the substrate using die attach adhesive. The chip is then gold wire-bonded to wire bond pads on the substrate. Traces from the wire bond pads take the signal to vias which carry them to the bottom side of the substrate and then to circular solder pads. The bottom side solder pads are laid out on square or rectangular grid with either a constant 1.5 mm, 1.27 or 1.0 mm pitch. An overmold is then performed to completely cover the chip, wires and substrate wire bond pads. Majority of BGAs will utilize a wire bond interconnection on the periphery of the IC (Fig. 1). Area array flip chip connections to BGAs are needed for high I/O counts or high power chips (Fig. 2). Laminate-based array packages will require the use of underfillers to reduce the shear stress load on the flip chip interconnections for large die, due to the large difference in the CTE between the silicon IC and the substrate. BGA packages will provide good solutions for the pincount range above 200.

5.2. CSP technology

Defined by a criterion of size, CSP are no smaller than 80% of the chips they house. For CSP, the pitch drops below 0.8 mm. CSPs normally contain an interposer on which the die is attached by either wire-bonding or flip chip interconnect [8]. A rewiring of the I/O pads is done on a flex interposer or directly on the wafer/die surface. The area distributed I/O contacts (solder ball height smaller

than 250 μm) for board assembly are satisfied by the PCB land pitches in size 250 μm to 500 μm. In case of wafer level redistribution the packaging cost can be reduced tremendously. CSP package with flexible interposer is shown in the Fig. 4. The relatively expensive underfilling process after assembling can be avoided because a stress compensation can be implemented into the "package".

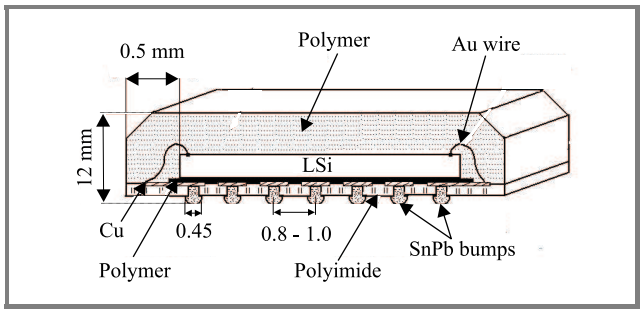


Fig. 4. Cross-section of CSP package.

The comparison of these two type of package is shown in Table 5. CSP packages will provide good solutions for the pincount range above 400. In CSP package the silicon-to-package area ratio is increased up to 100% in most versions. Having reduced the package body to equal the die size the next available step is the integration in the z-direction by stacking more dies in the same package [10].

Table 5
Development of BGA and CSP packages parameters in the next decade [3, 4]

Package/parameter	Year		
	2002	2006	2010
BGA – solder pitch [mm]	1.00	0.65	0.50
BGA – body size [mm]	23	18	17
– possible pincount	484	729	1089
CSP – area array pitch [mm]	0.40	0.30	0.30
CSP – body size 10 mm			
– possible pincount	320	540	540
CSP – body size 21 mm			
– possible pincount	572	1280	1280

The CSP packages provide a potential solution where low weight and small size are the requirements. These packages are only slightly larger than the chip itself, and are available in a variety of configurations and materials combinations. The size may range from 4 to 21 mm. Significant size reduction can only be realized by using CSP, which has pitches ranging from 0.4 mm to 0.3 mm. Four type of CSPs can be distinguished [8]:

– CSPs with rigid interposer (organic or ceramic),
– CSPs with a flexible interposer (usually polyimide),
– CSPs with a leadframe,
– wafer level CSPs.

The next level of wafer level packaging is reached by integration of passive (R, L) and active devices (Si, GaAs ICs). Such system are known as multi-chip modules (MCM).

5.3. MCM technology

The simplest definition of an MCM is that of a single package containing more than one IC. More precisely, in the MCM there are multiple bare dies mounted along with signal conditioning or support circuitry such as capacitors, resistors and other parts on a laminate or ceramic base material. MCMs introduce a packaging level between the application-specific IC (ASIC) and board level. Digital and analog functions can be mixed without serious limitations, and an ASIC can be composed easily with standard processors, devices and memories in one package. The next generation of MCMs could even have optical I/Os as an option. In such solution, high-speed components can be placed closer to each other, the load on the IC output buffers is lower and signal transmission properties are better. The new version of MCM is called system-in-package (SiP) where dies are mounted one on the other.
Multi-chip module technology has been previously used for high performance applications like mainframe CPU's. MCMs of this type are typically very complex and expensive. MCM-C applied in the IBM 4300 computer used a glass-ceramic substrate with more than 40 layers and containing more than 25 ICs [2].

5.4. SiP packages (multi-stacked LSI)

The term system-in-package is used to describe multi-chip IC, such as stacked die CSP or MCM-PBGAs. SiP is a functional block or module, which incorporates ICs, passive components, antennas, etc. Curent SiPs assemble 2–5 dies in a single package combining wire bond with peripheral or center pads, die-to-die- bonding, flip chip and SMT [12, 13]. SiP is used as a standard component in board level manufacturing (Fig. 5). It is based on a merger of mainstream, high volume and low cost IC assembly technology and surface mount technology. The IC assembly technology is used to interconnect the IC chips to the SiP substrate while the SMT technology is used to interconnect the IC chips to the SiP substrate while the SMT technology is used to connect passive components and other

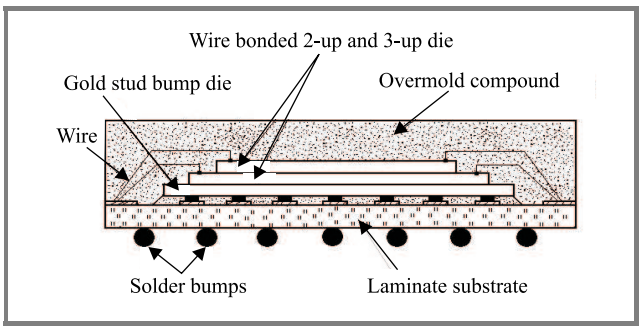


Fig. 5. Cross-section of SiP package.

SMT-compatible components (e.g., connectors) to the SiP substrate.
Current SiP solution is based on low cost IC assembly technology and surface mount technology. The SiP can be manufactured using ceramic, leadframe, organic laminate or even tape-based substrates. The following attributes are used to define SiPs:

- Includes chip-level interconnect technology. In other words, flip chip, wire bond, TAB or other interconnect directly to an IC chip.
- Quite often includes passive components. The passives may be either surface mounted discrete components or they may be embedded into or manufactured on the substrate material, as buried components.
- Typically includes more than one IC chip, so integration of various die technologies (e.g., Si, GaAs, SiGe, SOI, MEMS and optical) in the same package is possible.
- May include other components necessary to bring the SiP to more complete functional system or subsystem level – such as housings, lids, RF shields, connectors, antennas, batteries, etc.

The SiP thickness started at 1.4 mm for a 2-stack but has now been reduced to 1.2 mm for the 3-stack and already there is a demand for a 4-stack in the 1.4 mm package. The end applications continuously require thinner packages. Reduction in package thickness requires thinner substrate, thinner die and smaller solder balls with a shrinking pitch. The majority of common 2-L laminate substrates have been reduced to 0.21 mm and the copper/polyimide tape substrates today use 0.05 mm polyimide and ultimately will migrate to 0.025 mm. The thinnest die in high volume production today is 0.14 mm, but this will soon be reduced to 0.10 mm and ultimately to 0.05 mm.
The primary driver behind SiP is maximum functionality in the minimum footprint. Assembling multiple die side-by-side in one package reduces interconnect length and can increase electrical performance. Lower cost is the driver for most packaging. SiP cost is only fractionally higher than the single-die package cost because it utilizes only one substrate, the backend operations are the same and are assembled using the same infrastructure.
The need to mix ASIC or digital signal processor (DSP) chips with memory has pushed the development of a three-stack SiP with sequentially smaller die stack. Same-die-stack technologies all using the 1.4 mm thickness CSP family. Technically these options can be grouped into three SiP families.

- **SiP I**. Wire bonded die-up SiP assembles die with peripheral bond pads in 2-up or 3-up configurations in one or two stacks. Through the use of a thin flex tape-based interposer and thin backgrind wafer technology, the triple-chip stacked SiP still meets the 1.4 mm maximum thickness ceiling required for advanced handsets.
- **SiP II**. Wire bonded die-up and die-down SiP with peripheral and centre pads at the bottom. The most common application is with DRAM with center pads at the bottom. The wire bonding to the die is done through a slot in the substrate that is subsequently encapsulated before the remaining dice are assembled.
- **SiP III**. Flip chip and wire bond SiP with die-up configuration are shown in Fig. 5. The flip chip connection is the gold stud bump formed directly on the aluminium die pads using wire bond technology.

Though the majority of CSPs today use the 0.8 mm pitch with 0.48 mm ball diameter, higher pincount applications with smaller footprints are migrating to 0.5 mm pitch with 0.30 mm diameter ball.
SiP offers the flexibility of mixing analog and digital designs, each optimised separately for performance and cost, and also integrates technologies like CMOS and SiGe or GaAs that cannot be mixed in one die to produce a system-on-chip (SoC).
Integrating many ICs in a SiP format and using a high-density substrate can result in significant cost saving by reducing routing complexity and layer count in the motherboard. Also, SiP solution can save a significant amount of space on the motherboard, which will enable more functionality to be integrated in a system board. The SiP equates to be functional block of the system that can be standarized across a product family and dropped into a new system board design with minimal effort. The SiP solution usually results in system cost savings by reducing motherboard cost, reducing system size and simplifying system assembly and rework process.
Communication system uses SiP technology for RF and wireless devices, networking and computing, image sensors and memory applications such as flash cards. In today's systems, the ASIC and memory are packaged separately and mounted to the system board. Considerable cost, size and performance benefits can be realized by integrating the ASIC and memory devices into a system in package configuration. The ASIC device is mounted in a flip chip configuration along with memory devices mounted on the same substrate beside the ASIC. Decoupling capacitors and other passives can also be mounted to the substrate. Memory devices are packaged and tested in individual CSP packages before they are mounted to the substrate using a conventional SMT process. This eliminates the yield loss issues associated with traditional MCM designs. Since all routing between the ASIC and memory is done on the first or second layer of the substrate, signal integrity is optimised.

6. Conclusion

Packaging technology plays an important role in the realisation of a new generation of products. This requires

the development of new packaging technologies and materials. The estimation of packaging technologies can be done by an indicator the so called packaging efficiency. It is defined by the ratio of the silicon chip to the area of the package. The best level of efficiency for pheripheral packages, like QFP, can reach 50%; for area array package like BGA/CSP does not exceed 90% and for wafer scale or bare chip solutions can achieve 100% silicon efficiency. Through SiP technologies, 200% efficiency can be approached without having to tackle the more complex second level interconnect and reliability challenges associated with 0.5 mm and finer ball pitch.
Stacked chip interconnection in combination with ball grid array and chip scale package technology platforms are merging to create a new wave of 3D package integration.

References

[1] C. Beelen-Hendrikx and M. Verguld, "IC packaging: what's around the corner?", *OnBoard Technol*., p. 29, Oct. 2002.

[2] J. S. Hwang, *Environment-Friendly Electronics: Lead-free Technology*. Bristol: Electrochemical Publ., 2001, Chapter 2.

[3] International Technology Roadmap for Semiconductors, 2001 Edition, Assembly and Packaging, http://public.itrs.net

[4] International Technology Roadmap for Semiconductors, 2002 Update, http://public.itrs.net

[5] "Microelectronic packaging-ready for the next product generation", *IZM Ann. Rep*., p. 45, 2000.

[6] "TAB in volume", *Electron. Produc.*, pp. 17–18, Feb. 1996.

[7] C. E. Presley, "FCA: a critical advantage for designers", *Adv. Packag*., p. 38–42, Apr. 1999.

[8] C. Beelen, "Trends in assembly processes for miniaturised consumer electronics", *Glob. SMT & Packag.*, vol. 1, no. 1, p. 24–31, 2001.

[9] Z. Szczepański, "Nowe osiągnięcia konstrukcyjne miniaturowych obudów dla układów scalonych dużej skali integracji", *Elektronika*, no. 7–8, pp. 18–23, 1999 (in Polish).

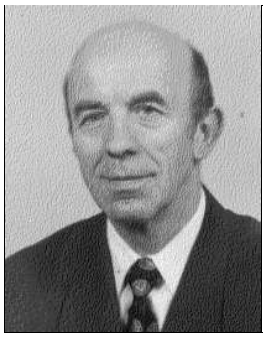
[10] M. Karnezos, F. Carson, B. Zahn, and T. K. Choi, "Benefits and technical issues of SiP", *Glob. SMT & Packag.*, vol. 1, no. 4, p. 29, 2001.

[11] K. Friedel, "Zaawansowane technologie w montażu aparatury elektronicznej", *Elektronika*, vol. XLIII, no. 7–8, pp. 25–27, 2002 (in Polish).

[12] M. Kada and L. Smith, "Advancements is stacked chip scale packaging, provides system-in-a-package functionality for wireless and handheld applications", Jan. 2000, http://www.amkor.com/products

[13] C. M. Scanlan and N. Karim, "System-in-package technology, application and trends", Jan. 2002, http://www.amkor.com

Ryszard Kisiel received the M.Sc. degree in material science from the Warsaw University of Technology, Poland, in 1974. His professional carrier began the same year as an academic teacher in Department of Electronics. He was responsible for the courses connected with the basic of electronic equipment construction and the thin as well as thick film hybrid circuits. He received the Ph.D. degree in electron technology from the Warsaw University of Technology in 1983. His recent research and development activities have been in the areas of surface mount technology, especially in applying lead-free technology and using electrically conductive adhesives in PCB assembly. He is currently employed at Warsaw University of Technology, Department of Electronics and Information Technology, as the Assistant Professor. He is a member of the IMAPS-Poland Chapter and IEEE CPMT Society.
e-mail: kisiel@imio.pw.edu.pl
Institute of Micoelectronics and Optoelectronics
Warsaw University of Technology
Koszykowa st 75
00-662 Warsaw, Poland



Zbigniew Szczepański is the Assistant Professor in the Institute of Microelectronics and Optoelectronics from the Warsaw University of Technology. He received the M.Sc. degree in mechanical engineering at the Warsaw University of Technology in 1962 and the Ph.D. degree in heat transfer engineering in 1971. In 1976 he initiated Microelectronics Technologies and Materials Laboratory at the Department of Electronics of Warsaw University of Technology. The main his research interests are: hybrid circuit technology, advanced interconnection technologies, thick film sensor technology, assembly techniques for sensors. He has published above 100 papers and student books and taken part in many international conferences. He is a member of IMAPS since 1982.
e-mail: szczepanskiz@imio.pw.edu.pl
Institute of Micoelectronics and Optoelectronics
Warsaw University of Technology
Koszykowa st 75
00-662 Warsaw, Poland

Paper

Ultra-shallow nitrogen plasma implantation for ultra-thin silicon oxynitride (SiO_xN_y) layer formation

Tomasz Bieniek, Romuald B. Beck, Andrzej Jakubowski, and Andrzej Kudła

Abstract— The radiation damage caused by low energy r.f. plasmas has not been, to our knowledge, studied so far in the case of symmetric planar plasma reactors that are usually used for PECVD processes. The reason is that, unlike non-symmetrical RIE reactors, such geometry prevents, basically, high-energy ion bombardment of the substrate. In this work, we present the results of experiments in which we have studied the influence of plasma processing on the state of silicon surface. Very low temperature plasma oxidation has been used as a test of silicon surface condition. The obtained layers were then carefully measured by spectroscopic ellipsometry, allowing not only the thickness to be determined accurately, but also the layer composition to be evaluated. Different plasma types, namely N_2, NH_3 and Ar, were used in the first stage of the experiment, allowing oxidation behaviour caused by the exposure to those plasma types to be compared in terms of relative differences. It has been clearly proved that even though the PECVD system is believed to be relatively safe in terms of radiation damage, in the case of very thin layer processing (e.g., ultra-thin oxynitride layers) the effects of radiation damage may considerably affect the kinetics of the process and the properties of the formed layers.

Keywords— MOS technology, plasma processing, shallow implantation, radiation damage.

1. Introduction

ITRS roadmap [1] forecasts an increase of the packaging density and improvement of the performance of silicon integrated circuits by the reduction of the of gate dielectric thickness, among other factors.

In mixed logic/memory circuits manufactured as a system on a chip, two different thicknesses of dielectric layers are required. The solution the most advanced technologically would be to form both dielectric layers during a single process. This may be possible through the oxidation of a nitrogen-implanted silicon layer, where the oxidation rate depends on the nitrogen implantation dose and profile (e.g., [2, 3]).

The experiments presented in this work are a part of a broader study that examines the possibility of conducting both stages of ultra-thin oxynitride formation (e.g., ultra-shallow nitrogen implantation and silicon oxidation) in one technological reactor.

The important issue here is the possibility of carrying out extremely shallow ion implantation by means of planar r.f. reactors. For the time being planar reactors have been believed to enable very limited bombardment only, especially in the case of symmetrical reactors – used for PECVD process. In non-symmetrical reactors used for reactive ion etching (RIE) the effects of ion bombardment have been employed practically to etch layers, no data or even estimation have, however, been presented (to our knowledge) so far on the depth of implantation and significance of this effect.

Such effects may not only influence the process of layer forming, but also influence their properties.

2. Experimental

In contrast to the methods presented so far in the literature where classical implanters or the ion immersion implantation in plasma (IIIP) method were used to carry out ultra-shallow implantation, in our work we used typical planar r.f. plasma reactors usually applied in PECVD. The processes of plasma implantation of nitrogen and subsequent plasma oxidation are both taking place at a very low temperature 350°C.

The nitrogen N_2 and ammonia NH_3 plasmas were used in this investigation (stage 1). This allowed the effect of easier dissociation and excitation of ammonia to be compared to that of nitrogen.

The subsequent process of plasma oxidation (stage 2) was carried out in conditions resulting from the studies of plasma oxidation at very low temperatures presented in [4]. The kinetics of oxidation are highly dependent on the state of the silicon surface. The disorder of this region (e.g., amorphisation) obviously results in a higher oxidation rate.

Spectroscopic ellipsometry was used to measure the thickness of the obtained layers. Since either nitrogen or ammonia plasma was used in the first process (stage 1), we could expect some nitride content in the layer – the appropriate optical model had to be incorporated. Our earlier experience has proved that even in the case of low nitride-phase content, we may obtain very good agreement of the EMA model with the spectrum measured by ellipsometer in a very wide range of wavelengths. An example of fitting the EMA model to the real ellipsometric spectrum is shown in Fig. 1. In the EMA model, we included two phases most likely to be present in our layers, namely:

Table 1
Complete matrix of experiments – two-stage processes performed in a PECVD system; the results obtained by means of spectroscopic ellipsometry measurement analysis are also included

Process name	Oxid. only		50 W Ar		100 W Ar		50 W N_2		100 W N_2		50 W NH_3		100 W NH_3	
Stage 1 process														
Gas type			Ar				N_2				NH_3			
Gas flow [ml/min]			50				50				50			
Pressure [Tr]			0.5				0.5				0.5			
Temperature [°C]			350				350				350			
Process time [min]			5				5				5			
Power [W]			50		100		50		100		50		100	
Stage 2 process														
Power [W]	50		50				50				50			
Gas type	O_2		O_2				O_2				O_2			
Gas flow [ml/min]	50		50				50				50			
Pressure [Tr]	0.5		0.5				0.5				0.5			
Temperature [°C]	350		350				350				350			
Process time [min]	2	10	2	10	2	10	2	10	2	10	2	10	2	10
Results of the spectroscopic ellipsometry measurements														
Optical thickness [Å]	44	66	59	91	73	130	36	68	55	171	36	60	32	63
Refractive index (@630 nm)	1.477	1.471	1.488	1.477	1.453	1.460	1.443	1.463	1.509	1.433	1.627	1.542	1.663	1.526
Si_3N_4 content [%]	–	–	5.92	3.78	0.1	0.64	0.1	1.14	9.98	0.1	31.6	16.1	37.9	13.2

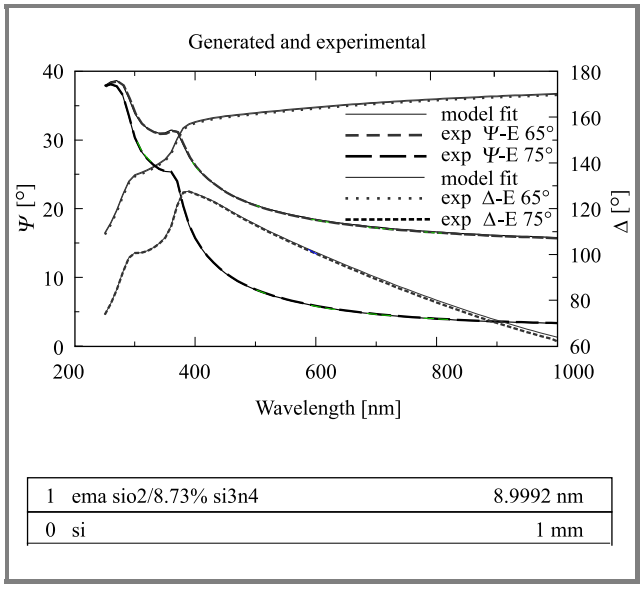


Fig. 1. An example of a fitting of the optical EMA model to measurement data in wide range of wavelength (from 250 nm to 1000 nm) for two incident angles 65° and 75°. Very good agreement allows both layer thickness determination and approximate evaluation of layer composition.

silicon dioxide (SiO_2) and silicon nitride (Si_3N_4), although we could also expect some "optically detected" consequences of the amorphisation of silicon. The fitting of this model to the measured spectrum fitting is very good (low values of mean deviation error) indirectly justifying our choice of phases included in the EMA model.

In order to isolate the impact of amorphisation of the silicon substrate, by implanted nitrogen ions, on the final kinetics of oxidation, reference experiments were done with argon implantation (stage 1) under similar conditions.

The complete set of experiments is presented in Table 1. It should be noted that each experiment was carried out for two process times (2 and 10 minutes) of plasma oxidation (stage 2), and two levels of r.f. power (50 W and 100 W) applied during nitrogen implantation (stage 1). The oxidation process was carried out with the power of r.f. oxygen plasma equal to 50 W.

The independent information about possible consequences of ion collisions with the silicon substrate was obtained from TRIM simulations. They allowed probable profiles of implanted atoms to be evaluated, as well as profiles and densities of the expected damage in the silicon crystal lattice – factors that obviously determine the kinetics of the subsequent process of plasma oxidation of silicon.

3. Results and discussion

The results of all ellipsometric measurements characterising all experiments performed in this study are also included in Table 1.

When looking at the obtained results one can immediately realise that the history (type of plasma process and

r.f. power used at stage 1) does have enormous impact on the final layer thickness and its composition. Basically, one could say that each of the process sets results in different oxide (oxynitride) thickness. However, studying the relative differences can bring more interesting information on the effects that are most probably responsible for those differences.

The undisturbed kinetics of the plasma oxidation process was observed by means of the reference process with oxygen plasma only (stage 2 only). The thickness dependence was comparable to the results obtained in our previous work on plasma oxidation [4]. The difference in the oxide thickness observed in the samples bombarded in argon plasma (stage 1) for both, lower and higher r.f. argon plasma power is clear (see Fig. 2). At the same time, for both oxidation times the obtained oxide layer thickness is higher in the case of samples exposed primarily (during stage 1) to Ar plasma. Furthermore, substantial difference in the obtained thickness is observed between 50 W and 100 W r.f. Ar plasma (see also Fig. 2).

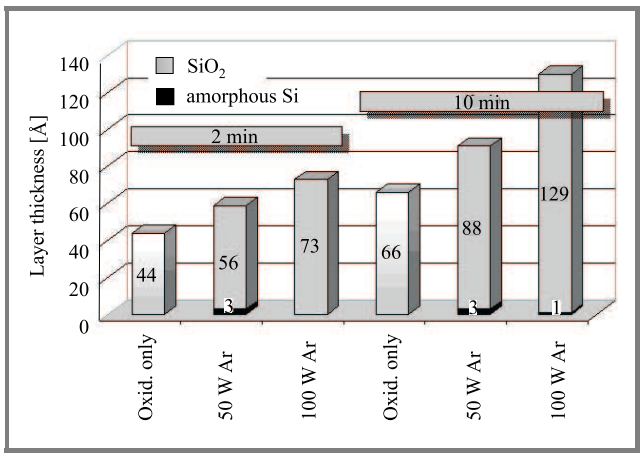


Fig. 2. The total and component-layer thickness (as evaluated using EMA model) of the layers obtained by plasma oxidation of samples exposed to argon plasma. The results of plasma oxidation only are used as a reference.

The observed results seem to be obvious, if we realise that due to chemical passivity of argon the only possible effect of silicon exposure to Ar plasma can be the damage to silicon surface due to energetic Ar ion bombardment. In these conditions the dependence of oxide thickness on r.f. power applied to plasma is simple – the higher the r.f. power, the higher ion bombardment is expected and thus, more damage is done to the silicon substrate. The damage obviously enhances oxide growth, therefore the final oxide thickness is the highest (in this set of experiments) for the process using the higher r.f. power (100 W) of argon plasma during stage 1. We have to keep in mind that, the enhancement of oxidation rate will end as soon as the whole damaged region of the silicon substrate is be consumed by plasma oxidation.

A different situation occurs when nitrogen or ammonia plasmas are used in the stage 1 process (see Fig. 3).

In the case of nitrogen plasma short oxidation time results in lower or slightly greater layer thickness (for lower and higher r.f. plasmas power, respectively). It should be noticed that layers formed in conditions of higher r.f. power contain also the nitride phase, as evaluated by means of ellipsometry. For longer (10 min) oxidation times the tendency is similar, although the layer thickness for higher power is exceptionally high.

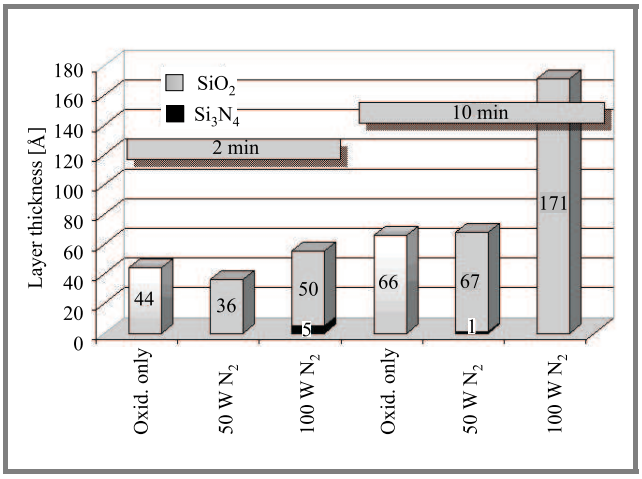


Fig. 3. The total and component-layer thickness (as evaluated using the EMA model) of the layers obtained by plasma oxidation of samples exposed to nitrogen plasma. The results of plasma oxidation only are used as a reference.

When analysing the possible reasons for such results we have to take a complex nature of the processes into consideration. One should be aware of the fact that apart from the amorphisation (damage) of the silicon subsurface region, the creation of SiN bonds (even at 350°C) is also possible in nitrogen containing plasma. The latter effect is possible due to the damage (broken bonds) resulting form nitrogen ion bombardment.

Thus, we have to consider two contradictory effects during plasma oxidation (stage 2). While oxide growth is slowed down by the presence of SiN bonds, at the same time it is also accelerated by the presence of broken silicon bonds in the subsurface region. Additionally, we have to remember that nitrogen atoms tend to be replaced by oxygen atoms in position bonded to silicon during oxidation even at such low temperatures.

The use of ammonia plasma results in a different behaviour from that observed in nitrogen plasma (compare Figs. 3 and 4). In all experiments of this series the final oxynitride layer thickness is lower than that obtained during reference oxidation, which proves the effective passivation of silicon surface by nitrogen implantation from ammonia. It is interesting to note that the nitride-phase content is much higher in these samples than in all those described before. Moreover, in contrast to previous cases, the oxidation time does not have much influence on the layer thickness. A careful look at Fig. 4 indicates that the dependencies of layer thickness on r.f. power are different for shorter and

longer oxidation times in this case. This obviously confirms the hypothesis presented above of different effects competing with one other during the stage 1 and stage 2 processes.

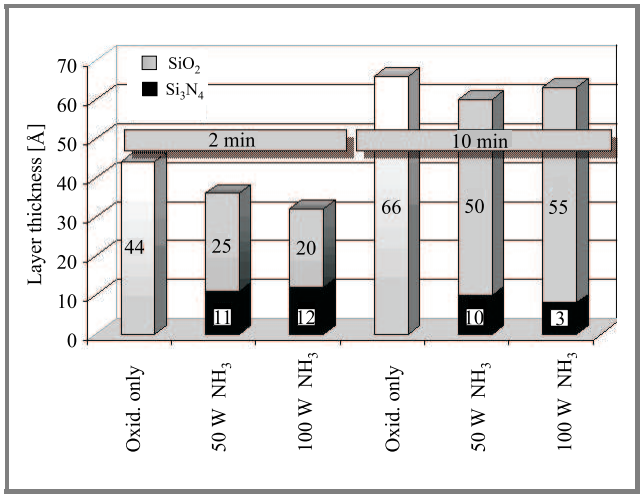


Fig. 4. The total and component-layer thickness (as evaluated using the EMA model) of the layers obtained by plasma oxidation of samples exposed to ammonia plasma. The results of plasma oxidation only are used as a reference.

Taking into consideration that both effects are caused by the phenomena taking place only in the ultra-shallow, subsurface region, we have to remember also that all these effects will disappear if an oxide layer thick enough is grown (thick enough to consume the affected silicon depth).

As a consequence, oxidation slow-down takes over for short oxidation times the, while oxidation rate is enhanced by broken silicon bonds for long enough oxidation times (longer than it takes for all or most of the nitrogen to be removed from the layer). This enhancement will continue until the whole silicon region affected by ion bombardment is consumed by the growing oxide.

Even during oxidation, at such low temperatures (350°C) the nitrogen atoms incorporated and bonded with silicon during the earlier process (stage 1) tend to be replaced by oxygen. As a result, the nitrogen content in the oxynitride layer is dependent on the oxidation time and inevitably tends to decrease during oxidation. The latter effect is very effective, as can be seen using the 100 W N_2 plasma process as an example. As the result of oxidation results in a ~170 Å oxide layer with no nitride phase present, as evaluated by spectroscopic ellipsometry (see Fig. 3).

The situation is little different, when ammonia plasma is used instead of nitrogen. As can be seen in Fig. 3, the nitride-phase content is much higher then, when compared with nitridation carried out in N_2 plasma. This observation can be explained by a well known fact that ammonia dissociates much easier in plasma than nitrogen. Consequently, for both short (2 min) and longer (10 min) oxidation times we can still observe the nitride phase in the layer.

It is also interesting to note that in the oxynitride layers formed by the oxidation silicon exposed previously to NH_3 plasma the nitride content does not change much between 2 min and 10 min oxidation times (compare Fig. 4). It seems that nitrogen replacement with oxygen is not so effective during oxidation in this case. The most probable hypothesis explaining this effect refers again to more effective formation of nitrogen in NH_3 plasma, which in turn allows thicker and denser nitride phase to be formed during this stage of experiment. Denser and more perfect nitride component may be less prone to oxygen replacement during plasma oxidation at low temperatures (< 350°C).

As it was already mentioned above, the oxidation rate is influenced by all the effects described above until the whole affected silicon subsurface region becomes oxidised. It is, therefore, crucial for understanding of the whole process to be aware of the range of the distances covered in the silicon substrate by ions implanted from r.f. plasma. As already mentioned before, this information may only be obtained by means of appropriate simulations or experimental measurements of the implantation depth (profile) and doses.

The energies of ions bombarding silicon samples in a PECVD planar reactor are expected to be very low – depending on the pressure and r.f. power used they could amount to a few hundreds of volts or even less. The only simulator that allows such low ion energies to be considered is TRIM [5] (a part of SRIM simulator). The problem in this case is, however, that no data is available on the accuracy of this simulator for such an energy range.

On the other hand, due to the fact that the expected penetration is within the ultra-shallow range, the measurements require extremely high in-depth resolution which, until recently, could not be reached. Only recently, the progress made in medium energy ion spectroscopy (MEIS) allowed some unique experimental data to be obtained.

In [6], the results of MEIS measurements of nitrogen profile implanted into silicon at very low energies (200 eV, 500 eV and 1000 eV) may be found. This gave us a unique chance to compare the MEIS profiles with those obtained by means of TRIM simulations assuming the same ion energies. The results of theoretical calculations by TRIM for all three energy values (i.e., 200 eV, 500 eV and 1000 eV) are compared with the MEIS profiles in Fig. 5.

Looking carefully at this figure one may notice obvious differences between the calculated and measured ion profiles. The TRIM profiles are certainly deeper than those obtained by MEIS. The relative differences between the individual ion ranges obtained by both methods are, however, very similar.

The simulated and experimental profiles differ also by the total number of nitrogen ions located within the silicon substrate per implanted ion. The lower total densities of implanted ions received for MEIS profiles seem to prove that more ions than expected by TRIM simulations are un-

able to penetrate the silicon substrate. This can possibly be explained by more effective backscattering from the silicon surface than calculated by TRIM.

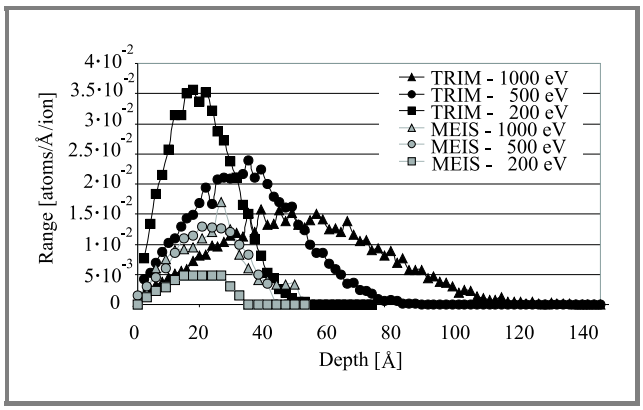


Fig. 5. Comparison of the nitrogen profiles in silicon determined experimentally by means of MEIS [6] and the results of TRIM simulation assuming the same energies of nitrogen ions (i.e., 200 eV, 500 eV and 1000 eV).

Despite the observed differences, it should be noted that in both cases the implanted nitrogen profile depths are of the same range (or higher) than the thickness of oxynitiride layers for state-of-the-art MOSFET gate stacks. This means, that ultra-shallow implantation of nitrogen from r.f. plasma will always play a very important role in the formation of ultra-thin oxynitride layers, especially when low thermal budget is at stake.

4. Conclusions

Within the scope of this work it has been proved that the process of nitrogen implantation takes place even in PECVD r.f. plasma. It influences the kinetics of oxynitride layer growth and the properties of the resulting layers. This ultra-shallow implantation affects the final oxide (oxynitride) layer thickness and its composition by three mechanisms. These are: damage of the substrate during ion implantation from r.f. plasma, formation of SiN bonds (in the case of nitrogen containing plasmas), and replacing the nitrogen in SiN bonds by oxygen during oxidation (even at low temperatures – 350°C). It should be also noted that all of these processes take place "dynamically" while consuming the substrate during oxidation.

The results of nintrogen implantation from nitrogen and ammonia plasmas are different. In the former case even short low-temperature oxidation results in the formation of almost pure oxide – the nitride phase is hardly seen there. The case is different for ammonia plasmas. We obtain true oxynitride layers, as the nitride-phase content may even amount to 30% for short oxidation times.

The depth range of the nitrogen implantation under low-energy conditions has been evaluated by means of TRIM simulations and compared to MEIS measurements found in [6]. Although the TRIM profiles exhibit deeper penetration range than that obtained by MEIS and TRIM seems to underestimate the backscattering from the silicon surface, in both cases the nitrogen implantation depth is higher than the thickness of oxynitride layers required in state-of-the-art CMOS ICs.

Acknowledgements

This work was supported by the State Committee for Scientific Research under project no. 8T11B 074 19.

References

[1] ITRS – International Technology Roadmap for Semiconductors, http://www.public.itrs.net

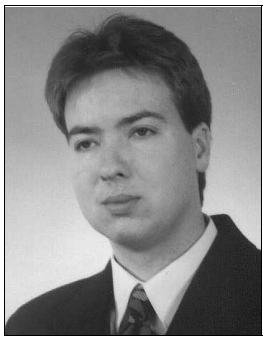
[2] J.-P. Carrere, A. Grouillet, F. Guyader, A. Beverina, M. Bidaud, and A. Halimaoui, "Triple gate oxide by nitrogen implantation integrated in a 0.13 μm CMOS flow", in *Proc. 32nd Eur. Solid-State Dev. Res. Conf. ESSDERC 2002*, Bologna, Italy, 2002, pp. 155–158.

[3] C.-C. Chen, M.-C. Yu , J.-Y. Cheng , M.-F. Wang , T.-L. Lee, S.-C. Chen, C.-H. Yu, M.-S. Liang, C.-H. Chen, C.-W. Yang, and Y.-K. Fang, "Characterization of plasma damage in plasma nitrided gate dielectrics for advanced CMOS dual gate oxide process", in *7th Int. Symp. Plas. Proc. Induc. Dam*., Santa Clara, USA, 2002; *Amer. Vac. Soc*., pp. 41–44, 2002.

[4] R. B. Beck, T. Bieniek, A. Wojtkiewicz, and A. Jakubowski, "Low temperature plasma oxidation of silicon – possibility of application in CMOS-ULSI technology", in "Advances in Plasma Chemistry" *Acta Agrophys*., vol. 80, pp. 247–256, 2002.

[5] J. F. Ziegler, http://www.srim.org

[6] E. P. Gusev, H. C. Lu, T. Gustafsson, E. Garunkel, M. L. Green, and D. Brasen, "The composition of ultra-thin oxynitrides thremally grown in NO", *J. Appl. Phys*., vol. 82, p. 896, 1997.



Tomasz Bieniek was born in Warsaw, Poland, in 1977. He received the M.Sc. degree from Warsaw University of Technology in 2002. He is currently working towards a Ph.D. degree. His research area concentrates on plasma processes for CMOS-VLSI technology.

e-mail: tbieniek@elka.pw.edu.pl
Institute of Microelectronics and Optoelectronics
Warsaw University of Technology
Koszykowa st 75
00-662 Warsaw, Poland

Romuald B. Beck graduated from the Faculty of Electronics, Warsaw University of Technology, in 1977 and joined the Institute of Microelectronics and Optoelectronics the same year. He received the Ph.D. and D.Sc. degrees in 1985 and 1996, respectively. His research activities are concentrated in the area of modeling, diagnostics and technology of metal-insulator-semiconductor structures, especially with very thin and ultra thin oxide. His work is focused on theoretical and experimental studies of different methods of dielectric-layer formation (including those involving DC and AC plasma) and their kinetics, as well as the yield, electrophysical properties and reliability of the resulting devices. His research interests include also dry etching methods and their implementation in modern IC and MEMS technology. He authored and co-authored more than 90 technical publications.

e-mail: beck@imio.pw.edu.pl
Institute of Microelectronics and Optoelectronics
Warsaw University of Technology
Koszykowa st 75
00-662 Warsaw, Poland

Andrzej Jakubowski received the M.Sc., Ph.D. and D.Sc. degrees from the Faculty of Electronics, Warsaw University of Technology. At present he is the Head of the Division of Microelectronics and Optoelectronics Devices in the Institute of Microelectronics and Optoelectronics (WUT). In the years 1989–1992 he held the position of the Director of the Institute of Electron Technology in Warsaw. From 1989 to 1991 he was the head of the Government Program for Polish Microelectronics. In the years 1990–1991 he headed Polish Government Committee on Applied Sciences. The main area of his research interests is modeling, characterization and technology of semiconductor structures and integrated circuits. He is also the author and co-author of approximately 400 technical papers.

e-mail: jakubowski@imio.pw.edu.pl
Institute of Microelectronics and Optoelectronics
Warsaw University of Technology
Koszykowa st 75
00-662 Warsaw, Poland

Andrzej Kudła – for biography, see this issue, p. 39.

Paper

Properties of Al contacts to Si surface exposed in the course of plasma etching of previously grown nanocrystalline c-BN film

Piotr Firek, Aleksander Werbowy, Jan Szmidt, and Andrzej R. Olszyna

Abstract—Properties of Al electric contacts to Si(p) surface exposed to fluorine-based plasma etching of nanocrystalline cubic boron nitride (c-BN) film grown previously were studied and compared to the properties of Al contacts fabricated on pristine or dry etched surface of Si(p) wafers. In addition, a part of the investigated samples was annealed in nitrogen atmosphere at the temperature of 673 K. Analysis of contract properties is based on current-voltage (*I-V*) measurements of the produced Al-Si structures. The presented investigations were performed in order to evaluate the efficiency of the applied plasma etching method of nanocrystalline c-BN from the viewpoint of its influence on the properties of metal contacts formed subsequently and thus on the performance of electronic devices involving the use of boron nitride.

Keywords*— *cubic boron nitride, plasma etching, electric contacts.

1. Introduction

Cubic boron nitride (c-BN) owing to its unique properties, such as wide bandgap, good thermal conductivity, high thermal stability and chemical resistance, is a promising material for numerous applications. These cover such areas as high-temperature and high-power electronics, UV (included deep UV) detection and light emitting devices or structures intended to work in chemically harsh environments [1–3].

Fabrication of metal-insulator-semiconductor (MIS) structures with boron nitride layer playing the role of an insulator is quite easy and offers a very useful tool for electronic characterization of a dielectric film (BN in this case) [4, 5]. MIS systems also form very convenient testing ground for verification of various processing technologies of the investigated insulating material, e.g., etching [5–7] or fabrication of metal contacts [8]. However, elaboration of the complete technological process leading to the fabrication of an advanced device making use of BN (e.g., MIS field-effect-transistor (MISFET), where boron nitride film acts as a gate insulator) is a much more demanding task. Among other things it requires the development of a selective etching technique of BN film, which allows desired features (gate area in this case) to be patterned on its surface. Due to the aforementioned chemical resistance, boron nitride is difficult to be removed by means of wet etching, typically used in electronics. Much more promising results are brought by experiments with dry plasma selective etching of BN [5–7]. However, in order to determine the real usability of the applied method, it is necessary not only to investigate the efficiency of material removal or selectivity of such a process, but also to examine the influence of the process on the characteristics of the final device. Plasma is an aggressive environment and comes into interaction with the surface (e.g., substrate) being exposed, undoubtedly affecting its state, even if the etch stop moment is selected correctly. In this work we have studied the properties of Al contacts to Si surface, which were formed after removal of nanocrystalline c-BN film grown previously. Any practical application requires the availability of contact fabrication method ensuring the contacts exhibit desired properties. This may also include the knowledge of how to bring such contacts into the right state if they do not posses the required parameters initially (immediately after formation). In the present study our goal was to obtain ohmic contacts to plasma exposed Si surface, as such terminals are necessary for, e.g., source and drain regions of the planned MISFET structures. Ohmic character is here considered primarily in the sense of small contact resistance with respect to the resistance of the substrate or device, rather than in terms of strict linearity of the observed current-voltage curves [9].

2. Experimental details

Nanocrystalline, 60–80 nm c-BN films were produced on p-type Si (<100>, $\rho = 6$–$8\ \Omega$cm) substrates using reactive pulse plasma (RPP) assisted CVD method [10]. Then BN layers were dry etched in fluorine based r.f. (13.56 MHz) plasma (PLASMALAB OXFORD 80+ station). Selectivity of etching was achieved by applying aluminum metallization as a mask. Etching process parameters are presented in Table 1. The results of etching were verified by ellipsometric measurements (GAERTNER 117, $\lambda = 632.8$ nm) as well as scanning electron microscopy (SEM). Subsequently, Al contacts were vacuum evaporated on such surfaces. In order to determine the potential influence of BN removal on electrical properties of Al contacts formed on

plasma-exposed Si surface, in parallel a comparative set of Al contacts was fabricated on pristine Si wafers (of the same type as above) that were not covered with boron nitride. Another set of contacts was evaporated on Si substrates that were initially plasma pre-etched (under the same conditions as shown in Table 1 except shorter – 0.5 min – process time). This step was intended to simulate the effect of the final stage of BN plasma etching of Si surface, when it becomes directly exposed to the plasma environment. Finally, a part of the obtained structures of all types was annealed at the temperature of 673 K for 30 min in nitrogen atmosphere. All wafers had Al backside contacts as well.

Table 1
Process parameters of nanocrystalline c-BN film etching

Parameters	Values
Gas mixture	CF_4 + Ar
Flow rates of gas mixture components [ml/min]; CF_4 to Ar	30 to 10
Power [W]	100
Pressure [mTr]	40
Etching time [min]	3 (0.5*)
* In the case of pristine Si surface pre-etching.	

Electrical properties of Al contacts were investigated by means of automatic current-voltage (*I-V*) measurements with Keithley SMU 238 within the range ±10 V with the current monitored every 0.1 V.

3. Results and discussion

Since ellipsometrical measurements of the thickness of nanocrystalline c-BN layers performed after completing the etching process gave ambiguous results (some readouts suggested 2.5–3 nm thick residue of boron nitride film), in addition SEM images of sample surface were obtained (see Figs. 1 and 2). They seem to indicate that the boron nitride layer was successfully removed, except some isolated islands of unetched material (separated on the average by a distance of 100 nm one from another), most likely composed of chemically resistant compounds like BC_x, BCN, C_xN_y, etc. (Fig. 1). Their local character suggests that they should have no significant influence on Al contact properties. However, even if BN layer residuum formed a continuous film, it should not be a problem since in the case of such a thin film quantum tunneling effects start playing a crucial role in carrier transport and thus that kind of interface cannot be a barrier for current flow. At the same time trenching effects were observed around the edges of Al metallization used as an etching mask, as well as erosion in the case of a particular edge (Fig. 2). These are phenomena characteristic of physical sputtering, which indicates most likely that the composition of reactant gases should be changed in favor of the component responsible for chemical etching (in our case CF_4 at the expense of Ar). This observation is supported by our previous studies concerning dry plasma etching of gallium nitride (GaN) by means of the same method, where improvement was observed if the content of chemically etching component was increased in the reactant gas mixture [11].

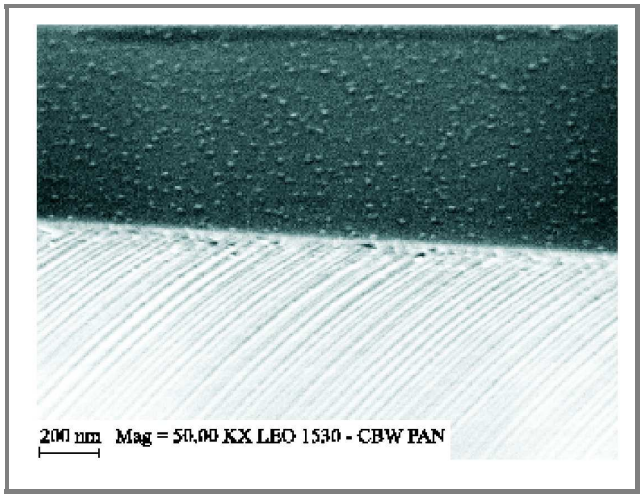


Fig. 1. SEM image of a fractured Si substrate showing the surface of the wafer exposed to plasma in the course of plasma etching of a nanocrystalline c-BN film grown previously.

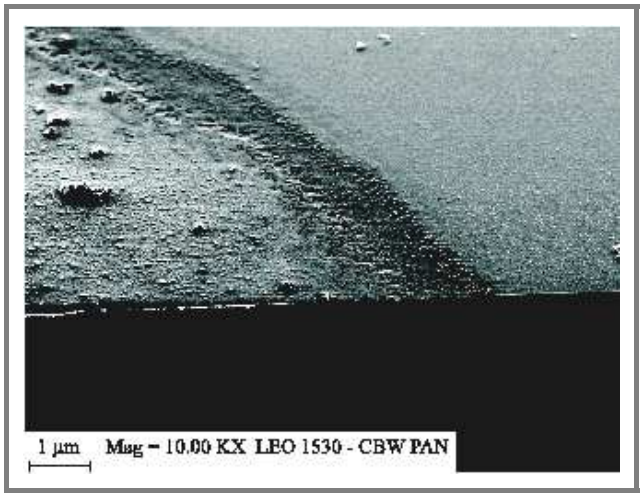


Fig. 2. SEM image of a fractured Si substrate with Al masking metallization after plasma etching process.

The results of *I-V* measurements of different sets of Al/Si structures are shown in Figs. 3–8.

Figure 3 presents current-voltage curves observed for Al contacts evaporated onto Si surface exposed to plasma during c-BN layer etching. As it can be seen, these characteristics are strongly asymmetric. However, the situation changes dramatically after annealing (673 K, N_2 atmosphere), as characteristics become symmetrical (Fig. 4). Now, if such curves are compared to the corresponding

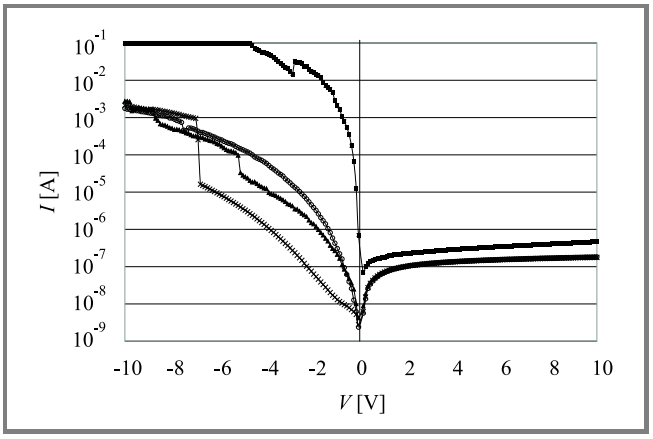


Fig. 3. *I-V* characteristics of Al contacts evaporated onto Si surface exposed to plasma during plasma etching of a nanocrystalline c-BN film grown previously.

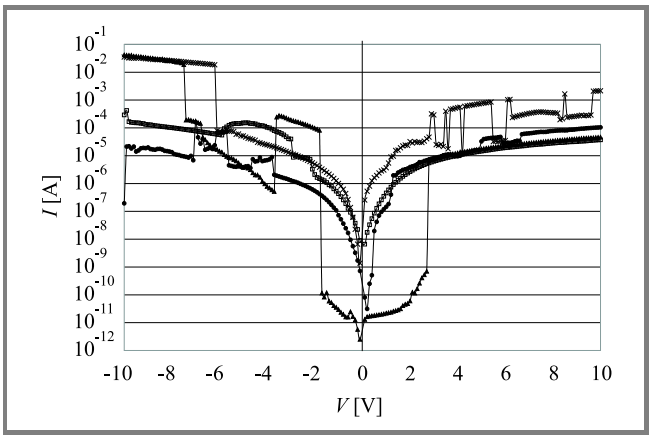


Fig. 4. *I-V* characteristics of Al contacts evaporated onto Si surface exposed to plasma during plasma etching of nanocrystalline c-BN film grown previously – results of contact annealing at 673 K in N_2 atmosphere.

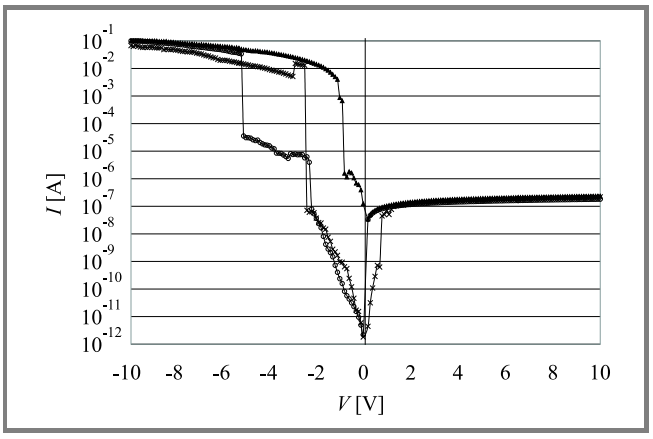


Fig. 5. *I-V* characteristics of Al contacts evaporated on pristine and briefly plasma etched Si surface.

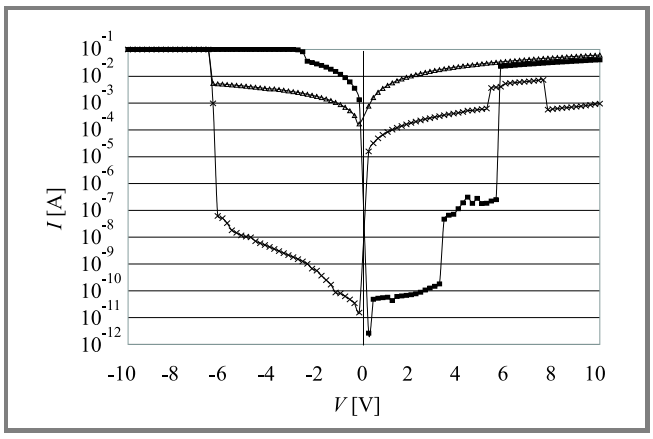


Fig. 6. *I-V* characteristics of Al contacts evaporated on pristine and shortly plasma-etched Si surface – results of contact annealing at 673 K in N_2 atmosphere.

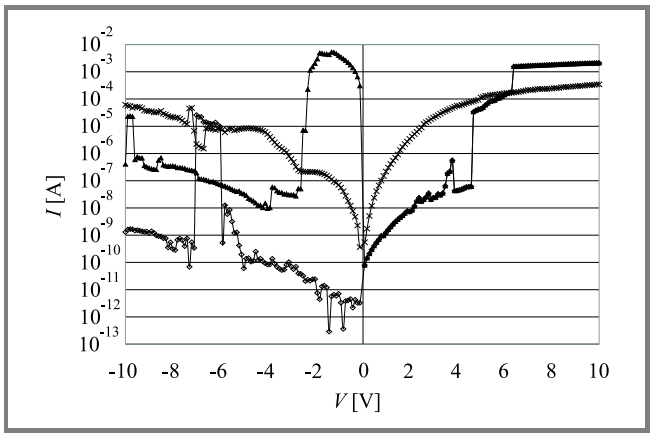


Fig. 7. *I-V* characteristics of Al contacts evaporated on pristine Si surface that was not subject to pre-evaporation plasma etching.

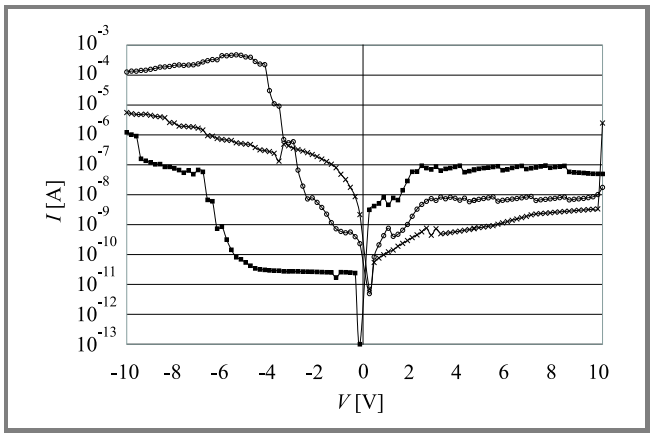


Fig. 8. *I-V* characteristics of Al contacts evaporated on pristine Si surface that was not subject to pre-evaporation plasma etching – results of contact annealing at 673 K in N_2 atmosphere.

characteristics obtained for contacts evaporated on pristine and plasma pre-etched Si substrates, it becomes evident that one has to deal with a similar trend. It is clear that non-annealed electrodes act as rectifiers (Fig. 5) but after annealing (Fig. 6), the characteristics of both structure types become symmetrical. Slightly lower current level in the case of contacts grown on Si surface truly exposed to plasma in the course of BN etching seems to be the obvious result of the remaining BN residues and other post-etching contaminants which effectively increase the contact resistivity. The similarity of the *I-V* characteristics of both sample sets seems also to be an indirect proof that plasma etching was successful since otherwise the obtained curves might be quite different.

Figures 7 and 8 present behavior of analogous *I-V* curves, obtained for Al contacts produced on pristine Si surfaces that were not exposed to plasma prior to Al evaporation. As it can be seen, they perform much worse, especially as far as their resistivity is concerned (definitely lower current levels). This seems to be in accordance with the well-known fact that better ohmic contacts are achieved if semiconductor surface is defected (resulting in the formation of high density efficient recombination centers causing significant decrease in the contact resistance, as well as generally promoting better metal adherence to the semiconductor) [9, 12]. Therefore annealing (Fig. 8) had apparently no influence on *I-V* curves observed for this set of samples.

4. Conclusions

The presented investigations show that fluorine based r.f. plasma etching of nanocrystalline c-BN films is an efficient way to selectively remove boron nitride and to pattern BN layers for the purposes of application in microelectronics (in this case – obtaining of gate dielectric in MISFET structures). A c-BN film may be removed from any unmasked substrate region areas. Studies of the properties of Al contacts to Si substrate regions exposed to plasma in the course of BN-layer etching prior to evaporation (and comparative studies of analogous structures evaporated on silicon surfaces never covered with BN) demonstrate the possibility of fabricating contacts that meet requirements for electrical ohmic terminals for semiconductor devices. Ohmic character means here rather a low contact resistance compared with the resistance of substrate or device and not essentially the linearity of its *I-V* characteristics. In other words, plasma processing does not affect the characteristics of Al electrodes that are to be formed later on in a negative sense.

Moreover, it has an even beneficial influence because of a certain level of surface damage that improves metal adhesion and introduces recombination centers decreasing the resistance of the metal-semiconductor system. As expected, annealing plays a substantial role in improvement of contacts fabricated in this way.

References

[1] P. J. Gielisse, "Wide bandgap materials in future electronic applications", in *Conf. IMAPS*, Rytro, Poland, 2000, www.cyf-kr.edu.pl/academic/OBRMHiR/imaps/data/txt/txt11.htm

[2] S. Noor Mohammad, "Electrical characteristics of thin film cubic boron nitride", *Solid-State Electron.*, vol. 46, p. 203, 2002.

[3] *High-Temperature Electronics*. R. Kirschmann, Ed., Piscataway: IEEE Press, 1999.

[4] A. Werbowy, J. Szmidt, A. Olszyna, A. Sokołowska, and P. Pawłowski, "Electronic properties of nanocrystalline BN and AlN films deposited on Si and GaAs – a comparison", *Diam. Rel. Mater.*, vol. 8, p. 391, 1999.

[5] A. Werbowy, P. Firek, J. Szmidt, A. Olszyna, and M. Gałązka, "Electric characterization and plasma etching of nanocrystalline c-BN layers", *J. Wide Bandgap Mater.*, vol. 9, p. 169, 2002.

[6] A. Werbowy, J. Szmidt, A. Sokołowska, and S. Mitura, "RF plasma selective etching of boron nitride films", *Diam. Rel. Mater.*, vol. 9, p. 609, 2000.

[7] M. Gałązka, J. Szmidt, and A. Werbowy, "The influence of plasma RF processes on nitride (BN, AlN, CN) layers", *J. Wide Bandgap Mater.*, vol. 9, p. 207, 2002.

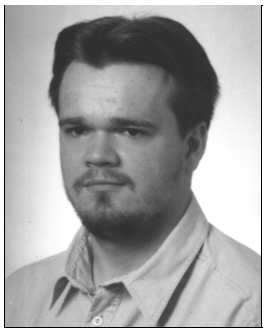
[8] A. Werbowy, J. Szmidt, A. Sokołowska, A. Olszyna, and S. Mitura, "Fabrication and properties of Mo contacts to amorphous cubic boron nitride (a-cBN) layers", *Diam. Rel. Mate*r., vol. 5, p. 1017, 1996.

[9] E. H. Rhoderick, *Metal-Semiconductor Contacts*. Oxford: Clarendon Press, 1978.

[10] M. Sokołowski, A. Sokołowska, Z. Romanowski, B. Gokieli, and M. Gajewska, "Properties and growth of BN (borazone) layers from a pulse plasma under a reduced pressure", *J. Cryst. Growth*, vol. 52, p. 165, 1981.

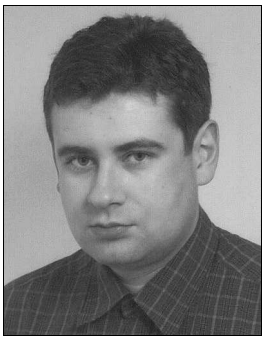
[11] A. Szczęsny, P. Śniecikowski, J. Szmidt, and A. Werbowy, "Reactive ion etching of novel materials – GaN and SiC", *Vacuum*, vol. 70, p. 249, 2003.

[12] A. G. Milnes and D. L. Feucht, *Heterojunctions and Metal-Semiconductor Junctions*. New York, London: Academic Press, 1972.



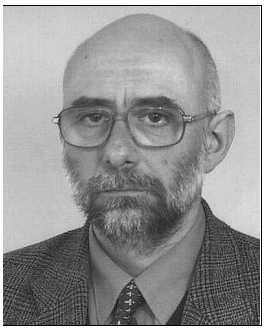
Piotr Firek was born in Rawa Mazowiecka, Poland, in 1977. He received the M.Sc. degree in microelectronics from the Faculty of Electronics and Information Technology, Warsaw University of Technology in 2004, where he is currently working towards a Ph.D. His research concentrates on fabrication, characterization, processing and application of wide bandgap (e.g., boron nitride, diamond) thin film materials in microelectronic devices.

e-mail: pfirek@elka.pw.edu.pl
Institute of Microelectronics and Optoelectronics
Warsaw University of Technology
Koszykowa st 75
00-662 Warsaw, Poland



Aleksander Werbowy received the M.Sc. and Ph.D. degrees in microelectronics from the Faculty of Electronics and Information Technology, Warsaw University of Technology in 1994 and 1999, respectively. His main research interests are synthesis, characterization, processing and application of thin films of wide bandgap materials in microelectronics and nanoelectronics.

e-mail: werbowy@imio.pw.edu.pl
Institute of Microelectronics and Optoelectronics
Warsaw University of Technology
Koszykowa st 75
00-662 Warsaw, Poland



Jan Szmidt received the M.Sc. degree in electronics from the Faculty of Electronics, Warsaw University of Technology in 1976. From the same university, he received the Ph.D. and D.Sc. degrees in 1984 and 1995, respectively. In 1999, he became an Associate Professor. Since 2002 he has been an Associate Dean for Development of the Faculty of Electronics and Information Technology, Warsaw University of Technology. His research interests concentrate on characterization of thin films for electronics, especially for microelectronics and nanoelectronics, as well as on their application in microelectronic and nanoelectronic structures.

e-mail: j.szmidt@elka.pw.edu.pl
Institute of Microelectronics and Optoelectronics
Warsaw University of Technology
Koszykowa st 75
00-662 Warsaw, Poland

Andrzej R. Olszyna is a Professor of the Faculty of Materials Science and Engineering at Warsaw University of Technology. His technical interests fall within the general areas of ceramics and their composites, and specifically within the subjects of synthesis of thin layers and bulk ceramic materials. He is a member of the Polish Materials Engineering Society (PTM).

e-mail: aolszyna@meil.pw.edu.pl
Faculty of Materials Science and Engineering
Warsaw University of Technology
Wołoska st 141
02-507 Warsaw, Poland

Paper

Application of scanning shear-force microscope for fabrication of nanostructures

Andrzej Sikora, Teodor Gotszalk, Anna Sankowska, and Ivo W. Rangelow

Abstract—In view of the rapid growth of interest in AFM technique in surface property investigation and local surface modification we describe here an AFM microscope with optical tip oscillation detection. The modular shear-force/tunneling microscope for surface topography measurement and nanoanodisation is described. The measurement instrument presented here is based on the fiber Fabry-Perot interferometer for the measurement of the conductive microtip oscillation that is used as nano *e*-beam for local surface anodisation. An advantage of this system is that quatitative measurements of tip vibration amplitude are easily performed.

Keywords*— *AFM, nanostructures fabrication, shear-force microscopy.

1. Introduction

Scanning force microscopy [1] is one of the many scanning probe techniques developed after the invention of scanning tunneling microscopy (STM) [2]. In atomic force microscopy the force interaction observed between the microtip mounted on the cantilever and the investigated surface is utilized to characterize the surface. Several measurement techniques can be applied to the detect of force interactions acting on the microtip. One of them is the shear-force microscopy (SHFM). In this technique a wire tapered either by means of electrochemical etching or by pulling is mounted perpendicularly to the sample. This tip is oscillating laterally to the surface by a few nanometers near to one of its mechanical resonant frequency. The oscillations are damped out by shear forces at a distance of a few nanometers from the surface. This effect may then be used to adjust the tip-sample distance. It may also be used as a basis for high-resolution topographic imaging. Accurate measurement of the true oscillation amplitude of the tip yields very important information that allows this tool to be used more confidently. Moreover, high measurement sensitivity is desired to enable the use as low oscillation amplitude as possible. This is because of two reasons: in order not to loose the lateral resolution of surface measurements and in order to extend the lifetime of the microtip.

In this paper optical tip oscillation detection setup will be presented that includes a Fabry-Perot optical fiber interferometer. The advantages of the presented setup are extreme sensitivity and compactness. Using the described measurement system, quantitative measurement of the probe dither motion with the resolution of 0.1 nm is possible in the bandwidth of 100 Hz. Moreover, optical detection system allows electrical voltage to be applied to the conductive microtip. In this case the microtip can be used as an electron beam (*e*-beam) source for nanolithography processes or as a collector of tunneling or field emission current flowing between the surface and the microprobe.

In parallel to fundamental studies [3–7], several nanometer-scale devices, such as single electron [8] and metal-oxide [9] transistors, quantum wires [10], high-density memories [3, 11], and Josephson junctions [12], have illustrated some of the nanoelectronic applications. Machining of silicon structures has also been demonstrated by AFM oxidation [13]. Furthermore, local oxidation lithography is compatible with the operation of parallel tip arrays [14]. This allows single-centimeter area patterning to be performed.

2. Instrumentation

In our setup, a single-mode optical fiber is fixed at the distance of a few microns from the reflecting surface, the deflection of which is monitored. The interference between the light reflected from the fiber-air interface and the backscattered light from the reflecting surface is monitored with a optical detector. In our interferometer design we use a single-mode pigtailed semiconductor laser as the light source. Since the light backscattered from any junction in the optical fiber system may disturb the operation of the laser diode, we use a solid-state optical Faraday isolator to ensure proper emission of the laser light. In our system we utilize an isolator built into the laser diode, which accurately fits the wavelength of the light source and is insensitive to mechanical vibrations. One of the fiber ends of the bidirectional coupler is connected directly with the optical isolator. The wire with the microtip is mounted close to the well cleaved fiber end of the coupler. The signal photodiode (Fig. 1) detects the interference of the light reflected from the interferometer fiber end and from the microscope wire.

To obtain high resolution and high fringe visibility the interferometer light source is supplied using a high-precision current source built specifically for this purpose, which ensures the stability of the current supply in the range of 50 ppm. The temperature of the laser is controlled by a Peltier cooler with the stability of 10 mK. We tested the interferometer sensitivity by changing the distance between

the wire and the interferometer fiber and simultaneously recording the interferometric fringes. If the amplitude of the fiber movement is bigger than $\lambda/4$ we observe the pattern shown in the Fig. 2.

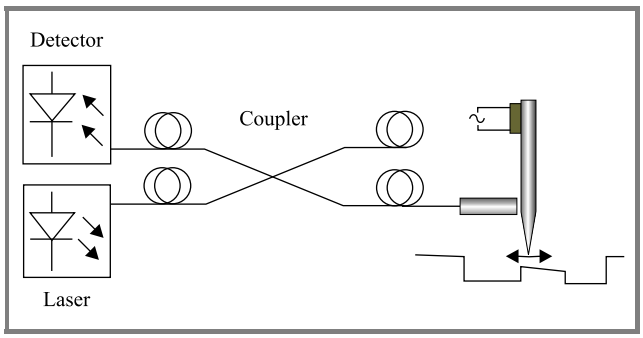


Fig. 1. Tip oscillation detection setup – optical fiber interferometer diagram.

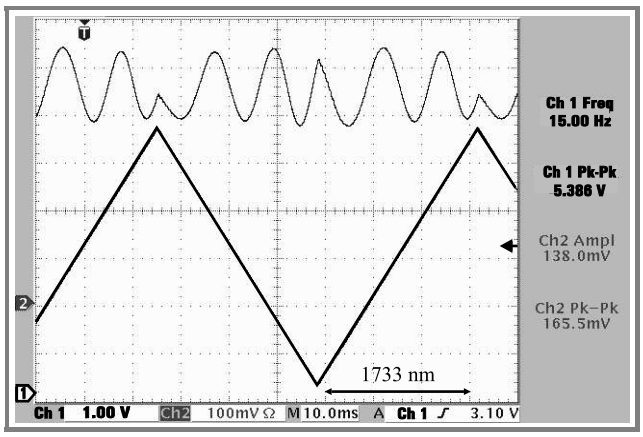


Fig. 2. Interferometric signal (thin line) versus tip movement (thick line).

The sensitivity of the developed interferometer may be obtained from the slope of the recorded signal – in our design we obtain the sensitivity 2 mV/nm. It should be noted that the distance between the interferometer fiber and the wire with the microtip must be adjusted using a piezoactuator, so that the working point is placed on the linear part of the slope. Our instrument reaches the resolution of 0.1 nm over a 100 Hz bandwidth.

The scanning tip is connected to a I/U converter built specifically for this purpose. The converter is placed next to the measuring head. Dielectric-insulation precision operational amplifier OPA111 was used as the main part of the converter. The I/U converter is connected to the amplifying/biasing module which may be controlled either remotely (scanning software) or manually.

The process controlling procedure is an integral part of the Topo-Scan program, which is used for scanning process control. The procedure was developed at the Laboratory of Scanning Probe Microscopy, Nanostructures and Nanometrology. The process data are placed in a .txt file and may be easily modified using, i.e., a standard word processor. The following parameters are available: X, Y are the start/stop point coordinates, t_1 – in-point waiting time, t_2 – tip moving time-constant, and U – voltage applied to the tip.

3. Experiment

In our experiments we used tungsten wires etched electrochemically (Fig. 3). The diameter of the tungsten tip was 120 μm and the length varied from 5 to 7 mm, which corresponds to the wire spring constant ranging from 1 N/m to 3 N/m.

Fig. 3. SEM photo of a tungsten tip.

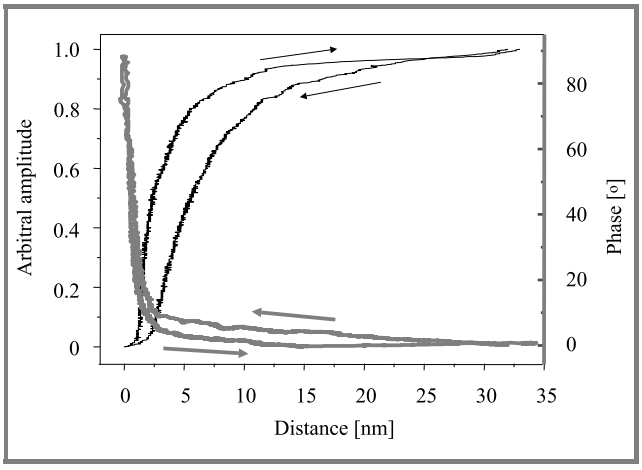


Fig. 4. Approach curves of the shear-force microscope.

The application of weaker wires enables detection of smaller forces but the scanning process is more time consuming and sensitive to acoustical disturbances.

The tip-sample distance is controlled using the shear-force detection method. When the tip approaches the surface-oscillation amplitude decreases and phase shift appears (Fig. 4). Highly accurate control of the tip-sample distance is possible when lock-in signal detection from the interferometer is used.

The emission-current module was tested using a sample of gold deposited on a glass substrate. The emission curve was measured for both bias polarizations. There is a clearly visible difference between the tip-emission surface-emission curves (Fig. 5). This setup offers a possibility to investigate local electrical material properties and correlate them with the surface topography [15].

Local oxidation was performed in the constant voltage mode. In contrast to the pulse mode – in the presented case

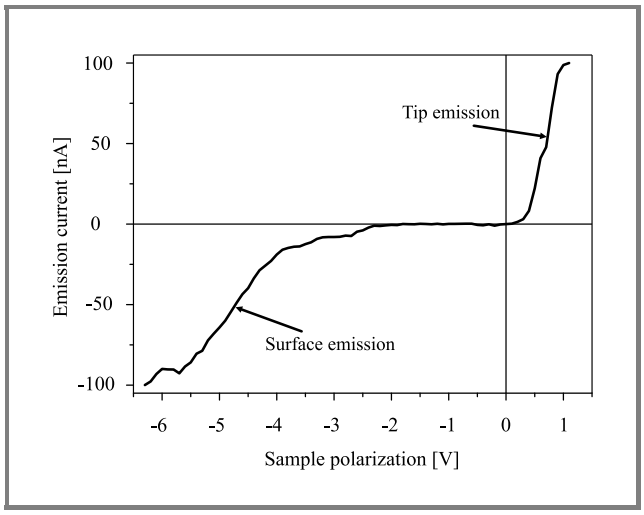


Fig. 5. Field emission current measured on gold film versus tip-sample voltage.

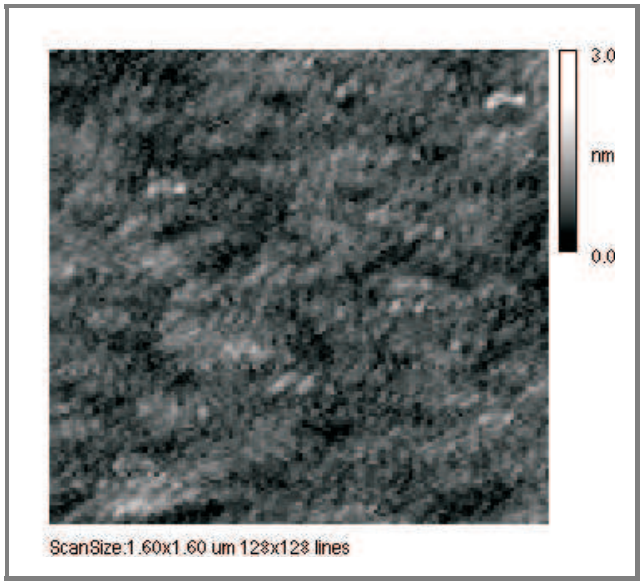


Fig. 6. Surface before process.

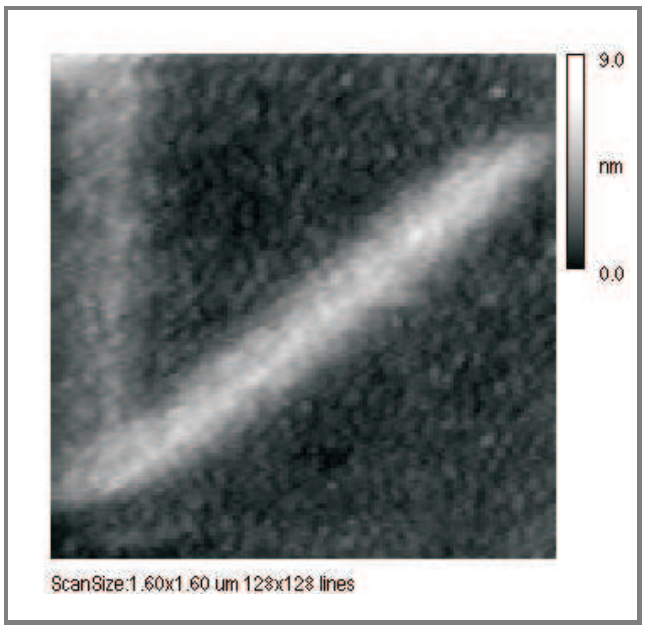


Fig. 7. Surface after process (Step 1).

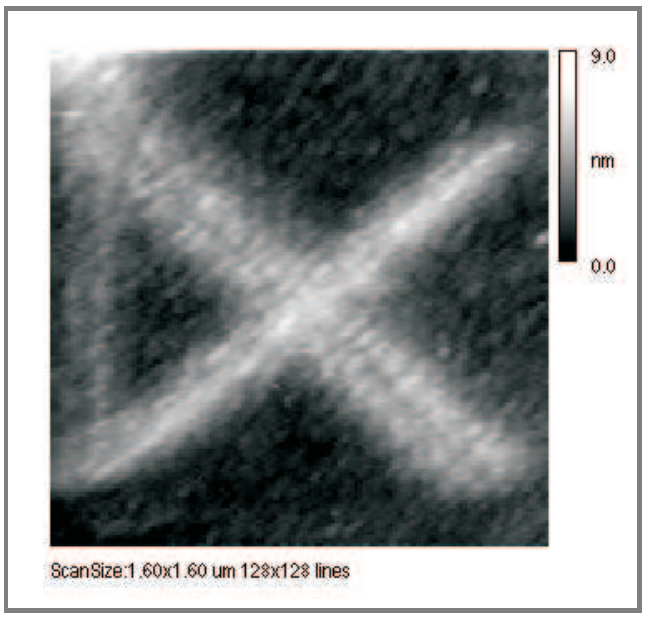


Fig. 8. Surface after process (Step 2).

the voltage was applied and the scanning tip was moved between the start and stop points.

Silicon surface was used for local anodisation. The process was performed in ambient conditions (air temp. 21°C and RH 42%).

The experiment was performed in five steps:

- scanning topography before the process (Fig. 6);
- process 1 (two lines: vertical and slanted line speed: 0.5 μm/s and 1.5 μm/s respectively, applied voltage 5 V);
- scanning topography after 1st process step (Fig. 7);

– process 1 (two parallel, slanted lines speed: 1.5 μm/s, applied voltage 5 V);

– scanning topography after 2nd process step (Fig. 8).

The obtained structures are typically 160 nm full width at half maximum (FWHM) and 5 nm high.

4. Summary

In this paper we presented experiments concerning the development and applications of SHFM based method for fabrication of nanostructures.
The major advantages of the described setup are low cost, possibility of handling very soft samples, and investigations of electrical surface parameters using the conductive tip.

Acknowledgments

We would like to thank our collaborators – Roman Szeloch, Jacek Radojewski, Paweł Janus, Robert Pędrak, Andrzej Marendziak, and Piotr Czarnecki – for fruithfull discussion and support. This work was partly supported by the project Wrocław University of Technology (project no. 342889).

References

[1] K. Karrai and R. D. Grober, "Piezoelectric tip-sample distance control for near field optical microscopes", *Appl. Phys. Lett.*, vol. 66, p. 1842, 1995.

[2] G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, "Tunneling through a controllable vacuum gap", *Appl. Phys. Lett.*, vol. 40, p. 178, 1982.

[3] R. Garcia, M. Calleja, and H. Rohrer J., "Patterning of silicon surfaces with noncontact atomic force microscopy: field-induced formation of nanometer-size water bridges", *Appl. Phys*., vol. 86, p. 1898, 1999.

[4] J. A. Dagata, T. Inoue, J. Itoh, K. Matsumoto, and H. Yokoyama, "Role of space charge in scanned probe oxidation", *J. Appl. Phys.*, vol. 84, p. 6891, 1998.

[5] M. Calleja and R. Garcia, "Nano-oxidation of silicon surfaces by noncontact atomic-force microscopy: size dependence on voltage and pulse duration", *Appl. Phys. Lett*., vol. 76, p. 3427, 2000.

[6] M. Tello and R. Garcia, "Nano-oxidation of silicon surfaces: comparison of noncontact and contact atomic-force microscopy methods", *Appl. Phys. Lett.*, vol. 79, p. 424, 2001.

[7] J. A. Dagata, F. Perez-Murano, G. Abadal, K. Morimoto, T. Inoue, J. Itoh, and H. Yokoyama, "Predictive model for scanned probe oxidation kinetics", *Appl. Phys. Lett.*, vol. 76, p. 2710, 2000.

[8] K. Matsumoto, Y. Gotoh, T. Maeda, J. A. Dagata, and J. S. Harris, "Room-temperature single-electron memory made by pulse-mode atomic force microscopy nano-oxidation process on atomically flat α-alumina substrate", *Appl. Phys. Lett.*, vol. 76, p. 239, 2000.

[9] E. S. Snow, P. M. Campbell, F. A. Buot, D. Park, C. R. K. Marrian, and R. Magno, "A metal/oxide tunneling transistor", *Appl. Phys. Lett.*, vol. 72, p. 3071, 1998.

[10] R. Held, T. Vancura, T. Heinzel, K. Ensslin, M. Holland, and W. Wegscheider, "In-plane gates and nanostructures fabricated by direct oxidation of semiconductor heterostructures with an atomic force microscope", *Appl. Phys. Lett.*, vol. 73, p. 262, 1998.

[11] E. B. Cooper, S. R. Manalis, H. Fang, H. Dai, S. C. Minne, T. Hunt, and C. F. Quate, "Terabit-per-square-inch data storage with the atomic force microscope", *Appl. Phys. Lett.*, vol. 75, p. 3566, 1999.

[12] V. Bouchiat, M. Faucher, C. Thirion, W. Wernsdorfer, T. Fournier, and B. Pannetier, "Josephson junctions and superconducting quantum interference devices made by local oxidation of niobium ultrathin films", *Appl. Phys. Lett.*, vol. 79, p. 123, 2001.

[13] F. S. Chien, C.-L. Wu, Y.-C. Chou, T. T. Chen, S. Gwo, and W.-F. Hsieh, "Nanomachining of (110)-oriented silicon by scanning probe lithography and anisotropic wet etching", *Appl. Phys. Lett.*, vol. 75, p. 2429, 1999.

[14] S. C. Minne, J. D. Adams, G. Yaralioglu, S. R. Manalis, A. Atalar, and C. F. Quate, "Centimeter scale atomic force microscope imaging and lithography", *Appl. Phys. Lett.*, vol. 73, p. 1742, 1998.

[15] A. Sikora, T. Gotszalk, R. Szeloch, A. Sankowska, and A. Marendziak "Modularny mikroskop tunelowy i sił atomowych do badań własności elektrycznych nanostruktur", in *Krajowa Konferencja Elektroniki*, Kołobrzeg, Poland, 2003 (in Polish).

Andrzej Sikora was born in Jawor, Poland, in 1975. He graduated from Electronic Faculty of Wrocław University of Technology in 2000. Since 2000 he has been working at the Faculty of Microsystem Electronics and Photonics (WUT) as a Ph.D. student. His research interests include scanning probe microscopy.
e-mail: andrzej.sikora@pwr.wroc.pl
Faculty of Microsystem Electronics and Photonics
Wrocław University of Technology
Janiszewskiego st 11/17
50-372 Wrocław, Poland

Anna Sankowska was born in Wrocław. In 1976 she received the M.Sc. degree from the Institute of Physics, Faculty of Fundamental Problems of Technology of Wrocław University of Technology. In 1980 she received the Ph.D. degree at the same faculty. Currently he is working in the laboratory of optoelectronics and fiber optics laboratory at the Faculty of Microsystem Technology at the Wrocław University of Technology. Her research interests include optoelectronics and fiber optic techniques.
e-mail: anna.sankowska@pwr.wroc.pl
Faculty of Microsystem Electronics and Photonics
Wrocław University of Technology
Janiszewskiego st 11/17
50-372 Wrocław, Poland

Teodor Paweł Gotszalk – for biography, see this issue, p. 44.

Ivo W. Rangelow – for biography, see this issue, p. 46.

Paper

TSSOI as an efficient tool for diagnostics of SOI technology in Institute of Electron Technology

Mateusz Barański, Krzysztof Domański, Piotr Grabiec, Mirosław Grodner, Bohdan Jaroszewicz, Andrzej Kociubiński, Wojciech Kucewicz, Krzysztof Kucharski, Jacek Marczewski, Halina Niemiec, Maria Sapor, and Daniel Tomaszewski

Abstract—This paper reports a test structure for characterization of a new technology combining a standard CMOS process with pixel detector manufacturing technique. These processes are combined on a single thick-film SOI wafer. Preliminary results of the measurements performed on both MOS SOI transistors and dedicated SOI test structures are described in detail.

Keywords— SOI CMOS technology, pixel detector, test structure.

1. Introduction

In typical silicon-on-insulator (SOI) technologies a thick substrate acts only as a mechanical support for the active silicon film. However in the specific application concerning the SUCIMA (Silicon Ultra Fast Camera for Electron and Gamma Sources in Medical Applications) project [1], the low-doped substrate detects ionized particles. This sensor is monolithically coupled to the readout electronics manufactured in the SOI device layer over the BOX (buried oxide) layer (Fig. 1). This is a solution unique in the world.

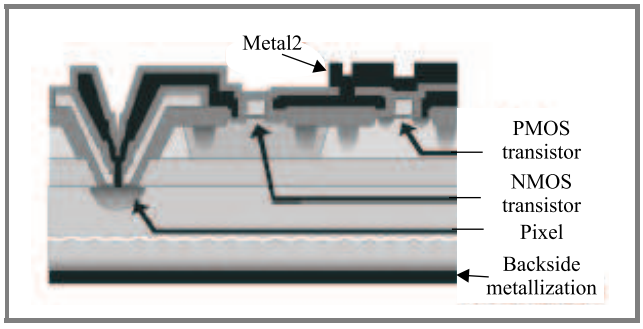


Fig. 1. A cross-section of a SOI pixel sensor.

So far sensors of ionizing radiation have not been fabricated on SOI wafers.

This challenging task requires a new technology of silicon processing at both sides of BOX layer [2, 3]. Interactions between these processes must be carefully taken into account. The following critical requirements for a sequence of individual operations can be mentioned:

- dielectric layers, doping profiles and junctions depths in the silicon body must be optimized for proper operation of read-out electronics;
- p^+-n pixel junctions should be shallow, because low-energy beta radiation must be detected;
- high quality of the substrate silicon, i.e. minority carrier life-time on the order of milliseconds, must not be degraded;
- metallization paths between pixel junctions (in deep cavities) and read-out electronics must be reliable.

A TSSOI test structure has been prepared and fabricated for development and characterization of the new technology, as well as for validation of the design. In this paper the individual elements implemented in the TSSOI structure will be described. Preliminary results of measurements of the test elements will be discussed.

2. Test structure specification

The TSSOI structure consists of two parts, namely a process-characterization substructure and a functional substructure. The process-characterization substructure of the TSSOI chip contains standard devices like resistors, capacitors, diodes, transistors and chains of contacts. Moreover, several special devices have been implemented. These are MOSFET arrays for threshold voltage variation measurements, and devices for the investigation of the influence of body contact configuration (layout) on the quality of the bias applied to the body of SOI MOSFETs. The functional substructure of the TSSOI structure contains sets of dedicated MOS transistors and numerous special devices. These are current mirrors dedicated to the mismatch measurements, devices dedicated to the capacitor mismatch measurements, specialized amplifying stages and read-out matrices. The two substructures of the TSSOI chip mentioned above are described in the next section.

2.1. Test elements for process characterization

The process-characterization substructure of the TSSOI chip was developed for the following purposes:

– extraction of parameters of SOI MOSFETs and other device models for SPICE-type simulator;

– characterization of SOI CMOS process;

– monitoring of quality and reliability of Metal1 and Metal2 connecting paths.

This substructure covers an area of 6.5×4.0 mm^2. It consists of 500×800 μm^2 modules, arranged in 5 rows and 13 columns. A 2×4 probe pad array is used throughout the chip for the purpose of automatic and/or manual testing. The size of several test modules exceeds the standard area of 500×800 μm^2. There are 46 probe pad arrays (i.e., 46 test modules) in the process-characterization vehicle.
The process-characterization substructure contains transistors with very long, very wide and standard-size channels, sets of diffused p$^+$-n junctions, poly-Si and p well resistors, capacitors with different area/perimeter ratios and different dielectric layers. A more detailed description of these modules is presented below.

2.1.1. Standard MOSFET arrays

These modules are illustrated in Fig. 2. There are two arrays of the so-called medium-size n- and p-channel MOSFETs. There are also another two arrays containing standard-size n- and p-channel devices, as well as devices with dimensions more critical from the point of view of the Institute of Electron Technology (IET) line.
The first set of wide-enough and long-enough transistors is used to extract the parameters of the DC models of standard wide- and long-channel transistors (threshold voltage VTO, substrate doping concentration NSUB, body factor GAMMA, Fermi voltage PHI) [4]. This set may also be used for the estimation of the variations of device dimensions *LD* (total channel shortening) and *DW* (channel narrowing) due to the limitations of the fabrication process (e.g., lithography). Standard DC parameters are extracted from transistor *I-V* characteristics, whereas *LD* and *DW* parameters are extracted through a comparison of *I-V* data obtained from devices with different dimensions.
Standard-size devices found in the second array are used for the extraction of DC model parameters of shorter transistors that are to be used in read-out electronics of the pixel detector in the SUCIMA project. These parameters include carriers mobility parameters (UO, UEXP, UCRIT), saturation velocity of carriers VMAX and several other parameters describing second-order effects. These parameters are estimated using the *I-V* characteristics of standard-size transistors [4]. The second array contains also devices of critical sizes. Electrical characteristics of these devices are

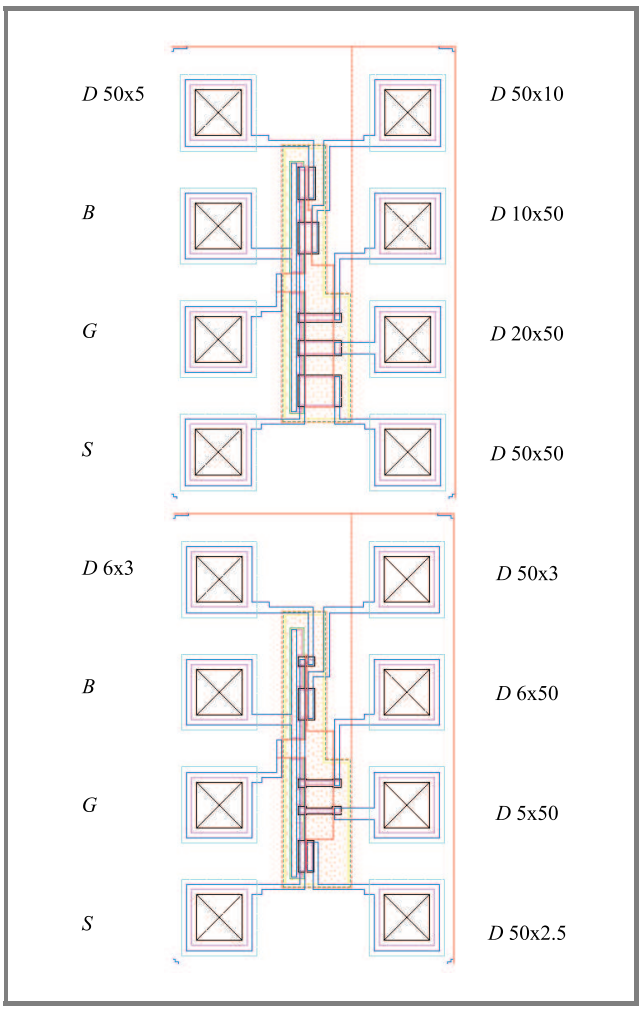


Fig. 2. A layout of n-channel MOSFET arrays used to extract DC model parameters; numbers following the drain pad symbol are channel dimensions (in micrometers).

used mainly for the improvement of devices performance in the presence of strong electric field.

2.1.2. Arrays of very wide and very long MOSFETs

The layout of the array of very wide MOSFETs is presented in Fig. 3. These transistors are used mainly for the extraction of the fringing gate-source (CGSO) and gate-drain (CGDO) capacitances. Moreover, they may be used for the estimation of gate oxide thickness (TOX) and/or gate oxide capacitance per unit area (COX). Also the channel shortening *LD* can be estimated using *C-V* data of these devices. The CGSO and CGDO parameters are extracted from the *C-V* characteristics in accumulation. The procedure of the extraction of the parameters mentioned above is based on a comparison of *C-V* data obtained from devices with different gates widths [5].
Very long MOSFETs have no source and drain areas, therefore they are capacitors. They are used to extract the fringing gate-substrate (CGBO) capacitance. Moreover, they

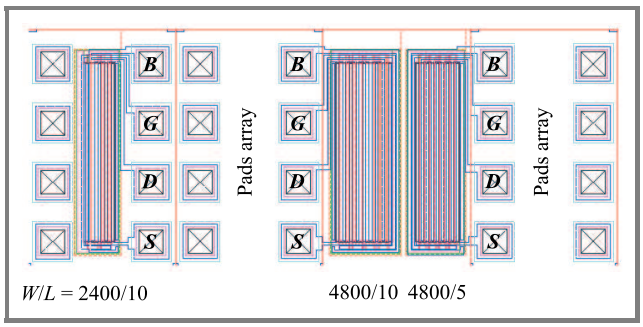


Fig. 3. A layout of wide n-channel MOSFETs array used for extraction of fringing capacitances CGSO, CGDO; the layout of p-channel MOSFETs is analogous.

may be used to estimate the values of TOX and channel narrowing *DW*. The CGBO parameter may be extracted from the *C-V* characteristics in inversion. COX and *DW* parameters may be obtained through a comparison of *C-V* data in accumulation and weak inversion obtained from devices with different dimensions. The extraction procedure is based on a comparison of *C-V* data obtained from devices with different gate lengths.

2.1.3. MOSFET arrays for the estimation of threshold voltage variations

The TSSOI structure contains also a set of twelve modules of n- and p-channel MOSFETs for the estimation of the threshold voltage mismatch. It is important to keep the value of this mismatch low, particularly for analog applications of MOS transistors. Each of these modules

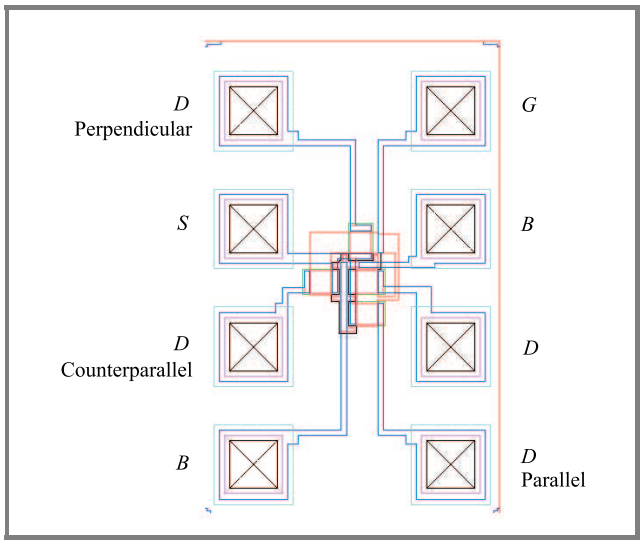


Fig. 4. The layout of one cell of p-channel MOSFETs for estimation of threshold voltage mismatch and its dependence on gate area.

contains four transistors of the same size placed very close to one another and arranged in parallel, counterparallel and perpendicular pairs (Fig. 4).

2.1.4. Cells for the estimation of the quality of the bias applied to the body of a SOI MOSFET

In the case of partially-depleted and thick-film SOI MOSFETs parasitic effects like "kink-effect" may be avoided if the silicon body is not floating. This solution requires an additional body contact. It may be expected that the position of this contact relative to the active part of the body may be relevant for MOSFET operation. This may be crucial particularly in the case of analog applications.

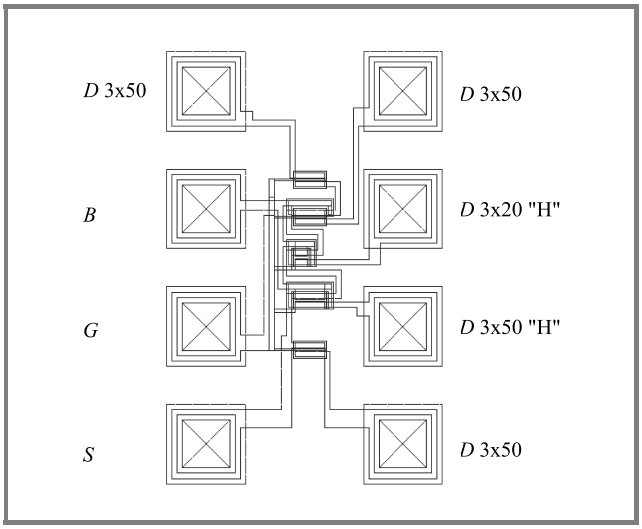


Fig. 5. The layout of one cell of p-channel MOSFETs for the optimization of body bias.

In order to investigate this problem two dedicated modules are included in the TSSOI structure (Fig. 5). They contain 5 n-channel MOSFETs with different p-well (body) contact configurations, as well as 5 p-channel MOSFETs with similar body contact configurations. The gates of two transistors are H-shaped. This will help to estimate the influence of parasitic edge transistors on the device characteristics.

2.1.5. Resistor modules for the extraction of the parameters of resistive paths

The TSSOI structure includes three modules containing sets of n^+-diffused, p^+-diffused and p-well resistors. These modules enable the parameters of resistive paths to be extracted, such as sheet resistance R_s, contact resistance R_c and width narrowing *DW*. These parameters can be extracted through a comparison of the resistance of resistors with different dimensions.

2.1.6. Diode modules for the extraction of p-n junction capacitance and leakage current parameters

Junctions p-n are extremely relevant for the operation of CMOS circuits. They strongly influence MOSFET operation in weak-inversion, as well as small-signal and transient operation. Thus the extraction of leakage currents

and capacitances of p-n junctions is important. The TSSOI structure contains three modules consisting of square and rectangle p well/n substrate, n^+/p-well and p^+/n-substrate diodes. The layout of the modules is illustrated in Fig. 6.

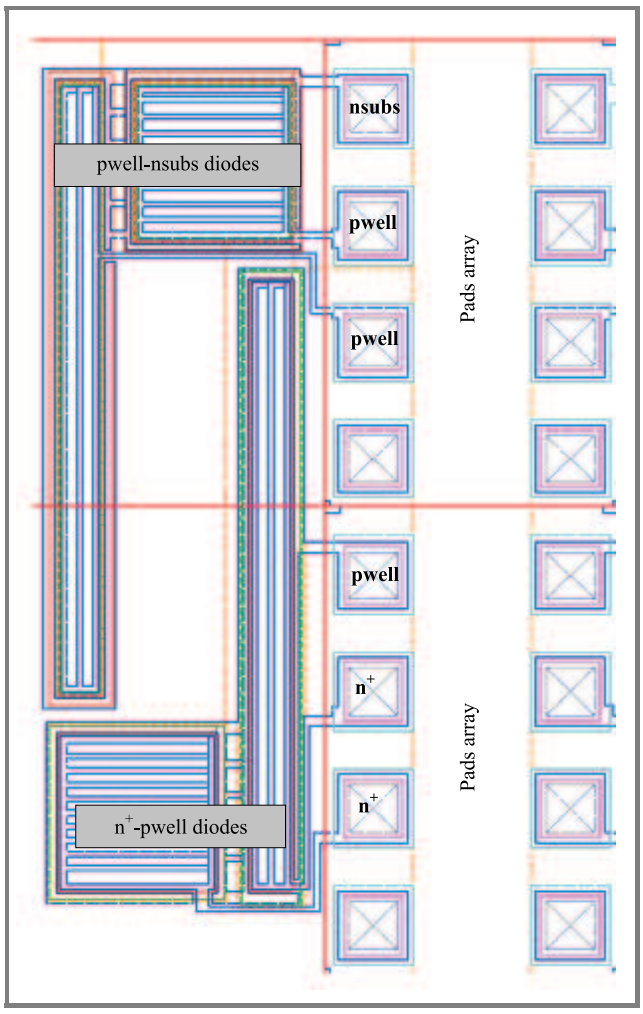


Fig. 6. A layout of modules containing different diodes with two area/perimeter ratios ($W \times L = 260 \times 260$, 65×1040 μm^2); they may be used to extract p-n junction capacitances and the volume and surface components of junction leakage.

A comparison of *I-V* and/or *C-V* characteristics of diodes with different dimension allows bulk and edge components of the considered parameter to be extracted. This may be very useful for the characterization and improvement of technology. It is also necessary for the proper extraction of p-n junction capacitance parameters for the SPICE simulator (CJ, MJ, PB, CJSW, MJSW).

2.1.7. Capacitor modules for the extraction of interlayer capacitances

The TSSOI structure contains a set of modules for the characterization of dielectric layers. The following capacitors (layout is the same as in the case of p-n junctions – see Fig. 6) are included in the structure:

- poly-Si/FOX/n-substrate, poly-Si/FOX/p-well,
- poly-Si/TOX/n-substrate, poly-Si/TOX/p-well,
- Metal1/FOX/n-substrate, Metal1/FOX/p-well,
- Metal1/DOX/poly-Si, Metal1/DOX/n-substrate,
- Metal2/DOX/Metal1.

where: TOX – thin (gate) oxide, FOX – field oxide, DOX – deposited (passivation) oxide.

The comparison of *C-V* data for capacitors of different dimensions allows vertical and lateral components of capacitances to be extracted. This may be important for the estimation of dielectric layer quality.

2.1.8. Modules for monitoring reliability of Metal1 and Metal2 conducting lines

The TSSOI structure contains a set of modules for monitoring of levels 1 and 2 (Metal1 and Metal2, respectively) of conducting metal lines, as well as vias between them. These modules consist of the following structures:

- a series of crossovers of Metal1 strip on poly-Si strips,
- a series of crossovers of Metal2 strip on Metal1and poly-Si strips,
- a series of double crossovers of Metal2 strip on Metal1 and poly-Si strip pairs drawn in accordance or violation of the design rules,
- a series of crossovers of Metal2 on Metal1 and poly-Si strip sandwich,
- a chain of VIA contact windows,
- chains of Metal1/p^+ and Metal1/n^+ contact windows,
- chains of Metal1/poly-Si(p^+) and Metal1/poly-Si(n^+) contact windows.

2.1.9. Modules for characterization of pixel p-n junctions and monitoring of contact reliability

The TSSOI structure contains modules for characterization of detector p-n junctions and for monitoring the reliability of contacts to these junctions. These modules are very important, when p-n junctions are to be formed in high-resistivity substrates below buried oxide layer, in deep cavities (see Fig. 1). Pixel properties and reliability of the connections are extremely important for detector operation. The structure contains the following modules for monitoring of contact reliability:

- a chain of contact windows for pixel detectors,
- a Metal1 serpentine over extremely deep detector contact windows.

These test elements as well as the elements mentioned above are very important as a tool for the diagnostics of conducting path reliability in the IET laboratory.
The structure contains also a module for the extraction of pixel p-n junction parameters in deep cavities. The module consists of 2 detector diodes of different area/perimeter ratios. A comparison of the *I-V* characteristics of these diodes is necessary for the extraction of lateral and vertical components of p-n junction current and thus for pixel characterization.

2.2. Test elements for verification of front-end electronics operation (functional substructure)

The functional test structure contains numerous elements. The main part consists of functional blocks that are described below. Among other devices they include several modules with individual MOSFETs (also with different body contact layout).

2.2.1. Blocks for mismatch measurements

The test structure includes two modules that are dedicated to the measurements of mismatch in the devices fabricated using SOI technology. The first one is a large set of PMOS and NMOS current mirrors (Fig. 7). In this unit some of the current mirrors have the transistors with the dimensions nominally exactly the same. As the consequence, the threshold voltage mismatch of the transistors can be evaluated by measuring the output currents I_{out} for a fixed value of the injected current I_{inj}. Additionally, this unit makes it possible to compare the output currents from transistors with the same W/L ratio but different dimensions and to observe the short channel effect in current sources.

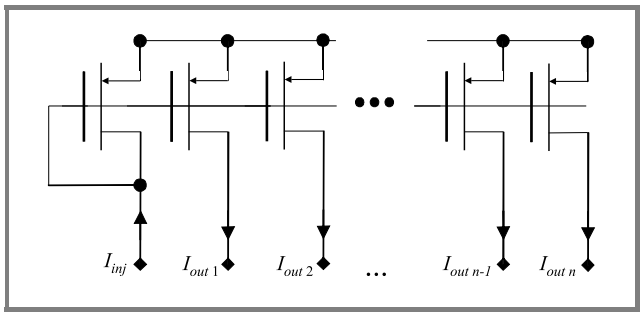


Fig. 7. Threshold voltage mismatch measurement in current mirrors unit.

The next module is dedicated to the measurements of capacitance mismatch. The designed test circuit allowing capacitance ratios to be measured consists of a couple of two capacitors (C_X and C_Y) that are nominally identical and a source follower.
The principle of capacitance mismatch measurements using this circuit is presented in Fig. 8. First the output signal is measured at the X pin connected to the signal generator and the Y pin grounded. The obtained value S_X is proportional to $C_X/(C_X+C_Y+C_p)$, where C_p is a parasitic capacitance. Next the same procedure is performed while the Y pin is connected to the signal generator and the X pin is grounded. This time the output signal value S_Y is proportional to $C_Y/(C_X+C_Y+C_p)$. Hence the ratio of S_X over S_Y signals is equal to the ratio of C_X over C_Y capacitances and does not depend on the input parasitic capacitance of the source follower.

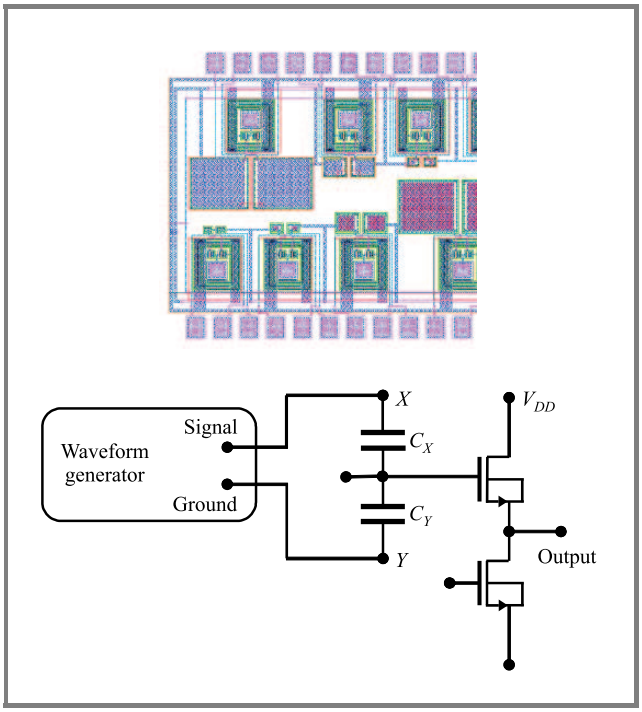


Fig. 8. Module and principle of capacitance mismatch measurements.

The next module of the test structure – two matrices of precise ratio capacitors – may be measured in a similar way. Apart from capacitance mismatch measurements, two different kinds of capacitors (Metal1-Metal2 and polySi-Metal1) may be compared using this module.

2.2.2. Blocks of amplifying stages

The block includes three simple amplifying stages: two common source OS stages and one cascode OS-OG stage. The layouts and schematics of those circuits are presented in Fig. 9. This part of the test chip not only allows the behavior of the new technology to be observed in an analog application, but also makes it possible to compare the measurement results with simulations.

2.2.3. Blocks with digital cells

Several basic digital gates and data flip-flops with asynchronous reset are implemented in the test structure (Fig. 10). Such important parameters, as rising and falling

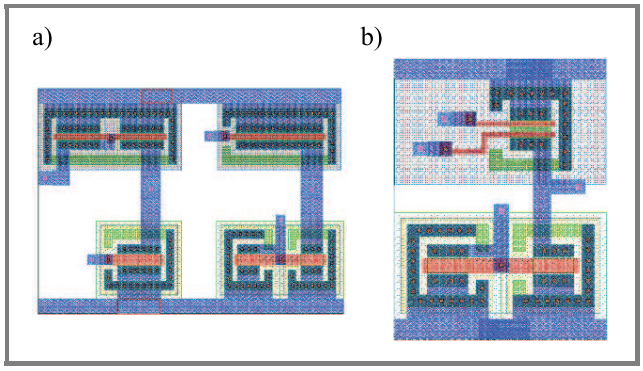


Fig. 9. Schematic layouts of amplifying stages: (a) OS; (b) OS-OG.

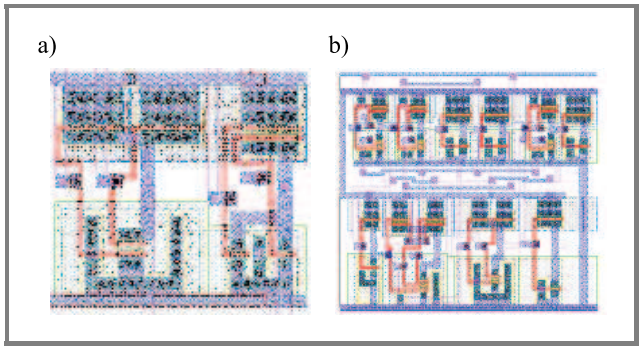


Fig. 10. Layouts of digital cells: (a) NAND, NOR gates; (b) data flip-flop cell.

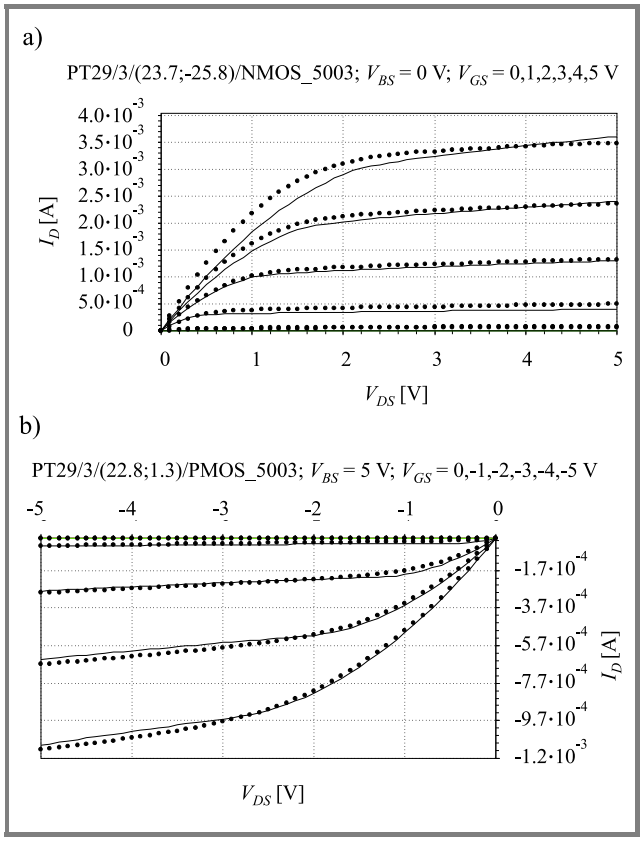


Fig. 11. Measured *I-V* characteristics of (a) n- and (b) p-channel MOSFETs (dots) and those simulated using SPICE LEVEL = 2 MOSFET model after parameters extraction (solid line).

times, propagation times, noise margins and other ones may be measured using these cells and compared with the simulation results. These measurements will also enable the maximum readout speed to be estimated for the readout electronics and the optimum bias voltage to be found for the digital circuits used in the SUCIMA project.

3. Measurements

In this section selected results of the measurements of individual transistors and functional blocks are shown.
The *I-V* characteristics of MOSFETs have been measured using both manual and semi-automatic probe-stations [6] combined with the METRICS software. Next the measured *I-V* characteristics were used to extract SPICE model parameters [4]. The results are shown in Fig. 11. Moreover, measurements of the functional blocks have been carried out. Certain results are shown in Figs. 12 and 13. These preliminary results suggest that the functional blocks work

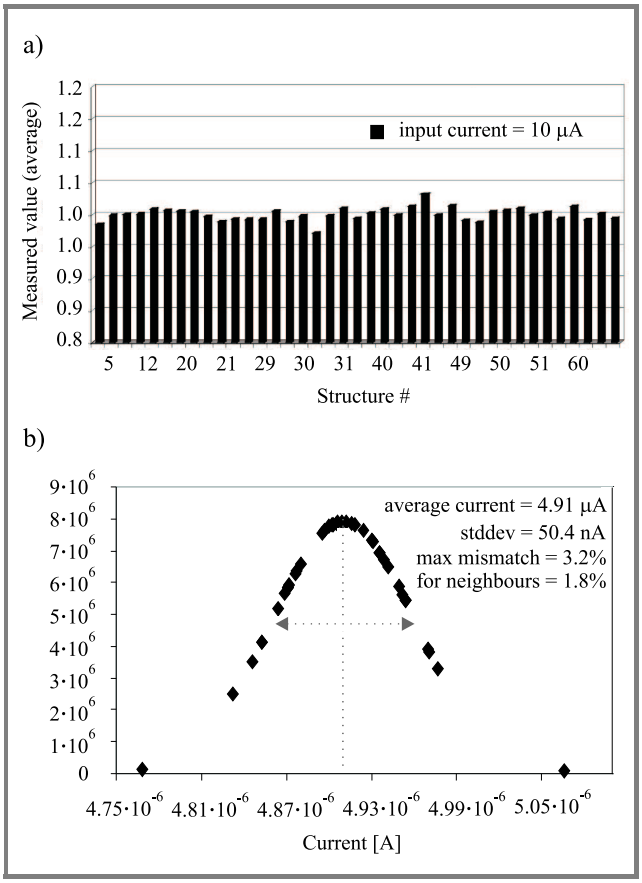


Fig. 12. Matching properties of the NMOS transistor with the dimensions of 50 μm/10 μm: (a) normalized output current at different wafer locations; (b) normal fit of measured current values – the current deviations between different structures on the same wafer are ca. 1% of the average value.

properly. The uniformity of the electrical parameters within wafers is satisfactory and proves that the technology under test is mature.

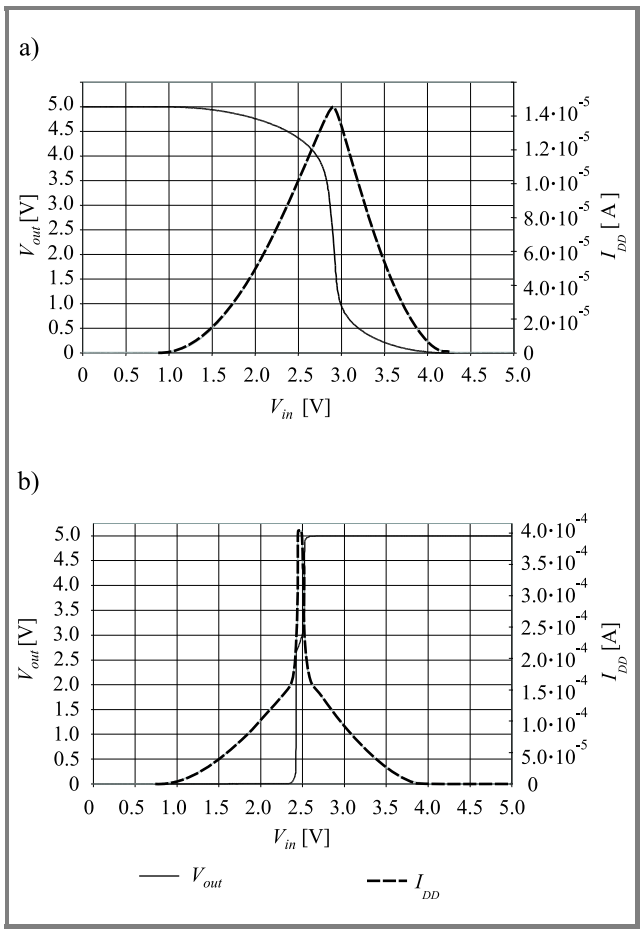


Fig. 13. Transfer characteristics and supply currents of different digital cells: inverter (a); digital buffer (b).

4. Conclusion

The TSSOI test structure is a versatile tool for future optimization of SOI-based pixel detectors. The proposed set of devices provides a basis for detailed studies of both process sequence and design.

Acknowledgment

Supported by E.C. Contract GRID-CT-2001-000561.

References

[1] M. Caccia *et al*., "Silicon ultra fast cameras for electron and gamma sources in medical applications", in *8th Top. Sem. Innov. Part. Radiat. Detec*., Siena, Italy, 2002 (*Nucl. Phys. B – Proc. Suppl.*, vol. 125, pp. 133–138, 2003).

[2] J. Marczewski *et al*., "Monolithic silicon pixel detectors in SOI technology", in *Lin. Coll. Worksh*., Prague, Czech Republic, 2002, http://hep2.fzu.cz/ecfadesy/Talks/ Vertex Detector.

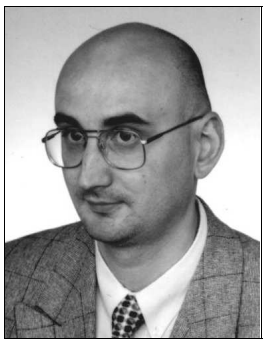
[3] H. Niemiec *et al*., "Technology development for silicon monolithic pixel sensor in SOI technology", in *10th Int. Conf. MIXDES*, Łódź, Poland, 2003, pp. 508–511.

[4] D. Tomaszewski *et al*., "A versatile tool for MOSFETs parameters extraction", in *6th Symp. Diagn. & Yield*, Warsaw, Poland, 2003.

[5] M. Sadowski and D. Tomaszewski, "An efficient approach to the measurement and characterization of MOSFETs capacitances", *Microelectron. Reliab*., vol. 40, pp. 1045–1049, 2000.

[6] J. Zając *et al*. "Semi-automatic test system for characterisation of ASIC/MPWs", in *6th Symp. Diagn. & Yield*, Warsaw, Poland, 2003.

Mateusz Barański received the M.Sc. degree in 2001 from Faculty of Electrical Engineering, Automatics, Computer Science and Electronics of University of Mining and Metallurgy, Kraków, Poland. Since 2001 he has been involved in development of silicon detectors of ionizing radiation for medical applications. Since 2002 he has been also involved in development of measurement devices for geophysical applications. He is an author of several scientific papers related to detection of ionizing radiation.
e-mail: mateusz.baranski@geofizyka.krakow.pl
AGH University of Science and Technology
A. Mickiewicza st 30
30-059 Kraków, Poland



Krzysztof Domański received the M.Sc. and Ph.D. degrees in electronic engineering from Warsaw University of Technology, Poland, in 1996 and 2003, respectively. In 1998 he joined the Institute of Electron Technology (IET), Warsaw, Poland, where he was involved in research on MEMS and porous silicon technology. His current research interests are silicon microfabrication and microsystems for medical applications.
e-mail: kdoman@ite.waw.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Mirosław Grodner received the M.Sc. degree in electrical engineering from Electronics Faculty of Warsaw University of Technology in 1970. Later he joined the Institute of Electron Technology (IET), Warsaw, Poland. He has been involved in the design of integrated circuits: contactless Hall-effect switches, TTLs, p-MOS gates, CMOS calculator circuits, and recently mixed-signal ASICs. At the beginning of the 90's and for the last three

years he has been working on the design of test structures for bulk and SOI CMOS processes.
e-mail: mgrodner@ite.waw.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Bohdan Jaroszewicz received the M.Sc. degree in electronics from the Warsaw University of Technology in 1968. He joined UNITRA-CEMI (Semiconductor Devices Factory) in 1968 and was involved in a variety of research, engineering and production activities related to MOS integrated circuits. As the engineering section manager he was responsible for optimisation of MOS ICs yield. He joined the Institute of Electron Technology (IET) in Warsaw, in 1997. His main research interests are focused on ion implantation engineering, as well as on design, fabrication and characterisation of ISFETs and ionising radiation detectors.
e-mail: bjarosz@ite.waw.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Andrzej Kociubiński received the M.Sc. degree in electrical engineering from Warsaw University of Technology in 2002, where he is currently working towards the Ph.D. degree. His research interests include diagnostics, characterization and modeling of silicon devices. His recent work is related to SOI devices, pixel detectors on SOI wafers and integrated microsystems.
e-mail: akociub@wp.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Wojciech Kucewicz received the M.Sc. degree in electronics engineering from Technical University, Wrocław, Poland, in 1977. The same year he joined as a scientist the Electronics Department of the University of Mining and Metallurgy in Kraków, Poland, where he has been working since. He received the Ph.D. degree in 1987, D.Sc. in 2001 and he became a Professor in 2003. W. Kucewicz was a Visiting Scientist at the Physics Department of Milan University in Italy (1987–1991), the Physics Department of UIC – Chicago (1998–1999), the Karlsruhe University in Germany (2002) and the Insubria University in Como, Italy (2003). His research interests are concentrated on the silicon position sensitive detectors and their application in the high energy physics and medicine. He is a participant of DELPHI and LHCb experiments in CERN, and PHOBOS experiment at BNL (USA).
e-mail: kucewicz@uci.agh.edu.pl
AGH University of Science and Technology
A. Mickiewicza st 30
30-059 Kraków, Poland

Krzysztof Kucharski received the M.Sc. and Ph.D. degrees in electrical engineering from Warsaw University of Technology in 1979 and Institute of Electron Technology (IET) in 1995, respectively. His research area covers IC yield modeling and intrinsic gettering in Cz silicon devices. His recent activities are related to CMOS technology on bulk and SOI substrates, SOI devices, and pixel detectors on SOI wafers.
e-mail: kucharski@ite.waw.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

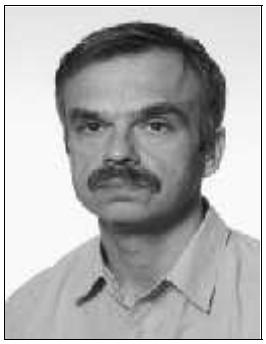
Jacek Marczewski received the M.Sc. degree in solid state electronics from Warsaw University of Technology in 1977. In 1982 he received the Ph.D. degree in electrical engineering from the Institute of Electron Technology (IET), Warsaw. He has been with IET since 1977. His research interest is concentrated on silicon devices and technology. Currently he is working on the development of monolithic pixel detectors of ionizing radiation manufactured in silicon-on-insulator technology.
e-mail: jmarcz@ite.waw.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Halina Niemiec received the M.Sc. in electronics and telecommunications from AGH University of Science and Technology, Kraków, Poland, in 2001. She is a Ph.D. student at the AGH University of Science and Technology, Kraków, in the field of electronics engineering. From 2000 till 2001 she was involved in radiation hardness tests of deep submicron integrated circuit technologies at the European Laboratory for Particle Physics CERN. Since 2001 she has been working on the development of a monolithic active pixel detector fabricated in silicon-on-insulator technology.

e-mail: manka@uci.agh.edu.pl
AGH University of Science and Technology
A. Mickiewicza st 30
30-059 Kraków, Poland

Maria Sapor received the M.Sc. degree in electronics and telecommunications from Wrocław Technical University, Poland, in 1976 and the Ph.D. degree in electronics from AGH University of Science and Technology, Poland, in 1986. Her main research interests are focused on transmission broadband systems, multimedia systems, integrated circuit design, DAQ, control systems and networks with distributed intelligence. Since 1994 she has been involved in international cooperation and EU projects. She is the co-author of more than 50 scientific publications and 1 patent.

e-mail: sapor@uci.agh.edu.pl
AGH University of Science and Technology
A. Mickiewicza st 30
30-059 Kraków, Poland



Daniel Tomaszewski received the M.Sc. and Ph.D. degrees in electrical engineering from Warsaw University of Technology in 1980 and Institute of Electron Technology (IET) in 1998, respectively. His research interests include diagnostics, characterization, simulation and modeling of silicon devices. His recent activities are related to SOI devices, pixel detectors on SOI wafers, ISFETs and parameter extraction methods.

e-mail: dtomasz@ite.waw.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Piotr B. Grabiec – for biography, see this issue, p. 45.

Paper

Silicon TCD for the methane and carbon monoxide detection

Jan M. Łysko, Bogdan Latecki, Marek Nikodem, Marianna Górska, Jerzy Mróz, and Marcin Małachowski

Abstract— Analytical model, design principles, technology and test results concerning a thermal conductivity detector (TCD) are presented. Prototype TCD units fabricated using the standard silicon IC VLSI and MEMS techniques are reported. The detectors are integrated with gas separation columns and micro-valve dosing systems. Initial tests were carried out in a gas mixture containing methane, carbon monoxide, oxygen, hydrogen and nitrogen.

Keywords— detector, silicon, gas mixture.

1. Introduction

Thermal conductivity detector (TCD) prototype units were designed and fabricated within the framework of the micro total analysis systems (μTAS) project, which is focused on the hazardous gas chemical detection-recognition in the coal-mine atmosphere. The device consists of two elements – silicon chip and glass plate, which are bonded together. At the silicon-to-glass interface there are two parallel capillary channels with a system of Pt resistors overhanged across them. There are also four external gas tubing connections to the channels outlets at the device edges. Resistors are connected into a Wheatstone bridge. One channel is used for the test mixture flow, while the second one for the reference flow of pure helium. Changes of the thermodynamic parameters of the gas mixture affect the resistor temperature and the temperature profile along the TCD channel. The most advantageous features of the presented device are low geometrical dimensions of thin film resistors and extremely low thermal capacity, which makes the device very sensitive and its response fast.

2. Model

The performance of a TCD unit may be analyzed using a very simple mathematical model. Sufficiently good results were obtained with the model deriveded from the one previously described by Koch *et al*. [1–3]. The differential Eq. (1) of the thermodynamic equilibrium:

$$\frac{\mathrm{d}(P-Q)}{\mathrm{d}x} = \lambda_f A \frac{\mathrm{d}^2 T}{\mathrm{d}x^2} = \rho_f c_f A v \frac{\mathrm{d}T}{\mathrm{d}x} - 2\lambda_f \frac{l_y}{l_z} T \qquad (1)$$

was solved with several simplifying assumptions and necessary boundary conditions.

The final Eqs. (2–6) of the model enable fast and easy PC calculations of the temperature distribution along the TCD channel for the given conditions:

$$T_1(x) = T_R \exp\left(\chi_1 (x+L)\right) \qquad (x < -L), \qquad (2)$$

$$T_2(x) = T_R \qquad (-L < x < L), \qquad (3)$$

$$T_3(x) = T_R \exp\left(\chi_2 (x-L)\right) \qquad (x > L), \qquad (4)$$

$$\chi_{1/2} = \frac{1}{2 l_z A \lambda_f} \left(l_z v A c_f \rho_f \pm \sqrt{l_z A \left(l_z v^2 c_f^2 \rho_f^2 A + 8 l_y \lambda_f^2 \right)} \right), \qquad (5)$$

$$T_R = \frac{l_z \left(\dfrac{I^2 \rho_{met} n l_y^2}{A_{met}} + 4 \lambda_{met} A_{met} T_B \right)}{\lambda_f l_y l_z A \chi_1 - \lambda_f l_y l_z A \chi_2 + 4 \lambda_f l_y^2 L + 4 \lambda_{met} A_{met} l_z}. \qquad (6)$$

Parameters applied in the analytical equation model: T – temperature, T_B – ambient temperature, T_R – resistor temperature, P – electrical power delivered to the resistor, Q – power loses caused by resistor thermal conduction, I – resistor current, n – number of the resistor meanders, w – resistor path width, g – resistor path thickness, s – separation between the resistor meanders, A_{met} – resistor cross section area = wg, $2L$ – resistor total width = $nw+(n-1)s$, ρ_{met} – metal resistivity, λ_{met} – metal thermal conductivity, v – gas flow, l_y – channel width, l_z – channel depth, A – channel cross section area = $2l_y l_z$, λ_f – gas thermal conductivity, c_f – gas thermal capacity, ρ_f – gas specific density.

3. Design and technology

The TCD unit consist of two elements: a silicon chip (15 mm × 15 mm × 380 μm) and a glass plate (15 mm × 15 mm × 2 mm) [4–6]. The glass plate has two parallel grooves (600 μm × 600 μm × 15 mm) formed with the diamond-disc milling technique. Silicon chips were fabricated in a standard CMOS IC technological facility using several non-standard MEMS techniques. Thin film Pt re-

sistors were patterned with the lift-off technique coupled with the application of negative photoresist (Figs. 1 and 2). After the silicon wafer dicing step, TCD chips had to come back to the "clean room" for the final chemical crystal anisotropic etching step. PECVD Si_3N_4/SiO_2 layers over the silicon substrate are sufficiently resistant to

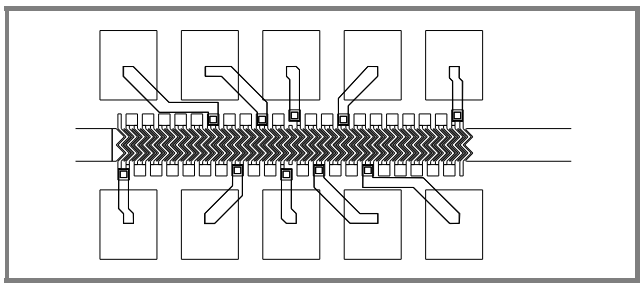


Fig. 1. Fragment of the TCD mask layout.

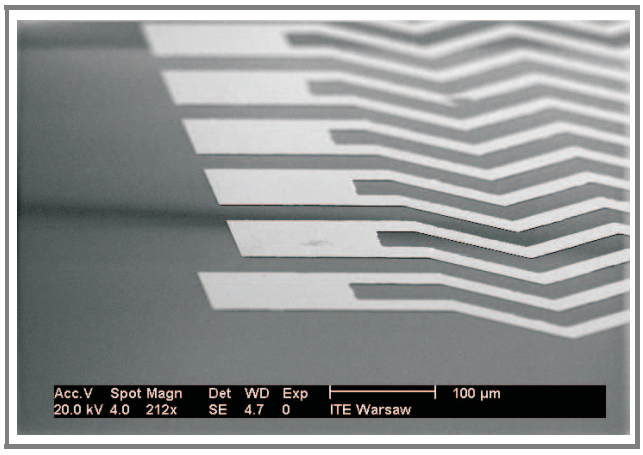


Fig. 2. Pt resistors before the channel photolithography and etching steps.

withstand the etching step in TMAH (tetra methyl ammonium hydroxide and water, $N(CH_3)_4OH + H_2O$) solution at 85°C for 200 minutes. During this etching step Pt resistors were laterally under-etched and suspended over the 400-μm wide and 100-μm deep grooves (Fig. 3). Finally, the silicon chip and the glass plate were aligned and bonded together.

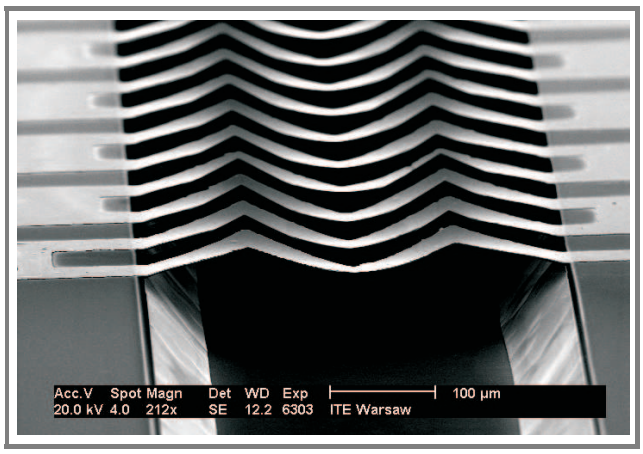


Fig. 3. Pt resistors after the channel photolithography and etching steps.

4. Results

Thermal conductivity detector units were tested as elements of a micro total gas analysis system (Figs. 4 and 5). The resistors were connected to form a fully active Wheatstone bridge, which was operated in the cooling mode. One pair of the resistors, located at the opposite bridge positions, was cooled by the pure carrier gas (helium) flowing through the reference channel. The other pair of the resistors was placed inside the stream of the carrier gas containing separated components of the gas mixture under test, flowing through the second TCD channel. The Wheatstone bridge was supplied from a standard battery source 9 V DC.

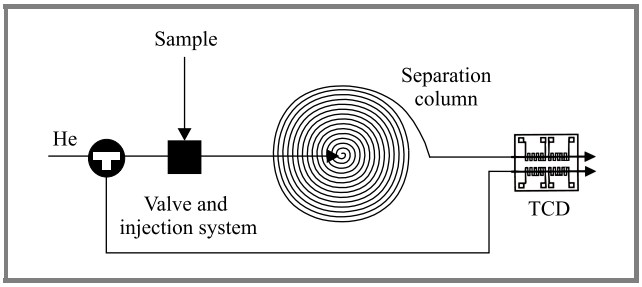


Fig. 4. TCD unit in the micro total analysis system.

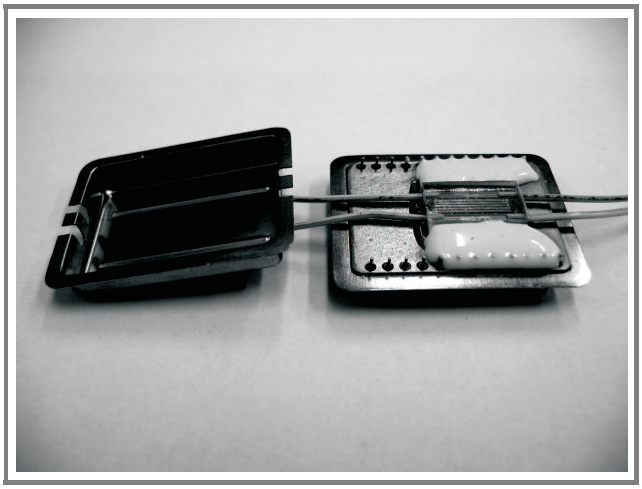


Fig. 5. TCD unit in the open package.

The following gas mixture of typical components of a coal-mine atmosphere has been used in the TCD tests: 20% CH_4, 8% CO, 15% O_2, 4% H_2, 55% N_2. The first

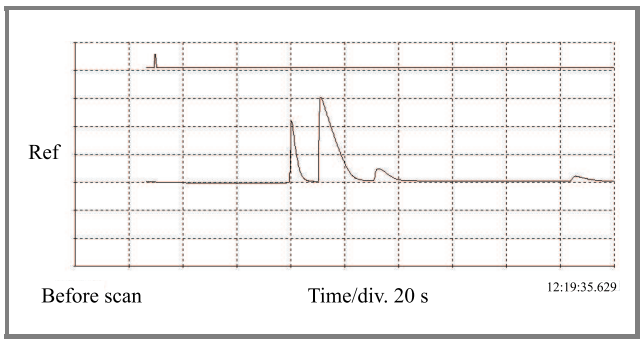


Fig. 6. TCD $U_{OUT}(t)$ plot in conditions of 26 cm/s (1.2 ml/min) carrier gas flow, upper line – sample injection, lower line TCD response: [O_2] peak 45 mV $\approx$ 50 s from the injection, [N_2] peak 60 mV $\approx$ 60 s from the injection, [CH_4] peak 10 mV $\approx$ 80 s from the injection, [CO] peak 5 mV $\approx$ 155 s from the injection.

experiment (Fig. 6) has been carried out with the helium carrier gas flow of 26 cm/s (1.2 ml/min). The reference flow rate chosen for the second experiment (Fig. 7) was 60 cm/s (2.8 ml/min). The output voltages of the Wheatstone bridge were plotted as a function of time. The horizontal axis scale was 20 s/div, while that of the vertical axis 20 mV DC/div. Pneumatically activated microvalve dosing system has been used for precise sample volume injection into the carrier gas stream – in both experiments 14 μl volume samples were used. The outlet of the dosing system was connected with the capillary sep-

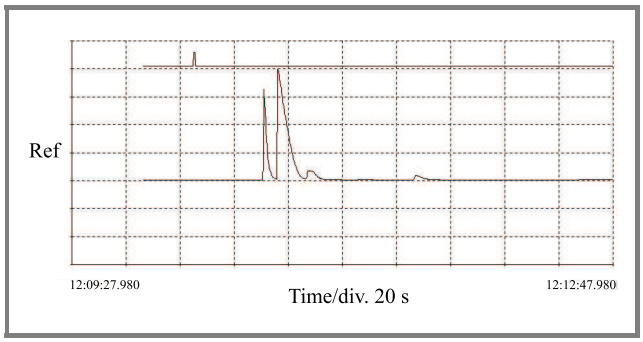


Fig. 7. TCD $U_{OUT}(t)$ plot in conditions of 60 cm/s (2.8 ml/min) carrier gas flow, upper line – sample injection, lower line TCD response: [O_2] peak 60 mV $\approx$ 25 s from the injection, [N_2] peak 80 mV $\approx$ 30 s from the injection, [CH_4] peak 10 mV $\approx$ 40 s from the injection, [CO] peak 5 mV $\approx$ 80 s from the injection.

aration column several meters long. Inside this column the gas mixture sample was diluted in the helium carrier gas and separated into the individual components, which were diluted in the carrier gas only. Every chemical component requires a specific time period to pass the separation column and reach the detector. It depends mainly on the carrier gas velocity and the internal surface coating of the capillary. Under stable thermodynamic stable the timing between the sample injection and electrical signals is used for the chemical identification of the mixture composition. The column, the tubing and the internal diameters of the TCD channel have to be very close to reduce gas stream disturbances and flow resistance.

A change in the chemical composition of the gas inside the TCD channel affects the thermodynamic properties of the system (mixture thermal conductivity, thermal capacity, specific density). The detector transforms these changes into a temperature profile shift along the TCD channel axis. In the cooling mode, decreased/increased heat transfer from the resistor to the gas stream results in an in-creased/decreased resistor temperature. The resultant material resistivity changes are registered by the Wheatstone bridge output voltage ΔU_{OUT}.

5. Conclusions

Thermal conductivity detector devices were successfully designed and fabricated. Test results in a well defined gas mixture of hazardous components of a coal-mine atmosphere were satisfactory. Further improvements of device sensitivity are possible if the Wheatstone bridge supply voltage is increased and the geometrical dimensions of the resistors are reduced. Using poly-Si resistors doped with P, As or B insted of Pt resistors seems to be a very promising idea seems to be implementation instead of the Pt resistors, because the latter ones are supposed to be chemically active in contact with the hydrocarbons.

References

[1] M. Koch, A. Evans, and A. Brunnschweiler, *Microfluidic Technology and Applications*. Baldock, Hertfordshire: RSP, 2000.

[2] T. S. J. Lammerink, N. R. Tas, M. Elwenspoek, and J. H. J. Fluitman, "Micro-liquid flow sensor", *Sens. Actuat.*, vol. A37-38, pp. 45–50, 1993.

[3] B. Latecki, J. Łozinko, Jan M. Łysko, "Krzemowy katarometr – model analityczny", *Elektronizacja*, no. 12, pp. 23–25, 2001 (in Polish).

[4] J. M. Łysko, J. Łozinko, B. Latecki, M. Górska, K. Hejduk, and H. Wrzesińska, "Pt resistors technology for the TCD application", in *Proc. Joint IMEKO TC-1 & XXXIV MKM Conf.*, Wrocław, Poland, 2002, vol. III, pp. 111–118.

[5] J. M. Łysko, M. Górska, A. Panas, K. Studzińska, and H. Wrzesińska, "BSC-type contacts formation in the katarometer technological sequence", in *Proc. Joint IMEKO TC-1 & XXXIV MKM Conf.*, Wrocław, Poland, 2002, vol. III, pp. 103–110.

[6] J. M. Łysko, M. Nikodem, B. Latecki, M. Górska, and K. Studzińska, "Elements of the silicon TCD design and technology", in *Proc. 6th Int. Symp. Microelectron. Technol. Microsyst.*, Lviv, Ukraine, 2002 also *J. Lviv Polytech. Nat. Univ.*, vol. 458, pp. 61–66.

Jan M. Łysko received the M.Sc. degree in 1979 from the Warsaw University of Technology, the B.Sc. degree in economics in 1983 from the High School of Planning and Statistics, and the Ph.D. degree in 1990 from the Institute of Electron Technology (IET). Since 1979 he has been working as an adjunct in the IET Sensor Department.
e-mail: jmlysko@ite.waw.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Bogdan Latecki received the M.Sc. degree in 1965 from the Warsaw University of Technology. Later he worked as an R&D engineer at the VLSI IC factory "Tewa" in Warsaw. In 1996 he joined the Sensors Department at the Institute of Electron Technology (IET), where he is currently employed. At present he is responsible for the mechanical sensors and MEMS testing.
e-mail: latecki@ite.waw.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Marek W. Nikodem received the M.Sc. degree in 1973 from the Warsaw University of Technology and joined the research team at the IC VLSI factory "Tewa" in Warsaw. Since 1981 has been working in the Institute of Electron Technology (IET) in the Microwave Devices Department. At present he is involved in several projects related to the microphotonic and laser devices, as well as the MEMS devices assembly.
e-mail: nikodem@ite.waw.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Marianna Górska received the M.Sc. degree from the Chemical Engineering Department, Silesian Technical University, Poland, in 1969. She worked at the factory "Tewa" as an engineer responsible for the VLSI IC photolithography. In 1988 she joined the microwave devices team of the Institute of Electron Technology (IET) in Warsaw. At present she is involved in the photonic and laser devices technology.
e-mail: gorska@ite.waw.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Marcin Małachowski received the M.Sc. degree from the Measurements and Automatic Control in Electrical Engineering Department, Silesian University of Technology in 2002. Since 2001 has been working in the Research and Development Center for Electrical Engineering and Automation in Mining (EMAG) in the Safety Systems and Ventilation Control Department. At present he is involved in the gas chromatography systems.
e-mail: malach@emag.katowice.pl
The Research and Development Center
for Electrical Engineering and Automation
in Mining (EMAG)
Leopolda st 31
40-189 Katowice, Poland

Jerzy Mróz – for biography, see this issue, p. 23.

Paper

Gas micro-flow-metering with the in-channel Pt resistors

Jan M. Łysko, Bogdan Latecki, and Marek Nikodem

Abstract— Standard thermo-conductive gas flow meters have the side-channel integrated with the temperature detector and heater coils, both winded around the tube. This design suffers from a high thermal capacity, reduced sensitivity in the lower limit flow range, high thermal inertia, long response time and necessity to amplify electronically the output signal. The newly designed TCD detector can be applied as a precise gas flow meter. To identify the composition of the unknown gas, the TCD unit requires a connection to the separation column, application of the reference channel and highly stable flow rate regulator. To measure flow rates, the same TCD unit requires only one flow channel application, with active resistors inside it, and a definition of the gas type. In this work principles of the TCD design, technology and flow rate sensitivity tests are presented.

Keywords*— *detector, silicon, gas, flow, thermo-conductivity.

1. Introduction

The thermal conductivity detector (TCD) unit was designed and fabricated for applications related to systems of multi-component gas mixture detection [1–4]. TCD performance requires initial separation of the gas mixture components, stable carrier gas flow rate and two parallel flow channels of highly symmetrical geometry. This device can also be used as a precise gas flow meter. In this mode the gas type has to be determined. The most advantageous feature of the presented device is the location of the resistors inside flow channels and highly reduced geometrical dimensions of metal paths. Resistors were patterned in a thin layer Pt/Cr sandwich by means of an advanced microelectronics technology. Direct contact between the resistors and the investigated gas, as well as extremely low thermal capacity of the resistors, improve sensitivity, output signal level and response time.

2. Design

The TCD unit was designed (Fig. 1) and fabricated with the application of standard microelectronics CAD software, tools and technology, supplemented by several technological steps that are typical of silicon MEMS [2–4]. The resistors were patterned using the Pt/Cr lift-off technique on a silicon substrate covered by PECVD SiO_2/Si_3N_4 layers. Grooves were etched in a (100) silicon substrate in a TMAH + water solution, with the SiO_2/Si_3N_4 layers serving as the etching mask. The TCD unit consists of a glass plate with parallel groves milled with a diamond-disk blade. During the alignment and bonding of the silicon chip and the glass plate, horizontal flow channels were formed at the silicon-glass interface. Flow-channel openings at the opposite edges of TCD were connected to the external gas tubing. The application of microelectronics technology

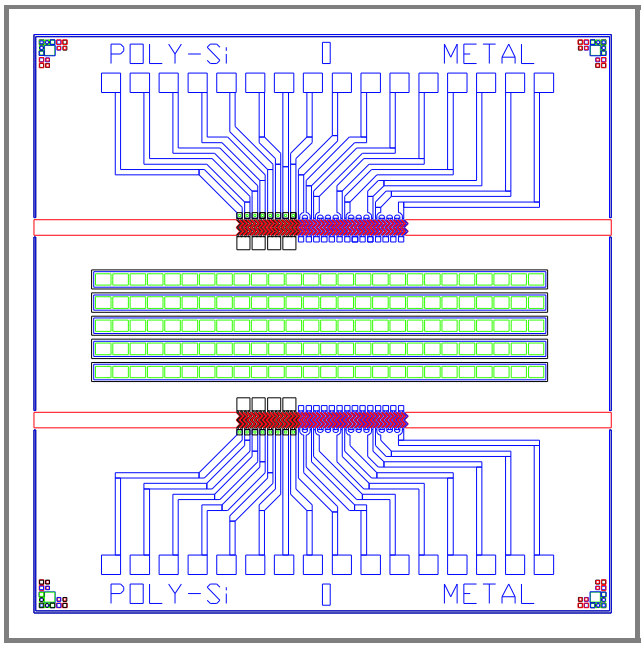


Fig. 1. TCD mask layout.

resulted in low dimensions of the device elements, symmetry, repeatability of physical and geometrical parameters, as well as the satisfactory yield of correctly manufactured device chips.

3. Technology

The most critical steps of the process of TCD silicon chip fabrication are lithography of Pt resistors and groove etching. Resistors were formed on a 3-inch, (100) Si wafer covered with SiO_2/Si_3N_4 PECVD layers. Lift-off technique and negative ma-N-1420 resist were used to pattern metal (4000 Å Cr/Pt layer) paths. A chromium pad of 50 Å is used to improve the adhesion of metal paths to silicon nitride. Grooves in the silicon substrate were defined by the next photolithography mask. The mask contained multiple shapes oriented in the directions [0 $\overline{1}$ 1] and [0 1 $\overline{1}$] and located between the resistor paths. Mask shapes were etched first in the silicon nitride and silicon dioxide layers (Fig. 2a). After resist removal, wafers were diced into

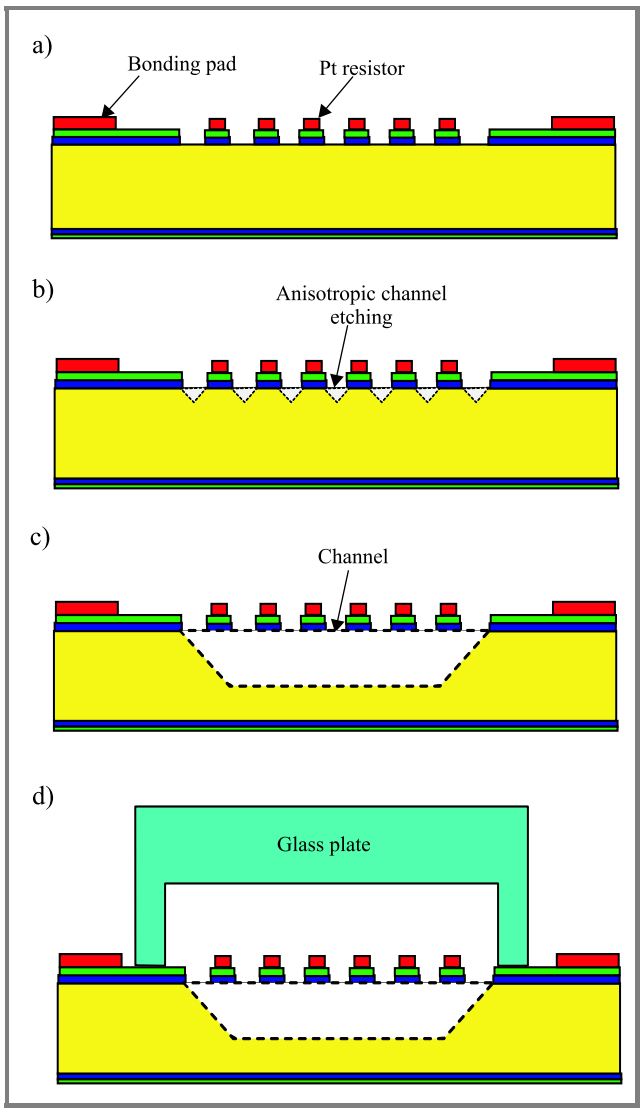


Fig. 2. Schematic drawings of the TCD chip cross section: (a) after the dielectric layer and Pt layer deposition, photolitography of the Pt resistor and photolitography of the channel; (b) initialization of the silicon substrate anisotropic etching; (c) final silicon chip with the resistor paths suspended accross the channel; (d) after silicon-to-glass anodic bonding step.

individual TCD silicon chips and these chips were subjected to anisotropic etching of the Si substrate through the windows opened previously in the double-layer dielectric mask (Fig. 2b). TMAH (tetra methyl ammonium hydroxide and water, $N(CH_3)_4OH + H_2O$) solution was used at 85°C. The duration of the etching process was 200 minutes. During this step SiO_2/Si_3N_4 mask layers were laterally under-etched and Pt resistors were suspended over the edges of grooves (Fig. 2c). Groove final dimensions are: width – 400 μm, depth – 100 μm, length – 15 mm. The side-walls correspond to the (111) crystal planes (Fig. 3). High lateral etch rate under the mask edge is caused by the presence of convex corners and high index crystal planes. Suspended resistors are extremely delicate, therefore even the steps of DI water rinsing and drying require special

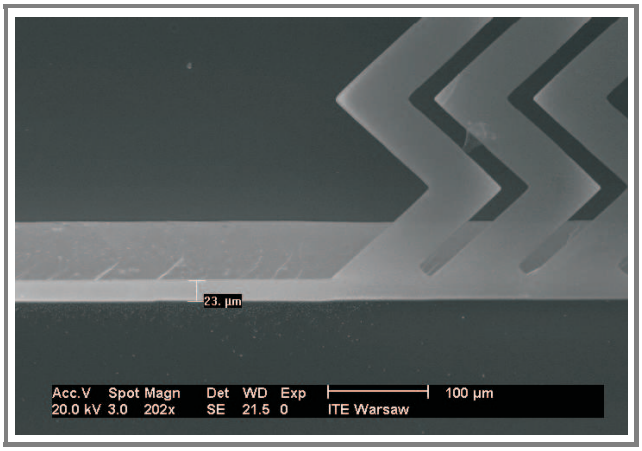


Fig. 3. SEM of a TCD silicon chip in the groove edge region. Lateral etch range in the direction [111] is 23 μm.

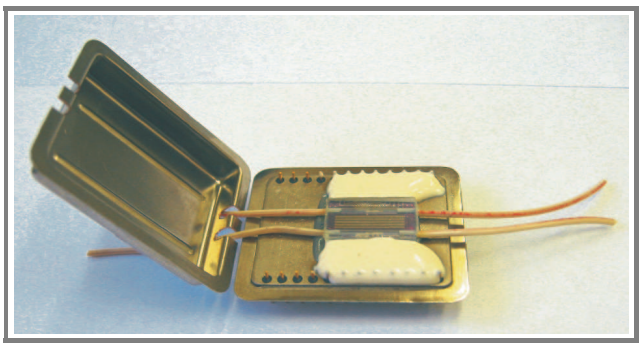


Fig. 4. TCD unit in an open package.

procedures to avoid mechanical damage. The glass plate was aligned and bonded with the silicon chip (Fig. 2d) and placed on the IC package platform. Channel outlets were connected to the external gas tubing and electrical wire connections were made between the TCD bonding pads and package pins (Fig. 4).

4. Results

Newly designed and fabricated TCD detectors are sensitive both to the gas flow velocity changes and gas mixture composition. In the chemical detection mode gas flow velocity should be very stable and each instability has to be compensated by the application of the reference flow channel. In the flow meter mode the gas type should be determined. Static, as well as dynamic, tests were done with the nitrogen gas and flow and electrical system set-up shown in Fig. 5.

Tests were done using a Keithley KPCI-3108 Series PCI Bus Data Acquisition Board and LabView 6.1 software installed on a standard PC. Two TCD units were used for dynamic tests – one of them was active, the second one was electrically passive and served to establish the same gas flow resistance, which is important for the gas flow switching. The active TCD unit was powered by a 8 $V_{(DC)}$ battery source. Flow switching from one channel to another was

performed with an electrically activated Humphrey valve of the MINI_MYTE 41E1 type. A Tylan FC-260 mass flow controller was used as the reference, supported by an additional precise regulator.

Flow resistance versus flow rate characteristics are presented in Fig. 6. Points around the 0 ml/min value are distributed in a slightly non-symmetric way. This can be explained by the non-symmetric bending of the resistors inside the channel (Fig. 3). The U_{OUT} versus nitrogen gas flow rate characteristics (Fig. 7) are nonlinear in the 0–50 ml/min range. Nonlinearity seems to be low and sensitivity high in the flow region below 10 ml/min. It was also observed that in the case of TCD outlet opened to the ambient, the results close to 0 ml/min were disturbed by the penetration of air into the detector flow channel, and the resultant changes of the thermodynamic properties of the gas. Dynamic response times of the new TCD were measured in comparison to those of a piezoresistive pressure sensor (Fig. 8). Sampling frequency to analyze both output signal slopes was 10 000/s for a 3 s period test. Nitrogen flow was switched from 0 to 10 ml/min and back. The pressure sensor and TCD required 26.6 ms and 46.3 ms to reach 0% → 90% ΔU_{OUT}, respectively. Switching from 10 to 0 ml/min revealed slower response times of the pressure sensor (66.6 ms) and slightly faster TCD action (40.7 ms). Dynamic test results show that TCD units are much faster than the classic flow meter devices.

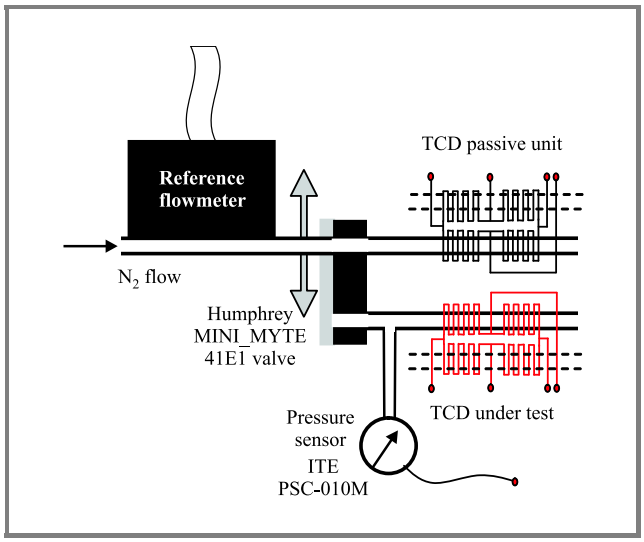


Fig. 5. Flow and electrical system used for static and dynamic tests of TCDs.

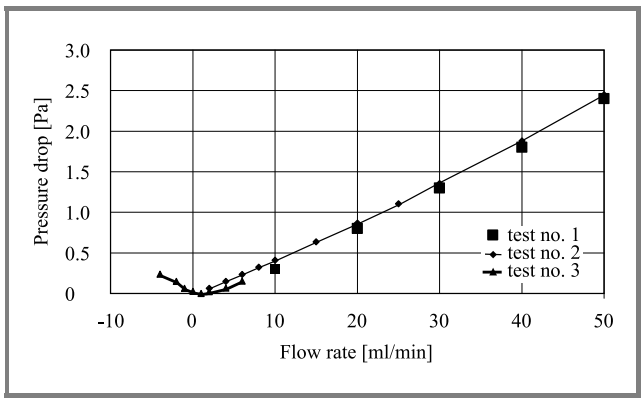


Fig. 6. Flow resistance of a TCD as a function of flow rate.

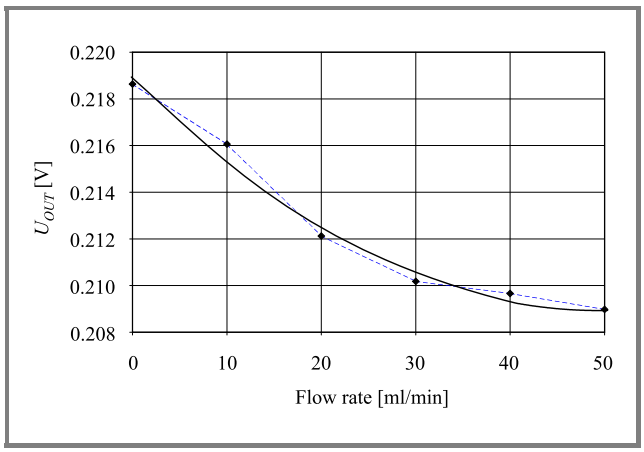


Fig. 7. Static U_{OUT} as a function of nitrogen flow rate.

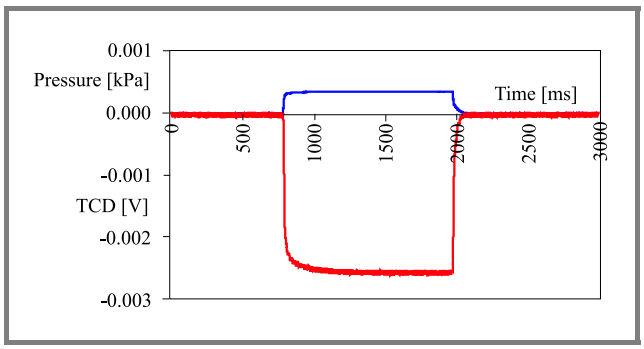


Fig. 8. Dynamic output of a TCD and a pressure sensor as a function of time. Switching of the nitrogen flow: 0 → 10 → 0 ml/min.

References

[1] B. Latecki, J. Łozinko, and J. M. Łysko, "Krzemowy katarometr – model analityczny", *Elektronizacja*, no. 12, pp. 23–25, 2001 (in Polish).

[2] J. M. Łysko, J. Łozinko, B. Latecki, M. Górska, K. Hejduk, and H. Wrzesińska, "Pt resistors technology for the TCD application", in *Proc. Joint IMEKO TC-1 & XXXIV MKM Conf.*, Wrocław, Poland, 2002, vol. III, pp. 111–118.

[3] J. M. Łysko, M. Górska, A. Panas, K. Studzińska, and H. Wrzesińska, "BSC-type contacts formation in the katarometer technological sequence", in *Proc. Joint IMEKO TC-1 & XXXIV MKM Conf.*, Wrocław, Poland, 2002, vol. III, pp. 103–110.

[4] J. M. Łysko, M. Nikodem, B. Latecki, M. Górska, and K. Studzińska, "Elements of the silicon TCD design and technology", in *Proc. 6th Int. Symp. Microelectron. Technol. Microsyst.*, Lviv, Ukraine, 2002; *J. Lviv Polytech. Natl. Univ.*, vol. 458, pp. 61–66.

Jan M. Łysko, Bogdan Latecki, Marek W. Nikodem – for biography, see this issue, p. 97.

Paper

DC and low-frequency noise analysis for buried SiGe channel metamorphic PMOSFETs with high Ge content

Sergiy Durov, Oleg A. Mironov, Maksym Myronov, Terence E. Whall, Evan H. C. Parker, Thomas Hackbarth, Georg Hoeck, Hans-Joest Herzog, Ulf König, and Hans von Känel

Abstract—Measurements of current drive in p-$Si_{1-x}Ge_x$ MOSFETs, with x = 0.7, 0.8 reveal an enhancement ratio of over 2 times as compared to a Si device at an effective channel length of 0.55 μm. They also show a lower knee voltage in the output *I-V* characteristics while retaining similar values of drain induced barrier lowering, subthreshold swing, and off current for devices with a Sb punch-through stopper. For the first time, we have quantitatively explained the low-frequency noise reduction in metamorphic, high Ge content, SiGe PMOSFETs compared to Si PMOSFETs.

Keywords*— *SiGe, metamorphic MOSFET, LF-noise, I-V, C-V, effective hole mobility.

1. Introduction

Strained-Si NMOS and PMOS devices have made remarkable strides in the last year or two and both IBM and Intel are developing full CMOS processes [1, 2]. On the other hand, while there is particular advantage [3] to be gained in increasing the performance of the p-channel current drive, enhancements in this case have been less than those in the area of n-channel. Sugii *et al*. [4], for example, find a current drive enhancement ratio in the n-channel of 1.7 compared to their Si control, but only 1.5 in the p-channel, which is of particular relevance to the current work. A strained $Si_{1-x}Ge_x$ layer capped with strained Si is an attractive alternative that offers higher effective hole mobility than strained silicon only [5] while being also compatible with a full N/P CMOS configuration. In this work we report on PMOSFET devices containing strained $Si_{1-x}Ge_x$ channel with $x = 0.7$, 0.8 and the effective channel length of 0.55 μm. The devices have the maximum effective hole mobilities in the range of 760–500 cm^2/Vs at a vertical effective fields $E_{eff} = 0.08-0.2$ MV/cm, compared to 170–130 cm^2/Vs in bulk Si and 110 cm^2/Vs in our epitaxial Si control. This leads to a current drive enhancement ratio of a factor of more than two whilst maintaining short channel characteristics similar to those of the Si control. The drain current is sub-linear in gate overdrive, implying advantageous high lateral field transport.

The reduction of low-frequency (LF) noise is crucial for achieving high performance in analogue Si-based MOSFET devices [12]. One solution to this problem is via the incorporation of strained SiGe buried layers. Recently there have been several contradictory reports concerning the LF-noise properties of SiGe pseudomorphic FET devices [13–16]. The authors of [14,16] reported, for example, a decrease of the normalised drain current noise power spectral density (NPSD) of pseudomorphic SiGe MOSFETs in comparison with Si controls, others reported an increase of NPSD [12, 13]. LF-noise characteristics and mechanisms of LF-noise reduction in buried channel p-SiGe metamorphic MOSFETs are described.

2. MOSFET fabrication

The MOSFETs fabricated on multilayer SiGe heterostructures grown by two epitaxial techniques are compared. The first structure (Fig. 1) has a $Si_{0.3}Ge_{0.7}$ p-channel and was

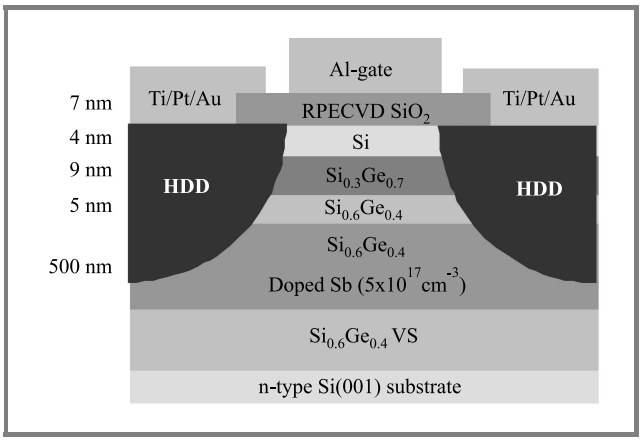


Fig. 1. Schematic cross-section of p-$Si_{0.3}Ge_{0.7}$ MOSFET.

grown by solid-source molecular beam epitaxy (SS-MBE) on an n-type ($1 \cdot 10^{15}$ cm^{-3}) Si(001) wafer. It consist of a 2.5 μm relaxed $Si_{1-y}Ge_y$ virtual substrate (VS) linearly graded to the final Ge composition $y = 0.4$, 500 nm of $Si_{0.6}Ge_{0.4}$:Sb doped at $5 \cdot 10^{17}$ cm^{-3} acting as a "punch-through stopper" to avoid short channel effects, a 5 nm $Si_{0.6}Ge_{0.4}$ spacer layer, a 9 nm compressively strained $Si_{0.3}Ge_{0.7}$ channel, and 4 nm tensile-strained Si cap layer.

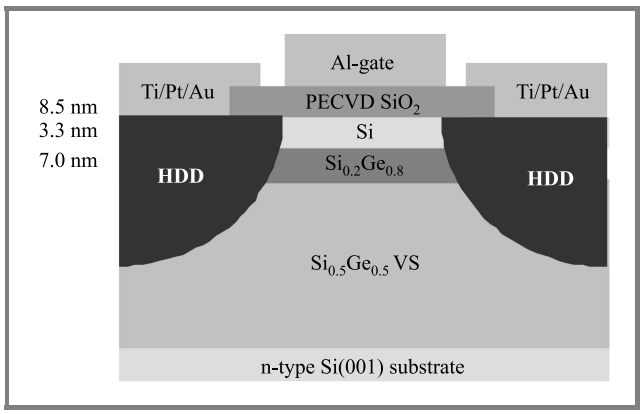


Fig. 2. Schematic cross-section of p-$Si_{0.2}Ge_{0.8}$ MOSFET.

The second structure (Fig. 2) was grown by low energy plasma enhanced CVD (LEPECVD) and differs in that the VS terminates at $y = 0.5$, there is no punch through stopper and the p-channel is 7 nm of strained $Si_{0.2}Ge_{0.8}$. As $x - y = 0.3$ in both structures the strain in the p-channel will be the same. The PMOSFET devices were fabricated using reduced thermal budget processing at 650°C, to minimize Ge out-diffusion from the strained $Si_{1-x}Ge_x$ channel [7] and to avoid Sb penetration to the channel, with 200 nm of plasma enhanced CVD (PECVD) SiO_2 deposited as a field oxide. In the active transistor area the field oxide was removed by wet chemical etching. After a cleaning step the gate oxide on the first SS-MBE grown structure was deposited by remote plasma enhanced CVD (RPECVD) as a 7 nm SiO_2 layer at 300°C [18]. The gate oxide on the second LEPECVD grown structure was 8.5 nm PECVD deposited at 370°C, followed by annealing in a N_2O atmosphere at 650°C for 1 min. Source and drain contacts were fabricated by BF_2^+ implantation at 40 keV, with a dose of $4 \cdot 10^{15}$ cm^{-2} and activated at 650°C for 30 s. The surface of contact areas was etched for a short time to remove impurities increasing contact resistance. Finally, the Al gate and Ti/Pt/Au contact metallization were evaporated. The second p-$Si_{0.2}Ge_{0.8}$(2) device of Table 1 was made using the same process as for MBE-grown p-$Si_{0.3}Ge_{0.7}$ device [5], but the thickness of the SiO_2 layer is 11 nm.
The p-Si MOSFET devices were fabricated on SS-MBE grown 100 nm Si epilayer, grown on n-type ($1 \cdot 10^{17}$ cm^{-3}) Si(001) wafers using a self-aligned gate process, with 9 nm dry SiO_2 thermally grown at 800°C for 120 min and 300 nm p-type ($5 \cdot 10^{19}$ cm^{-3}) poly-Si gate. The row of geometrical gate lengths for all fabricated transistors was in the range $L = 0.4 - 50$ μm with the same gate width $W = 50$ μm.

Table 1
Electrical and structural properties of 0.55 μm p-$Si_{0.3}Ge_{0.7}$, p-$Si_{0.2}Ge_{0.8}$ and p-Si MOSFETs

Parameter	Si	$Si_{0.3}Ge_{0.7}$	$Si_{0.2}Ge_{0.8}$	$Si_{0.2}Ge_{0.8}$ (2)
SiO_2 thickness [nm]	9	7	8.5	11
$g_{m\,(sat)}$ [mS/mm]	40	84	95	63
S [mV/decade]	85	95	130	200
V_{TH} [V]	–0.2	–0.84	–0.26	–0.95
I_{ON}/I_{OFF} [$V_{DS} = -50$ mV]	10^6	10^6	$2.5 \cdot 10^3$	$2.5 \cdot 10^3$
I_{ON}/I_{OFF} [$V_{DS} = -3$ V]	10^4	10^4	15	26

3. DC characteristics

Current-voltage (*I-V*) and quasistatic capacitance-voltage (*C-V*) characteristics were measured using an Agilent 4156C parameter analyzer for all devices at a temperature of 293 K (the basic parameters are given in Table 1). The input *I-V* characteristics for the $Si_{0.3}Ge_{0.7}$ PMOSFET in Fig. 3 show reduced drain induced barrier lowering (DIBL)

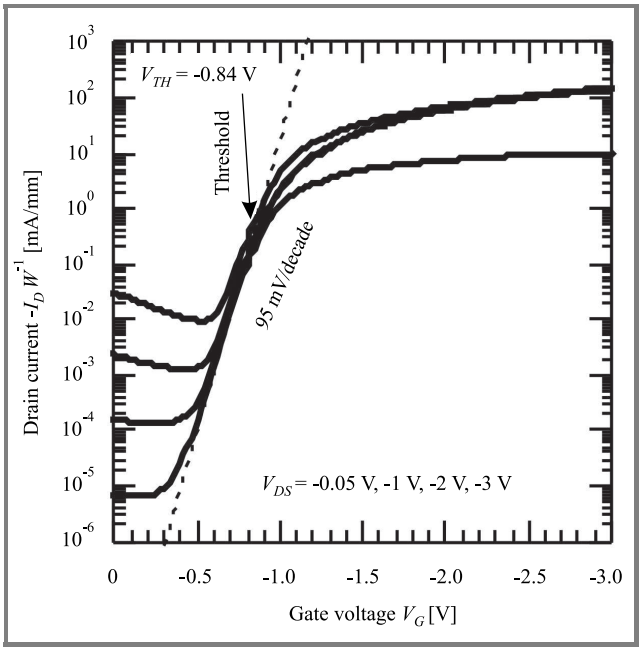


Fig. 3. Input *I-V* characteristics for p-$Si_{0.3}Ge_{0.7}$ MOSFET with $L_{eff} = 0.55$ μm.

and an excellent subthreshold swing $S = 95$ mV/decade at $V_{DS} = -50$ mV, which demonstrates the efficiency of the "punch-through stopper" for sub-micron MOSFET operation. This device has an excellent I_{ON}/I_{OFF} ratio of 10^6 in the linear region ($V_{DS} = -50$ mV) and in saturation ($V_{DS} = -3$ V) $I_{ON}/I_{OFF} \approx 10^4$. The threshold voltage V_{TH} is -0.84 V at $V_{DS} = -50$ mV.
The input *I-V* characteristics for the p-Si MOSFET are shown in Fig. 4. In this case, the subthreshold swing is 85 mV/decade and I_{ON}/I_{OFF} is 10^6 in the linear region and 10^4 in saturation. The threshold voltage V_{TH} is –0.2 V. Comparison of these two devices shows the metamorphic MOSFET is operating in an acceptable way and provides a competitive device at this technology node. The slight increase in S in the p-$Si_{0.3}Ge_{0.7}$ device can be completely accounted for by the added capacitance of the strained Si overlayer.
The electrical characteristics of the p-$Si_{0.2}Ge_{0.8}$ MOSFET, which does not have a punch through stopper, are

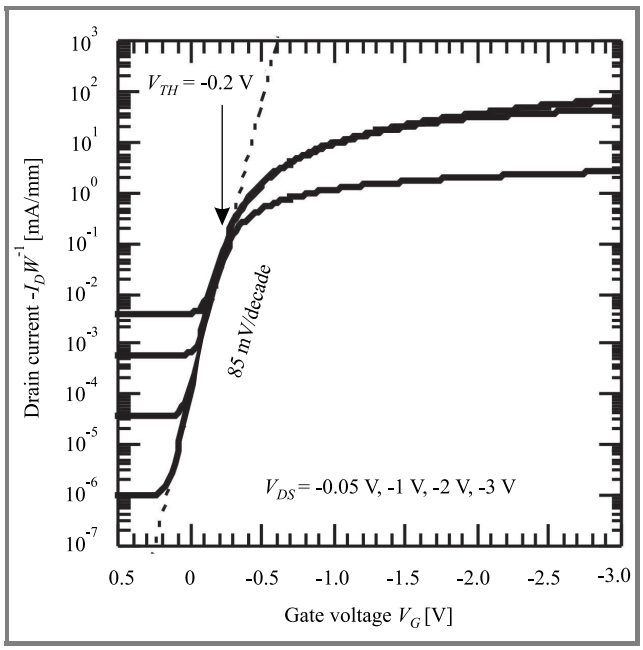


Fig. 4. Input *I-V* characteristics for p-Si MOSFET with $L_{eff} = 0.55$ μm.

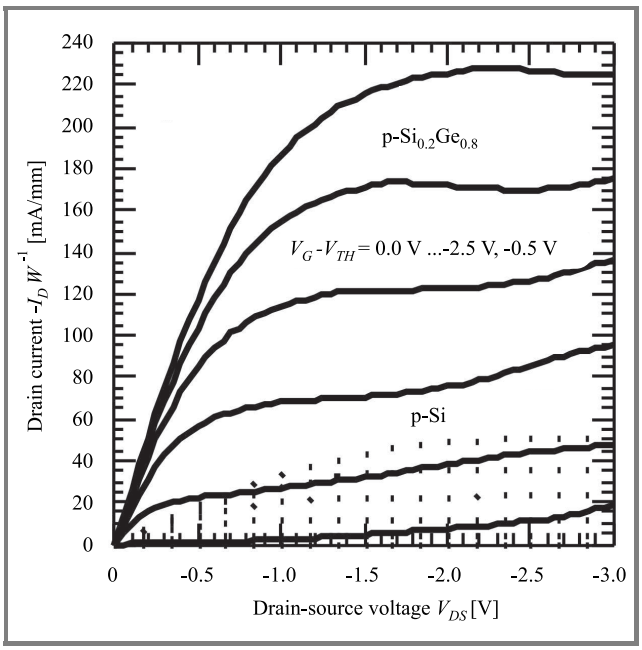


Fig. 6. Output *I-V* characteristics of p-Si and p-Si$_{0.2}$Ge$_{0.8}$ MOSFET with $L_{eff} = 0.55$ μm at the same V_G–V_{TH}.

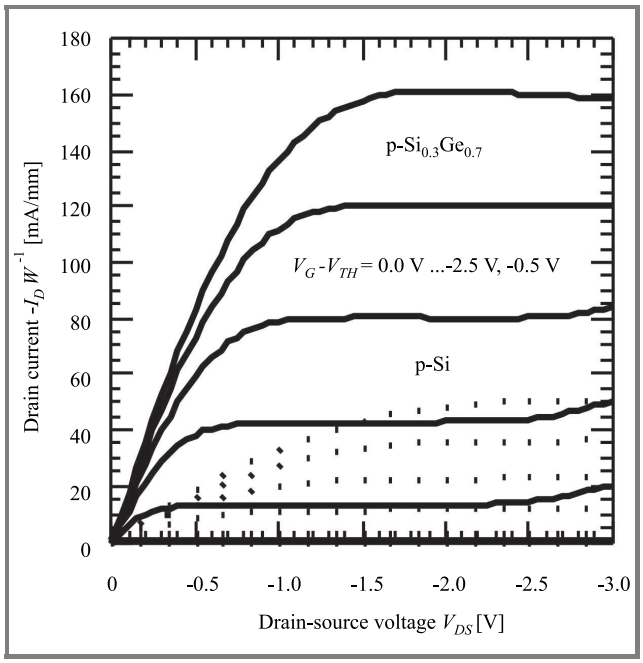


Fig. 5. Output *I-V* characteristics of p-Si and p-Si$_{0.3}$Ge$_{0.7}$ MOSFET with $L_{eff} = 0.55$ μm at the same V_G–V_{TH}.

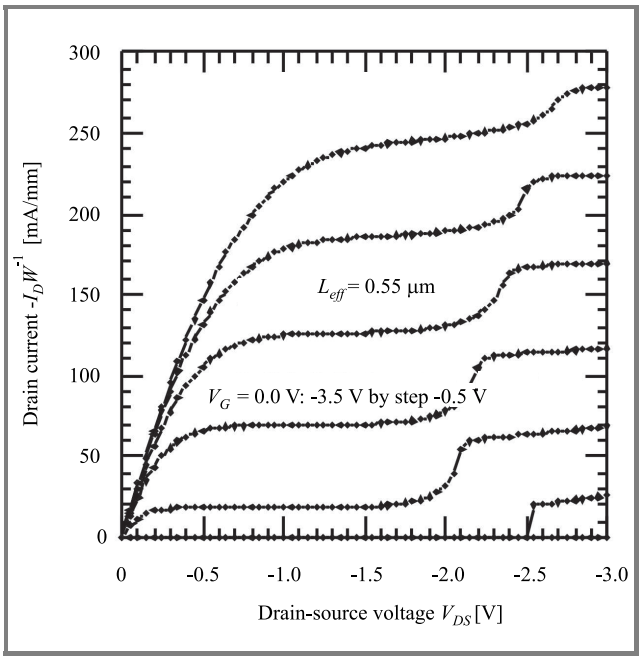


Fig. 7. Kink effect on output *I-V* characteristics of p-Si$_{0.3}$Ge$_{0.7}$ at low temperatures $T = 77$ K.

not as impressive. The subthreshold slope is increased to 130 mV/decade and the device does not switch off well resulting in a much reduced I_{ON}/I_{OFF} ratio. This is a consequence of the vertical architecture not being optimised for sub-micron device operation rather than any inherent problem with the channel material, as will be seen in the mobility measurements and output characteristics below.

The maximum transconductance in the saturation region is $g_m = 84$ and 95 mS/mm respectively compared to 40 mS/mm in the Si control. The maximum drain current at $V_G - V_{TH} = -2.5$ V is 165 mA/mm for p-Si$_{0.3}$Ge$_{0.7}$ and 230 mA/mm for p-Si$_{0.2}$Ge$_{0.8}$.

The output *I-V* characteristics measured on both the p-Si$_{0.3}$Ge$_{0.7}$ and p-Si$_{0.2}$Ge$_{0.8}$ MOSFETs are shown in Fig. 5 and Fig. 6, with comparisons to the p-Si device. Enhance-

ment in the saturated drain current by a factor of 2.5–3 is clearly visible in the output *I-V* characteristics of the $Si_{0.3}Ge_{0.7}$ PMOSFET (Fig. 5) at $V_{DS} = -2.5$ V, in comparison with the silicon control. Similar enhancement is seen at all values of drain bias. For the $Si_{0.2}Ge_{0.8}$ PMOSFET the enhancement factor in the normalised saturation drain current is actually higher than for the $Si_{0.3}Ge_{0.7}$ PMOSFET and is more than a factor of three above the control.

The *I-V* characteristics of the p-$Si_{0.2}Ge_{0.8}$(2) MOSFET are similar to those of p-$Si_{0.2}Ge_{0.8}$ MOSFET and differ only in slightly lower drain current values due to thicker gate dielectric.

The self-heating effect, which is responsible for the mobility degradation, threshold voltage lowering and negative differential conductance, was observed in all high Ge content metamorphic SiGe MOSFETs with gate length below 2 μm at high V_{DS}. The "kink" effect (Fig. 7) was clearly observed at low temperature (77 K) for devices with a punch-through stopper (p-$Si_{0.3}Ge_{0.7}$). This is due to the majority carriers generated by impact ionization that are collected in the body and increase the body potential (lower threshold voltage). For devices without a punch-through stopper (p-$Si_{0.2}Ge_{0.8}$) the "kink" effect was not observed. This behavior of our devices is similar to partially depleted silicon-on-insulator (SOI) MOSFETs [10].

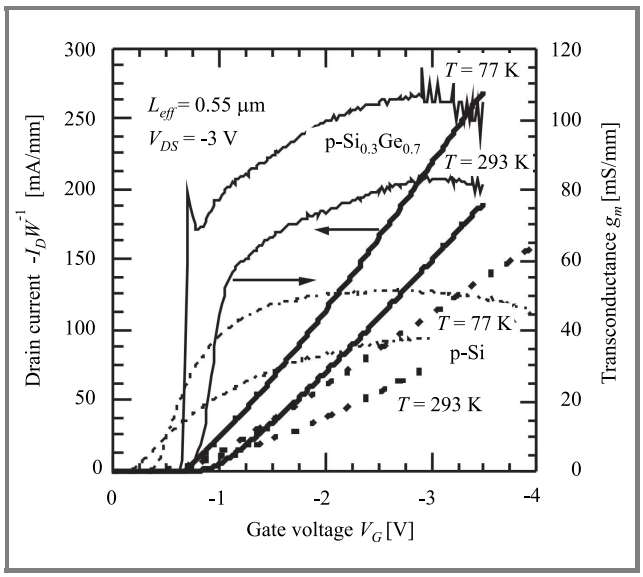


Fig. 8. Drain current I_D (thick lines) and transconductance g_m (thin lines) versus gate voltage for p-$Si_{0.3}Ge_{0.7}$ (solid lines) and p-Si (dashed lines) MOSFETs with effective gate length 0.55 μm at room ($T = 293$ K) and nitrogen ($T = 77$ K) temperatures.

The low temperature measurements have been carried out in liquid nitrogen ($T = 77$ K). The input *I-V* characteristics for the p-$Si_{0.3}Ge_{0.7}$ MOSFET at $T = 77$ K in comparison with the characteristics obtained at room temperature $T = 293$ K are shown in Fig. 8. The threshold voltage V_{TH} increases slightly with the temperature decreasing to 77 K. The maximum transconductance g_m and maximum drain current I_D in the linear regime increased 2.8 and 1.6 times, respectively, at 77 K when compared to the corresponding values measured at 293 K. The maximum transconductance g_m and maximum drain current I_D in saturation increased 1.4 and 1.3 times, respectively, at 77 K when compared to the respective values measured at 293 K.

The *C-V* characteristics were measured on devices with gate length $L = 50$ μm and gate width $W = 50$ μm. The p-$Si_{0.2}Ge_{0.8}$(2) MOSFETs with gate oxide thickness of 11 nm operate at the gate voltage range $-6\text{ V} \le V_G \le 6\text{ V}$ before breakdown. Figure 9 clearly shows that the Si cap starts to fill with carriers only at a gate overdrive voltage of 3.5 V. Very small changes in *C-V* curves measured at 77 K and at 293 K indicate low level of impurities in the heterostructure and low concentration of mobile charge inside the gate dielectric.

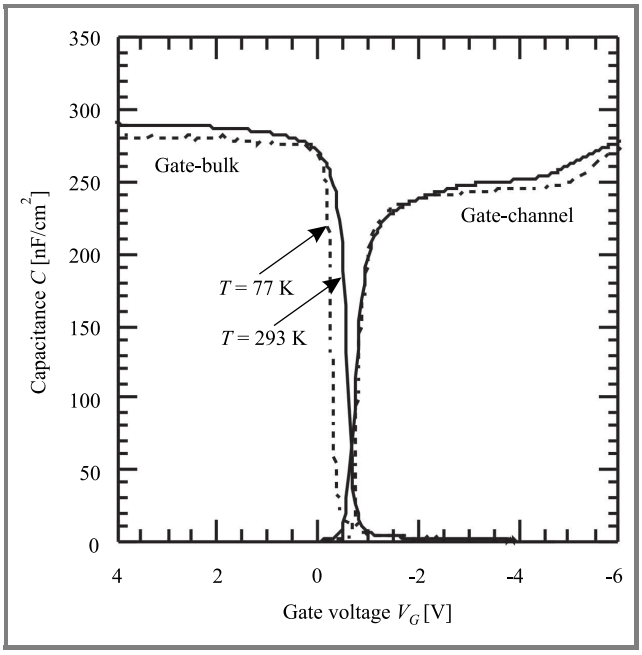


Fig. 9. High frequency split *C-V* characteristics for MOSFET p-$Si_{0.2}Ge_{0.8}$(2) with effective gate length 50 μm at room temperature 293 K (solid lines) and at liquid nitrogen temperature 77 K (dashed lines).

The *C-V* characteristics for p-$Si_{0.3}Ge_{0.7}$ MOSFETs with SiO_2 thickness of 7 nm (Fig. 10) show that the Si cap is not filled up to 2 V gate overdrive voltage. The oxide breakdown limit for p-$Si_{0.3}Ge_{0.7}$ devices is around 4 V.

The depletion charge was extracted from quasistatic and high frequency *C-V* characteristics [11]. Figure 11 shows increasing of the depletion charge for p-$Si_{0.3}Ge_{0.7}$ heterostructure from the depth of 25 nm. The value of the depletion charge is $\sim 3 \cdot 10^{17}$ cm^{-3} at maximum. It corresponded to the n-type doped SiGe buffer layer of 15 nm thickness that lays 5 nm beneath the channel. The depletion charge curve is broken abruptly at the depth of 40 nm, which points to a limitation of the depletion approximation. The depletion regime is changed to the inversion regime (holes accumulation) at this depth. No peculiarity was observed on the depletion charge curves for

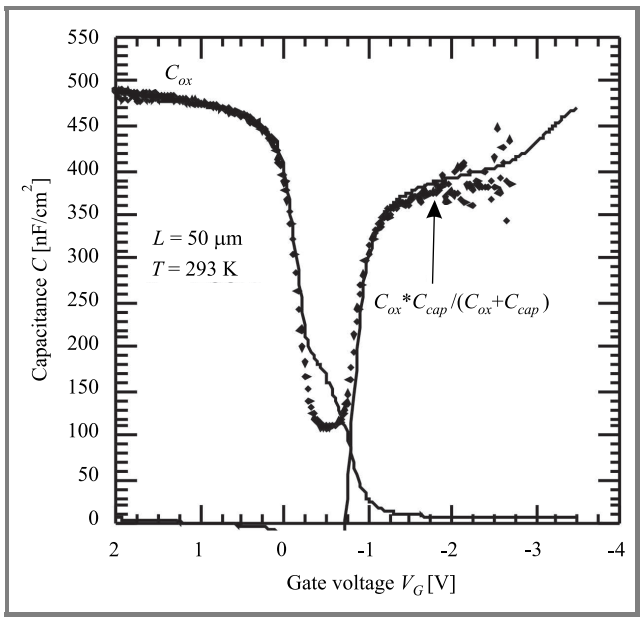


Fig. 10. High frequency (solid line) and quasistatic (dots) *C-V* characteristics for MOSFET p-$Si_{0.3}Ge_{0.7}$ with effective gate length 50 μm at room temperature.

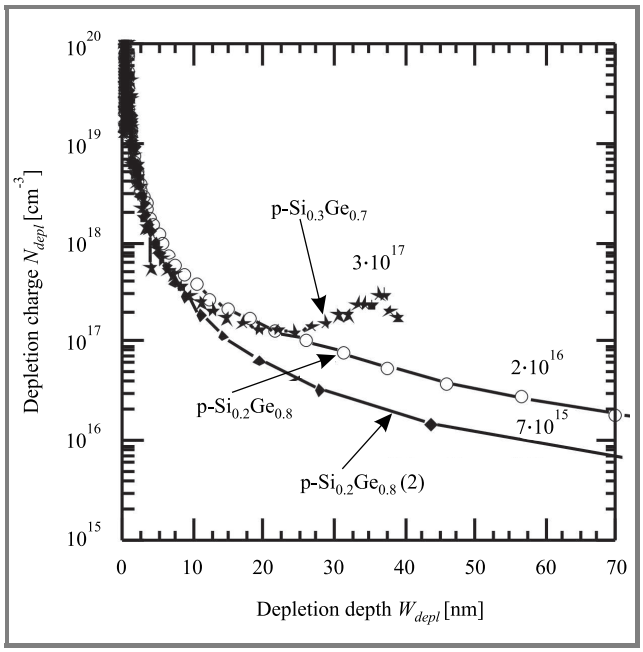


Fig. 11. Depletion charge profiles extracted from quasistatic and high frequency *C-V* characteristics for p-$Si_{0.3}Ge_{0.7}$, p-$Si_{0.2}Ge_{0.8}$, and p-$Si_{0.2}Ge_{0.8}$(2) heterostructures.

p-$Si_{0.2}Ge_{0.8}$ and p-$Si_{0.2}Ge_{0.8}$(2) MOSFETs. This indicates that the under-channel area of these samples was not doped or was doped with background donor concentration less than $1\cdot10^{16}$ cm^{-3}. Also Fig. 11 shows lower impurity background of the p-$Si_{0.2}Ge_{0.8}$(2) heterostructure when compared to the p-$Si_{0.2}Ge_{0.8}$ heterostructure.

The effective mobility μ_{eff} (Fig. 12) has been determined as a function of E_{eff} on large area MOSFETs (gate width W and length L both equal to 50 μm), from the input *I-V* at low $V_{DS} = -50$ mV and from quasistatic and high frequency split *C-V* characteristics [11]. 1D Poisson-Schrodinger simulation was used to obtain correct sheet densities inside the structure and recheck the parameters extracted from the depletion approximation by fitting the measured *C-V* data.

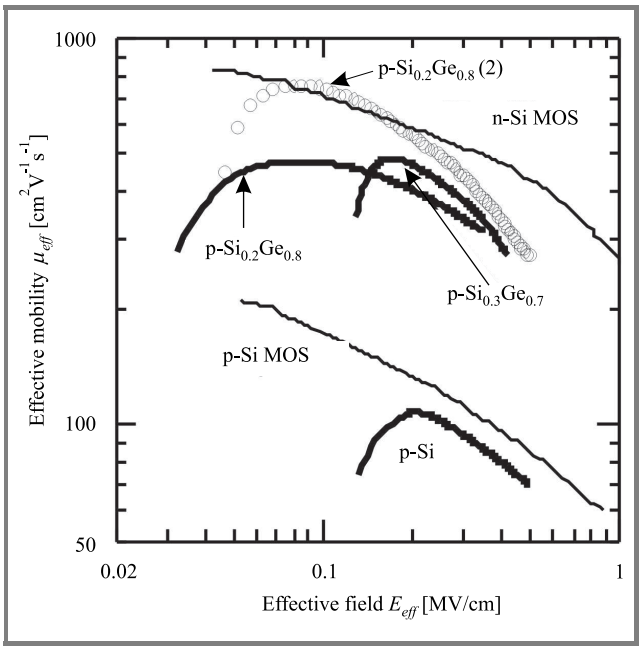


Fig. 12. Effective mobility as a function of effective field for p-$Si_{0.3}Ge_{0.7}$, p-$Si_{0.2}Ge_{0.8}$, p-$Si_{0.2}Ge_{0.8}$(2) and p-Si MOSFETs, as well as universal curves for p-Si MOS and n-Si MOS after S. Takagi [9].

The p-$Si_{0.2}Ge_{0.8}$(2) heterostructure has the highest mobility 760 cm^{2}V^{-1}s^{-1} at the field $E_{eff} = 0.08$ MV/cm due to the lowest background of ionized impurities. On the other hand, p-$Si_{0.2}Ge_{0.8}$(2) MOSFETs have the worst subthreshold slope, that could be explained by MOSFET short channel effects due to the absence of n-type "punch through" stopper (ionized impurities) underneath of p-$Si_{0.2}Ge_{0.8}$ channel.

4. Low-frequency noise

Conventional MOSFET characterisation techniques, such as the combination of *I-V* (current-voltage) and *C-V* (capacitance-voltage) measurements, are very problematic as device size decreases down to the deep sub-μm (DS-μm) scale. "Average per square" characteristic parameters obtained from large-scale devices cannot be suitable for DS-μm MOSFET analysis due to statistical uncertainty of fabrication technology together with the importance of mesoscopic quantum effects. Low-frequency noise measurements could be a powerful diagnostic technique for DS-μm MOSFET characterization in a wide range of device operation regimes [17]. Unfortunately, the commercially available current preamplifiers such as ITHACO-1211, SR-570, EG&G-181 have been optimised only for limited ranges

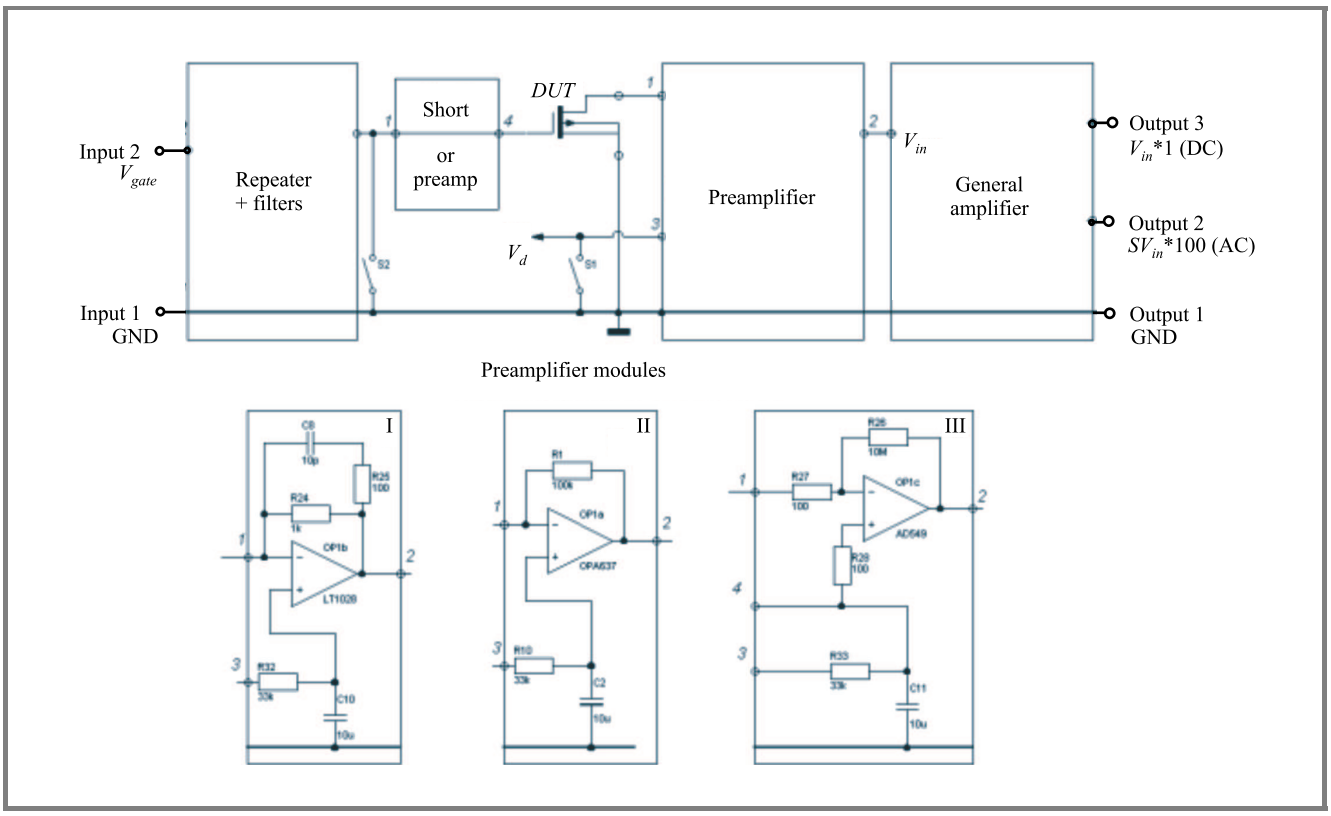


Fig. 13. Schematic of the current preamplifier with modular design and interchangeable first stage for LF-noise measurements.

of device input impedance and their conventional "all-in-one" desktop design also introduces extra problems when long cables are used to connect the equipment to the sample test fixture. To overcome all the above problems we have used the optimised preamplifier modules as the first stages for gate leakage and drain current noise measurements of MOSFETs with input impedance 50 Ω–10^8 Ω in the frequency range of 1.0 Hz–10^5 Hz. A three-box modular design with interchangeable first stage preamplifiers (Fig. 13) was chosen to improve the reliability and to reduce the influence of the connection cables on measurement results. The best operational amplifiers (OAMPs) currently available with optimal voltage v_n and current i_n noise, AD549 ($v_n = 200$ nV Hz$^{-1/2}$, $i_n = 0.15$ fAHz$^{-1/2}$), OPA637 ($v_n = 3.7$ nVHz$^{-1/2}$, $i_n = 2.0$ fAHz$^{-1/2}$) and LT1028A ($v_n = 0.85$ nVHz$^{-1/2}$, $i_n = 1.0$ pAHz$^{-1/2}$) were used for the first stage module at each of the three chosen impedance ranges.

The LF-noise was measured using an HP 35670A dynamic signal analyzer and the custom-made preamplifier described above. Characteristics *I-V* and LF-noise were measured simultaneously to account for possible offset of the applied gate voltage V_G. All measurements were done on MOSFETs with a geometrical gate length of 1.0 μm (an effective gate length was extracted as 0.55 μm) and 10 μm in an electrically shielded room at 293 K. The SiGe MOSFETs show enhancement in the drain current and transconductance at the same gate overdrive voltages in comparison with p-Si devices. LF-noise has been measured in the linear regime of the output *I-V* characteristics ($V_{DS} = -50$ mV), from the sub-threshold through weak to strong inversion ($V_G - V_{TH}$ from 0.5 to −3 V) of the input *I-V*, in a wide range of drain-source conductance $g_d = I_D/V_{DS}$.

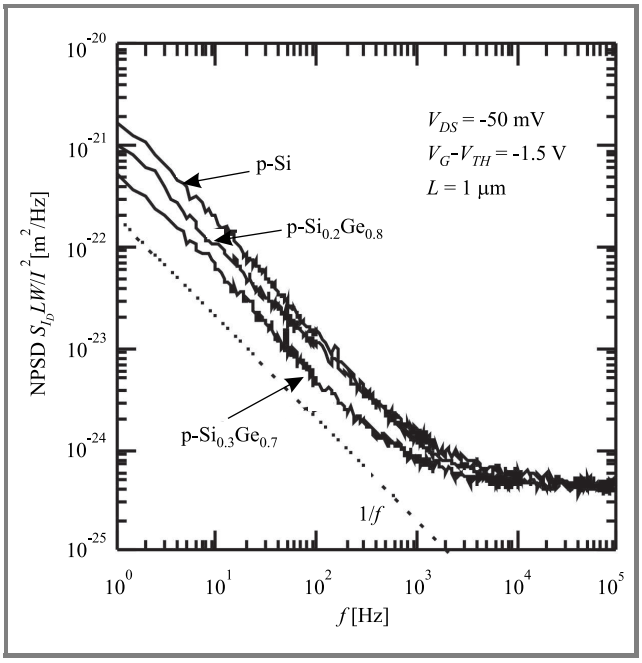


Fig. 14. Normalized power spectral density of drain current fluctuations as a function of frequency for p-$Si_{0.3}Ge_{0.7}$, p-$Si_{0.2}Ge_{0.8}$ and p-Si MOSFETs.

A typical normalized power spectral density (NPSD) S_I/I^2 of drain current fluctuations versus frequency in the range 1–10^5 Hz is presented in Fig. 14. Flicker, $1/f$ component, at low frequencies and thermal noise at high frequency range, dominate the spectra. In Fig. 14 the $1/f$ noise for the p-$Si_{0.3}Ge0.7$ MOSFET is clearly seen to be over three times lower than that for Si. We have not observed a generation-recombination (GR) noise component at any gate overdrive voltage. This is usually manifested as bumps in the spectra. GR noise could appear in the spectra due to Sb diffusion into the $Si_{0.3}Ge_{0.7}$ channel from the Sb-doped "punch-through" stopper or the existence of deep levels in the heterostructure. Thus we can confirm the absence of these defects and contaminations after the full MOSFET fabrication process.

The NPSD S_{I_D} in the $1/f$ region is described in terms of carrier number fluctuations (CNF), correlated mobility fluctuations (CMF) and source-drain series resistance fluctuations (SDRF) [17]:

$$S_{I_D}/I_D^2 = \left(1+\alpha\mu_{eff}CI_D/g_m\right)^2\left(\frac{g_m}{I_D}\right)^2 S_{V_{fb}} + \left(\frac{I_D}{V_{DS}}\right)^2 S_{R_{SD}}, \tag{1}$$

where α is the Coulomb scattering coefficient, μ_{eff} is the effective mobility, $S_{V_{fb}} = S_{Q_{it}}/(WLC^2)$ with $S_{Q_{it}}$ being the interface charge spectral density per unit area, C is the gate oxide capacitance C_{ox}.

The flat band voltage spectral density is defined by [17]:

$$S_{V_{fb}} = \frac{Q^2 k_B T N_{st}}{WLC_{ox}^2 f^\gamma} = \frac{q^2 k_B T \lambda N_t}{WLC_{ox}^2 f^\gamma}, \tag{2}$$

where f is the frequency, γ is the characteristic exponent close to unity, k_BT is the thermal energy, N_{st} is the density of traps near the Si/SiO_2 and/or Si/SiGe interface, λ is the tunnel attenuation distance to Si cap and/or SiO_2, and N_t is the volumetric trap density in the Si cap and/or SiO_2. The spectral density of the source-drain series resistance we defined by:

$$S_{R_{SD}} = \alpha_{H\_SD}\frac{R_{SD}^2}{fN_{SD}} \sim \frac{R_{SD}^3}{f}, \tag{3}$$

where α_{H_SD} is the Hooge parameter for $1/f$ noise in the series resistance, N_{SD} is the total number of free carriers and R_{SD} is the source-drain series resistance.

The CMF can be important in both the weak and strong inversion regions of MOSFET operation. Typically, SDRF can appear at the highest gate voltages for the shortest channel lengths, when the channel resistance becomes comparable to the source-drain series resistance.

Figure 15 shows how measured and calculated power spectral density (PSD) varies with device conductance for the p-Si MOSFET. This curve was very well fitted by CNF, CMF and SDRF using Eq. (1). The Coulomb scattering coefficient $\alpha = 8\cdot10^4$ Vs/C extracted from the fitting of the experimental data for p-Si MOSFET is close to the predicted value of 10^5 Vs/C for holes [17]. It is comparable to $\alpha^{PM\ \mathrm{Si\ cap}}$ for the Si cap of pseudomorphic p-SiGe devices and much higher than that for SiGe channels of the same pseudomorphic p-SiGe MOSFETs $\alpha^{PM\ \mathrm{SiGe}} = \sim 0.1\,\alpha^{PM\ \mathrm{Si\ cap}}$ [12].

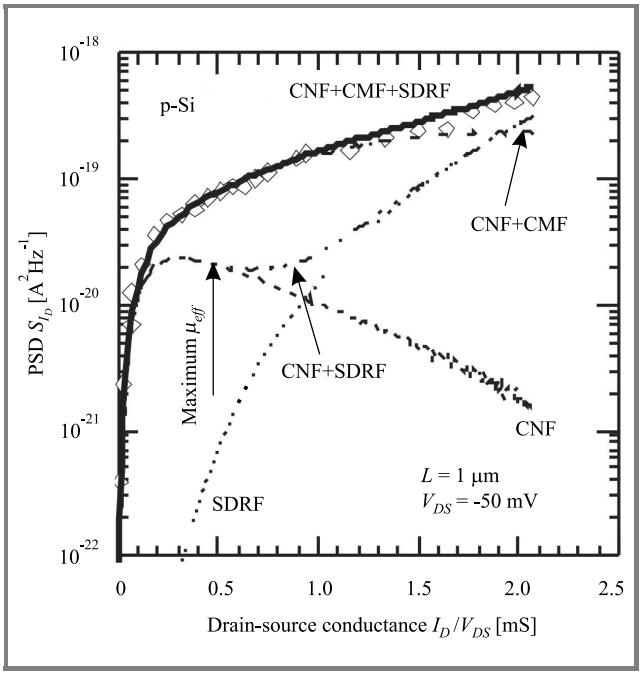


Fig. 15. Power spectral density dependence on device conductance for p-Si MOSFET.

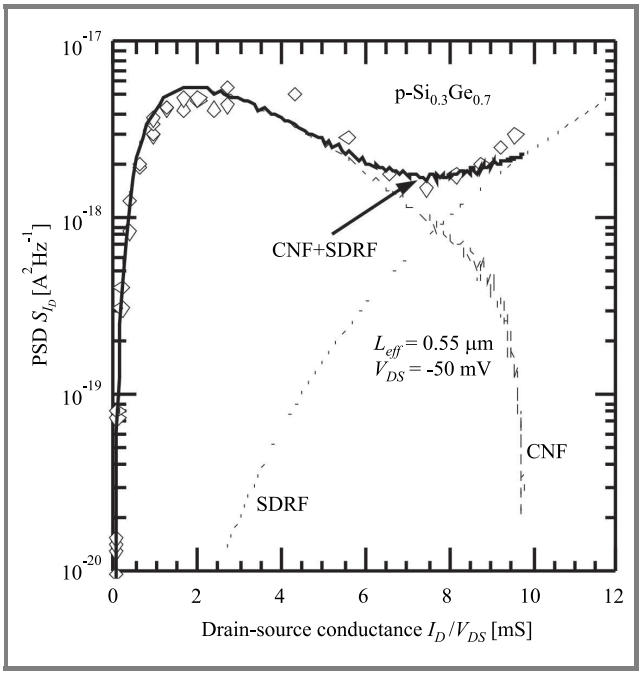


Fig. 16. Power spectral density dependence on device conductance for p-$Si_{0.3}Ge_{0.7}$ MOSFET.

Figures 16 and 17 show the variation of PSD with device conductance for the p-$Si_{0.3}Ge_{0.7}$ and p-$Si_{0.2}Ge_{0.8}$ MOSFETs, respectively.

The variation is explained completely by CNF and SDRF, which reduce Eq. (1) for the NPSD to:

$$S_{I_D}/I_D^2 = \left(\frac{g_m}{I_D}\right)^2 S_{V_{fb}} + \left(\frac{I_D}{V_{DS}}\right)^2 S_{R_{SD}}\,. \qquad (4)$$

In the case of our metamorphic p-SiGe MOSFETs the CMF component was not observed ($\alpha << 5\cdot 10^2$ Vs/C) due to the presence of a thin Si cap layer (4–5 nm) between the SiGe channel and the Si-SiO_2 interface. The CMF component is more important as carriers locate closer to the SiO_2/Si interface. Thus, the signal to noise ratio in conventional MOSFET structure could be significantly improved in the case of heavily doped substrates or introduced punch-through stopper doping if SiGe buried channel heterostructures are used.

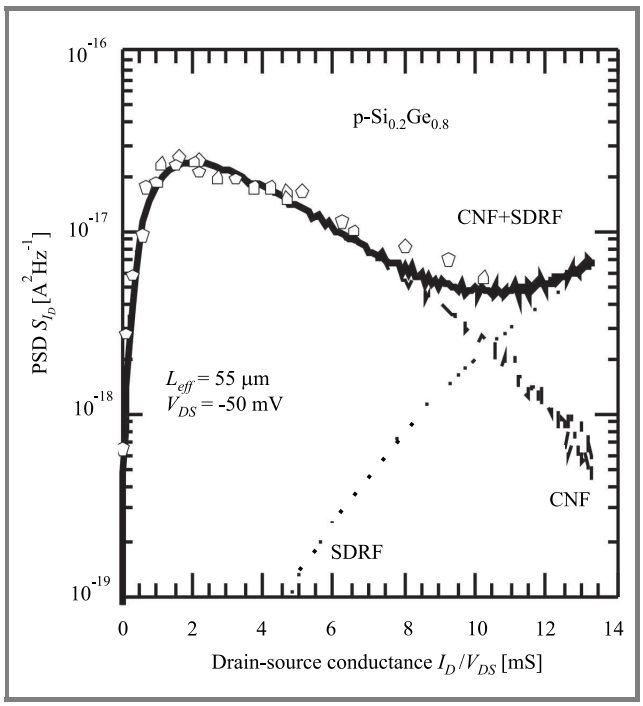


Fig. 17. Power spectral density dependence on device conductance for p-$Si_{0.2}Ge_{0.8}$ MOSFET.

The SDRF component dominated in strong inversion for all the measured devices, and its value is 10–100 times lower in metamorphic p-SiGe MOSFETs than in p-Si due to their lower source-drain access resistance. Contact resistance estimated from the SDRF component decreased with Ge content increasing (Table 2).

Source-drain resistance R_{SD} (noise) was calculated using equation (3) with suggestion that shape of contact areas (L_{SD} – effective length of contacts area) is the same or very similar, and the values of Hooge parameter multiplied to mobility of carriers in contact area $\alpha_{H\_SD} \times \mu_{SD}$ are similar. This product also known as "noise reduced mobility" can be used as quality factor of the material in contact area [20]. The R_{SD} (*I-V*) was obtained as a cross-over point of the lines drawn through points $R(V_G)$ for devices with different gate lengths L at several fixed gate voltages V_G (Terada-Muta method) [19].

Table 2
Source-drain resistance R_{SD} extracted from *I-V* and R_{SD} estimated from LF-noise results

Sample	$S_{R_{SD}}$ [Ω^2Hz^{-1}]	R_{SD}/R_0	R_{SD} (noise) [$\Omega\mu$m]	R_{SD} (*I-V*) [$\Omega\mu$m]
p-$Si_{0.3}Ge_{0.7}$	$1.2\cdot 10^{-7}$	1.00	2025*	2025
p-$Si_{0.2}Ge_{0.8}$	$8.0\cdot 10^{-8}$	0.87	1769	2300
p-Si	$7.0\cdot 10^{-6}$	3.88	7854	4680

* Resistance of 2025 $\Omega\mu$m of the sample p-$Si_{0.3}Ge_{0.7}$ measured using the Terada-Muta method [19] was used as a reference to estimate the resistance from LF-noise measurements.

The R_{SD} extraction procedure from LF-noise requires just one device to be measured and one reference device. The results estimated from LF-noise are applied to the individual measured devices as opposed to the set of devices needed in the Terada-Muta method.

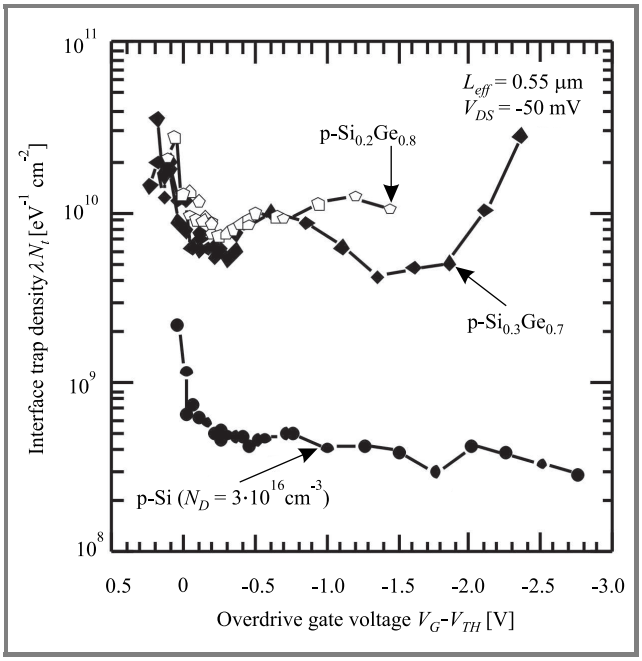


Fig. 18. Interface trap density extracted from fitting of the data supplied by LF-noise measurements versus gate overdrive voltage for p-$Si_{0.3}Ge_{0.7}$, p-$Si_{0.5}Ge_{0.8}$ and p-Si heterostructures at room ($T = 293$ K) temperature.

The average densities of traps λN_t in SiO_2 involved in the trapping-detrapping process and presented in $1/f$ noise are extracted from LF-noise after fitting of all noise components. Figure 18 shows the lowest λN_t for conventional p-Si MOSFET and values higher by an order of magnitude

for p-SiGe MOSFETs. This could be explained by the difference in Si and SiGe fabrication technologies. The quality of SiO_2 for p-SiGe MOSFETs is worse, due to the lower thermal budget required for the whole processing. Also, average densities of traps λN_t extracted from LF-noise are less than values usually obtained from *C-V* characteristics. This could be explained if it is assumed that either not all the traps inside SiO_2 are involved in the trapping-detrapping process or other processes also affect the LF-noise.

5. Conclusions

In conclusion, all these results demonstrate the advantages of metamorphic MOSFETs with a high Ge content and strained $Si_{1-x}Ge_x$ p-channel grown on a relaxed $Si_{1-y}Ge_y$ buffer in comparison with a bulk p-Si MOSFET. Both SS-MBE and LEPECVD grown material shows very significant hole mobility improvement over bulk Si, with peak values of $\mu_{eff} = 760$ cm^{2}V^{-1}s^{-1}. Of the three types devices studied the SS-MBE grown $Si_{0.3}Ge_{0.7}$ structure produces the best performance as a sub-micron MOSFET device, mainly due to the incorporation of an Sb-doped punch-through stopper. This results in a SiGe device with similar to the Si control short channel properties at an effective channel length of 0.55 μm to the Si control. The current drive enhancement ratio of 2.0 over p-Si MOSFET is found in the p-$Si_{0.3}Ge_{0.7}$ MOSFET, and is due to higher hole mobility $\mu_{eff} = 500$ cm^{2}V^{-1}s^{-1} in the $Si_{0.3}Ge_{0.7}$ quantum well. The highest drive current is found in the p-$Si_{0.2}Ge_{0.8}$ MOSFET, with a current drive enhancement ratio of more than 3.0 over the p-Si MOSFET. These studies demonstrate clearly the potential of using strained $Si_{0.3}Ge_{0.7}$ and $Si_{0.2}Ge_{0.8}$ heterostructures for the PMOSFETs in CMOS structure.

Also, the results presented in this paper demonstrate a significant reduction in LF-noise NPSD, achieved in metamorphic p-$Si_{0.3}Ge_{0.7}$ and p-$Si_{0.2}Ge_{0.8}$ MOSFETs compared to bulk p-Si. This advantage is observed in sub-micron devices relevant to the current Si-CMOS technology. In the linear region of MOSFET operation the reduction in $1/f$ noise is higher than a factor of three. The reduction is attributed to the existence of the Si cap layer in the p-SiGe MOSFETs, which further separates the holes in the buried $Si_{0.3}Ge_{0.7}$ and $Si_{0.2}Ge_{0.8}$ channels from the traps near the Si/SiO_2 interface, and an immeasurably low influence of traps at the Si/SiGe interface. LF-noise performance of p-SiGe MOSFETs could be significantly improved after technology of gate dielectric fabrication will be improved.

The influence of a "punch-through" stopper on the device reliability was analysed. It reduces short channel effects in sub-micron developed MOSFETs and provides perfect performance of devices especially in the subthreshold region as it is most important for switching devices (CMOS logic). Also $1/f$ noise is not significantly increased in buried channel p-SiGe devices with a "punch-through" stopper as in conventional p-Si MOSFETs with heavily doped substrate due to a 4–5 nm Si cap used. On the other hand, the introduced "punch-through" stopper slightly decreases the maximum current of the device and increases the influence of the negative effects due to impact ionization in the drain depletion area. These effects could be reduced through the optimization of the contact shape and doping profile. Better results could possibly be obtained using p-SiGe buried channel heterostructures together with SOI technology (analogue of fully depleted SOI MOSFETs [10]).

References

[1] K. Rim *et al.*, "Characteristics and device design of sub-100 nm strained Si N- and PMOSFETs", in *Symp. VLSI Technol.*, Honolulu, USA, 2002, pp. 98–99.

[2] S. Thompson *et al.*, "A 90 nm logic technology featuring 50 nm strained silicon channel transistors", *IEDM*, p. 50, 2002.

[3] A. Sadek, K. Ismail, M. A. Armstrong, D. A. Antoniadis, and F. Stern, "Design of Si/SiGe heterojunction complementary metal-oxide-semiconductor transistors", *IEEE Trans. Electron Dev.*, vol. 43, no. 8, pp. 1224–1232, 1996.

[4] N. Sugii, D. Hisamoto, K. Washio, N. Yokoyama, and S. Kimura, "Performance enhancement of strained-Si MOSFETs fabricated on a chemical-mechanical-polished SiGe substrate", *IEEE Trans. Electron Dev.*, vol. 49, no. 12, pp. 2237–2243, 2002.

[5] G. Höck, E. Kohn, C. Rosenblad, H. von Känel, H.-J. Herzog, and U. König, "High hole mobility in $Si_{0.17}Ge_{0.83}$ channel metal-oxide-semiconductor field-effect transistors grown by plasma-enhanced chemical vapor deposition", *Appl. Phys. Lett.*, vol. 76, no. 26, pp. 3920–3922, 2000.

[6] C. W. Leitz, M. T. Currie, M. L. Lee, Z.-Y. Cheng, D. A. Antoniadis, and E. A. Fitzgerald, "Hall mobility enhancement in strained $Si/Si_{1-y}Ge_y$ p-type metal-oxide-semiconductor field-effect transistors grown on relaxed $Si_{1-x}Ge_x$ ($x < y$) virtual substrates", *Appl. Phys. Lett.*, vol. 79, no. 25, pp. 4246–4248, 2001.

[7] M. Myronov, P. J. Phillips, T. E. Whall, and E. H. C Parker, "Hall mobility enhancement caused by annealing of $Si_{0.2}Ge_{0.8}/Si_{0.7}Ge_{0.3}/Si(001)$ p-type modulation-doped heterostructures", *Appl. Phys. Lett.*, vol. 80, pp. 3557–3559, 2002.

[8] T. Koster, J. Stein, B. Hadam, J. Gondermann, B. Spangenberg, H. G. Roskos, H. Kurz, M. Holzmann, M. Riedinger, and G. Abstreiter, "Fabrication and characterisation of SiGe based in-plane-gate transistors", *Microelectron. Eng.*, vol. 35, no. 1–4, pp. 301–304, 1997.

[9] S. Takagi, A. Toriumi, M. Iwase, and H. Tango, "On the universality of inversion layer mobility in Si MOSFET's: Part I – Effects of substrate impurity concentration", *IEEE Trans. Electron Dev.*, vol. 41, pp. 2351–2356, 1994.

[10] S. Cristoloveanu, "Silicon on insulator technologies and devices: from present to future", *Solid-State Electron.*, vol. 45, pp. 1403–1411, 2001.

[11] D. K. Schroder, *Semiconductor Material and Device Characterization*. 2nd ed. Wiley, 1998, pp. 540–547.

[12] G. Ghibaudo and J. Chroboczek, "On the origin of the LF noise in Si/Ge MOSFETs", *Solid-State Electron.*, vol. 46, pp. 393–398, 2002.

[13] A. D. Lambert, B. Alderman, R. J. P. Lander, E. H. C. Parker, and T. E. Whall, "Low frequency noise measurements of p-channel $Si_{1-x}Ge_x$ MOSFET's", *IEEE Trans. Electron Dev.*, vol. 46, pp. 1484–1486, 1999.

[14] S. Okhonin, M. A. Py, B. Georgescu, H. Fisher, and L. Risch, "DC and low-frequency noise characteristics of SiGe p-channel FET's designed for 0.13-μm technology", *IEEE Trans. Electron Dev.*, vol. 46, pp. 1514–1517, 1999.

[15] A. Asai, J. S. Iwanaga, A. Inoue, Y. Hara, Y. Kanzawa, H. Sorada, T. Kawashima, T. Ohnishi, T. Takagi, and M. Kubo, "Low-frequency noise characteristics in SiGe channel heterostructure dynamic threshold PMOSFET (HDTMOS)", *IEDM Tech. Dig.*, pp. 35–38, 2002.

[16] T. Tsuchiya, T. Matsuura, and J. Murota, "Low-frequency noise in $Si_{1-x}Ge_x$ p-channel metal oxide semiconductor field-effect transistors", *Jpn. J. Appl. Phys*., vol. 40, pp. 5290–5293, 2001.

[17] G. Ghibaudo and T. Boutchcha, "Electrical noise and RTS fluctuations in advanced CMOS devices", *Microelectron. Reliab.*, vol. 42, pp. 573–582, 2002.

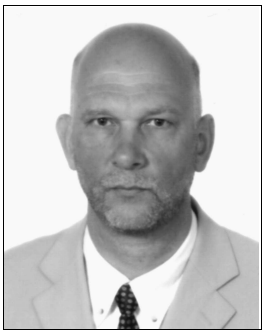
[18] M. Glück, J. Hersener, H. G. Umbach, J. Rappich, and J. Stein, "Implementation of low thermal budget techniques to Si and SiGe MOSFET device processing", *Solid-State Phen.*, vol. 57–58, p. 413, 1997.

[19] K. Terada and H. Muta, "A new method to determine effective MOSFET channel length", *Jpn. J. Appl. Phys.*, vol. 18, p. 953, 1979.

[20] M. M. Jevtic, Z. Stanimirovic, and I. Stanimirovic, "Evaluation of thick-film resistor structural parameters based on noise index measurements", *Microelectron. Reliab.*, vol. 41, p. 59, 2001.

Sergiy Durov was born in Kiev, Ukraine, in 1976. He received the B.Sc. and M.Sc. degrees in physics from National Taras Shevchenko University, Kiev, Ukraine, in 1998. From 1998 to 2000 he worked on theoretical research of carbon nanoclusters. Since 2001 he has been working towards Ph.D. in semiconductor physics in the University of Warwick, Coventry, UK. His main research interests concentrated in nanotechnology, physics of SiGe heterostructures and carbon nanoparticles, quantum chemistry calculations, electrical noise measurements and analysis.

e-mail: snl_@mail.ru
Department of Physics
University of Warwick
Coventry CV4 7AL, United Kingdom



Oleg A. Mironov received the M.Sc. degree in radiophysics and electronics from Kharkov State University, Kharkov, Ukraine, in 1971 and the Ph.D. degree in solid state physics from Kharkov Technical University in 1986. During 1971–2001 he was a Researcher/Senior Researcher at the Institute for Radiophysics and Electronics National Academy of Sciences of Ukraine, Kharkov. In 1993 he joined SiGe/Nano Silicon Research Group of Professor E. H. C. Parker at the Department of Physics, University of Warwick, Coventry, UK. His main research interests are focused on the physics of semiconductor heterostructures and devices, including high magnetic field and low temperature transport, cryogenic magnetic field and temperature sensors, structural and optical characterisation. Since 2000 he is engaged in *I-V*, *C-V* and low frequency noise studies of Si and pseudo-metamorphic SiGe MOSFETs. He is the co-author of more than 87 scientific publications and 3 patents.

e-mail: O.A.Mironov@warwick.ac.uk
Department of Physics
University of Warwick
Coventry CV4 7AL, United Kingdom

Maksym Myronov was born on March 2, 1974 in Kharkov, Ukraine. He was awarded M.Sc. in physics and mathematics from Kharkov State University, Kharkov, Ukraine, in 1996 and Ph.D. in physics from University of Warwick, Coventry, UK in 2001. His main research activities concerned epitaxial growth of SiGe heterostructures and their electrical, structural and optical characterization. From 2001 he has been working as a Research Fellow at the Physics Department, University of Warwick, Coventry, UK. His main research interests are electrical, structural and optical characterization of nanoscale SiGe heterostructures and SiGe FET devices.

e-mail: Maksym_Myronov@yahoo.co.uk
Department of Physics
University of Warwick
Coventry CV4 7AL, United Kingdom

Terence E. Whall was awarded an honours B.Sc. from City University (London 1965) and a Ph.D. from Sussex University (1970). His major field of study has been the electrical properties of solids, including Kondo/spin glass alloys, molecular solids and transition metal oxides. Professor Whall holds a personal chair in physics at Warwick University, where he heads the device physics activity looking at Si/SiGe MOSFETs. He is the co-author of ca 200 research publications and has given 15 invited papers at international conferences. He is a member of the (UK) Institute of Physics.

e-mail: T.E.Whall@warwick.ac.uk
Department of Physics
University of Warwick
Coventry CV4 7AL, United Kingdom

Thomas Hackbarth was born on August 28, 1958 in Stade, Germany. He received the diploma and Ph.D. degrees in electrical engineering from Technical University of Braunschweig, Germany, in 1986 and 1991, respectively. In 1991, he joined the Research and Technology Department of DaimlerChrysler in Ulm, Germany, where his work was molecular beam epitaxy for vertical cavity lasers and high electron mobility transistors in the III-V material system. His current interest is layer growth and device processing of SiGe modulation doped field effect transistors and integrated circuits for ultra-high frequency applications.

e-mail: thomas.hackbarth@daimlerchrysler.com
DaimlerChrysler Research Center
Wilhelm-Runge Str. 11
D-89081 Ulm, Germany

Georg Hoeck was born in Krumbach, Germany, in 1968. He received the diploma in physics from the Technical University of Munich, Germany, in 1996 and the Ph.D. degree in electrical engineering from the University of Ulm, Germany, in 2004. His research activities concerned manufacturing and characterization of p-channel SiGe heterostructure field effect transistors. Currently he is with Siemens AG in Ulm, Germany, working in the field of mobile communication networks.

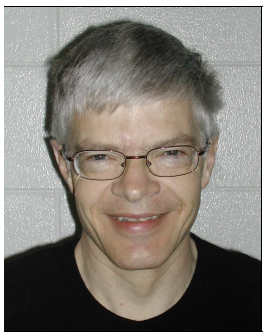
e-mail: georg.hoeck@siemens.com
Siemens AG
Lise-Meitner-Str. 5
D-89081 Ulm, Germany

Hans-Joest Herzog was born on September 15, 1943 in Rottweil, Germany. He received the diploma in physics from the University Stuttgart, Germany, in 1973. In 1973, he joined the AEG Telefunken Research Center Ulm, Germany, which merged in the Daimler-Benz Research Center Ulm in 1990, and in the DaimlerChrysler Research Center Ulm in 1998. His main activities concerned TEM, XRD, and electrical characterization of heteroepitaxy structures. Currently, he is engaged in the Si/SiGe material system and the preparation of novel and high speed devices. In 1994, Hans-Joest Herzog received the SSDM paper award.

e-mail: hans-joest.herzog@daimlerchrysler.com
DaimlerChrysler Research Center
Wilhelm-Runge Str. 11
D-89081 Ulm, Germany

Ulf König received the diploma in physics in 1970 and his Ph.D. in electrical engineering in 1973 at the RWTH Aachen, Germany. In 1975 he joined the AEG which was later on taken over by DaimlerChrysler. Since 1992 he is heading a department with focus on SiGe growth, processes, devices and circuits. He was responsible for the early transfer of the SiGe hetero bipolar transistor technology into production and is presently promoting the introduction of the strained-silicon hetero field effect transistors into Si-lines. Novel activities in his group concern the use of nanotechnology for automotive applications.

e-mail: ulf.koenig@daimlerchrysler.com
DaimlerChrysler Research Center
Wilhelm-Runge Str. 11
D-89081 Ulm, Germany



Hans von Känel received the diploma in physics and the Ph.D. degree in natural sciences at the ETH Zürich, Switzerland, in 1974 and 1978, respectively. From 1979 to 1981 he was a post-doctoral associate at the Physics Department of the Massachusetts Institute of Technology in Cambridge, USA. He was a member of the Research Staff of the Solid State Physics Laboratory, ETH Zürich from 1981 to 2002. In 2002 he assumed a faculty position at the Politecnico di Milano. In 1990 he completed a habilitation on "Epitaxial metal silicides on Si(111)" at the ETH Zürich. He was the scientific leader of a research group on Si-heterostructures at the ETH from 1985 to 2002. His current interests are scanning tunneling microscopy and spectroscopy on low-dimensional systems and new epitaxial deposition techniques, among which he invented the record breaking technique of LEPECVD for ultrafast deposition of thin semiconducting films. Professor Hans von Känel co-authored more than 250 scientific publications and 4 patents.

e-mail: vkaenel@solid.phys.ethz.ch
INFM and L-NESS Dipartmento di Fisica
Politecnico di Milano
via Anzani 52
I-22100 Como, Italy

Paper

Photoelectric measurements of the local values of the effective contact potential difference in the MOS structure

Andrzej Kudła, Henryk Przewłocki, Danuta Brzezińska, and Lech Borowicz

Abstract—We have shown that using focused UV laser beam in photoelectric methods it is possible to measure local ϕ_{MS} values over the gate area of a single MOS structure. The ϕ_{MS} distribution is such that its values are highest far away from the gate edges regions, lower in the vicinity of gate edges and still lower in the vicinity of gate corners. Examples of measurement results and description of the measurement system are presented. The dependence of the ϕ_{MS} value on the exposure time and the power density of UV light is discussed.

Keywords*— *MIS structure, photoelectrical methods, internal photoemission, contact potential difference.

1. Introduction

It is known that mechanical stress in the metal-oxide-semiconductor (MOS) system affects its electrical parameters (e.g., [1, 2]), and that its distribution under the gate of a MOS structure is not uniform [2–4]. Therefore one can expect that local values of MOS parameters in the vicinity of gate edges and are different from those measured far away from the edge. These expectations are confirmed by measurement results showing that the values of both effective contact potential difference (ECPD) [5], referred to as the ϕ_{MS} factor, and flat band voltage V_{FB} depend on the perimeter-to-area ratio.

2. Theory

The contact potential difference in a metal-insulator-semiconductor (MIS) structure is defined by the following formula:

$$\phi_{MS} \stackrel{def}{=} \phi_M - \left(X + \frac{E_G}{2q} + \phi_F \right), \tag{1}$$

where ϕ_M is the potential barrier height for internal photoemission from metal to dielectric, X is the electron affinity at the semiconductor – dielectric interface, E_G is the band gap energy, q is the electron charge and ϕ_F is the Fermi level in the semiconductor.

The gate voltage V_G is divided between the individual regions of the MIS structure according to:

$$V_G = V_I + \phi_S + \phi_{MS}, \tag{2}$$

where V_I is the voltage drop across the dielectric and ϕ_S is the surface potential in the semiconductor.

The voltage V_G^0 corresponding to null photocurrent may be precisely determined using the photoelectric method [5]. When V_G^0 is applied to the gate V_I is close to zero if the diffusion photocurrents from the substrate and gate are equal. Usually ϕ_S may be calculated from *CV* characteristics; however, if the substrate is highly doped, the surface potential is close to zero, and V_G^0 is equal to ϕ_{MS}.

3. Measurement setup

Traditionally a xenon lamp with monochromator is used as the source of UV radiation. In such conditions it is difficult to exceed 1 mW at $\lambda = 244$ nm, simultaneously keeping the spot diameter below 1 mm and the band pass below 5 nm. Application of a laser as a UV source in photoelectric measurements of MIS structures has opened new possibilities. With UV power $P = 100$ mW at $\lambda = 244$ nm it is possible to measure samples with gates exhibiting higher absorption (thicker or made of materials with higher extinction coefficient). Moreover, it is possible to focus UV light better and measure local internal photoemission. However, higher power densities create numerous phenomena that have to be taken into account, such as electron trapping in insulator or photoemission from the semiconductor conduction band.

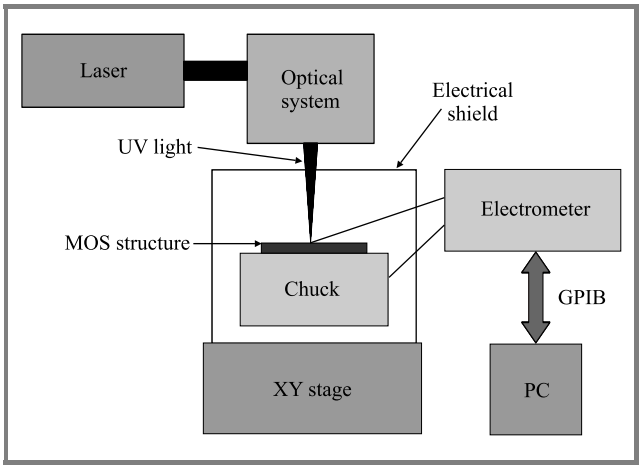


Fig. 1. Block diagram of the system for laser-based photoelectric measurements.

A system for local ϕ_{MS} photoelectrical measurement is shown in Fig. 1. In this system argon laser with a frequency doubler is used as a high-intensity UV-radiation source.

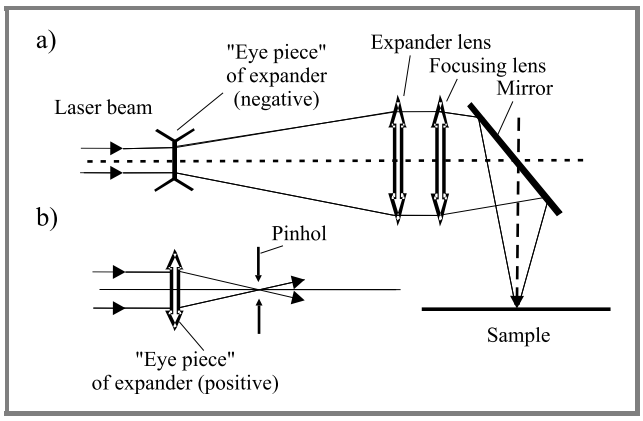


Fig. 2. Diagram of the optical subsystem: (a) with expander, Galileo configuration; (b) eye piece of the Kepler configuration with space filter.

To obtain a low diameter of the light spot a special optical subsystem has to be used (Fig. 2). One simple solution is the Galileo configuration that enables allows the spot diameter to remain in the range of 20 μm. The Kepler configuration reduces the spot diameter below 1 μm.

4. Experimental

Samples containing Al-SiO_2-Si MOS structures were investigated. Highly doped n-type wafers with the orientation of (100) were used. The oxide thickness t_{ox} was 60 nm, while the aluminum gate thickness t_{Al} was either 35 nm or 400 nm.

Argon ion laser with a frequency doubler was used as a UV-radiation source with the wavelength of 244 nm. The UV light spot diameter was 20–40 μm. The power of the UV beam at laser output was adjustable up to 100 mW.

5. Results

Lateral distribution of the ϕ_{MS} on MOS structure can be measured using high intensity, focused laser light (Fig. 3). It was found that the ϕ_{MS} values were the highest far away from the gate-edge regions (e.g., in the middle of a square gate), lower in the vicinity of gate edges, and still lower in the vicinity of gate corners. This result observed directly in photoelectric measurements was confirmed indirectly by purely electrical measurements.

The values of the parameters depend also on the structure dimensions (Fig. 4).

The dependence of the ϕ_{MS} on the structure size supports the assumption pointing mechanical stresses as the origin of the ϕ_{MS} changes. At higher light-power densities electron trapping in SiO_2 and emission from the conduction band in the semiconductor should be considered.

The influence of UV power on V_G^0 is visible in Fig. 5. With higher UV power V_G^0 increases. This could be the effect of photoemission from the silicon conduction band, which decreases electron affinity X (see formula (1)). Other important phenomena are charge trapping in the SiO_2 layer, resulting in V_G^0 increase, and resistance changes during measurements shown in Fig. 6.

Since the photocurrent decreased with increasing V_G^0 during the measurements illustrated in Fig. 6, it is reasonable to assume that ϕ_M has increased because electrons were trapped near the Al-SiO_2 interface. Electron trapping caused by UV radiation is a local property and it has

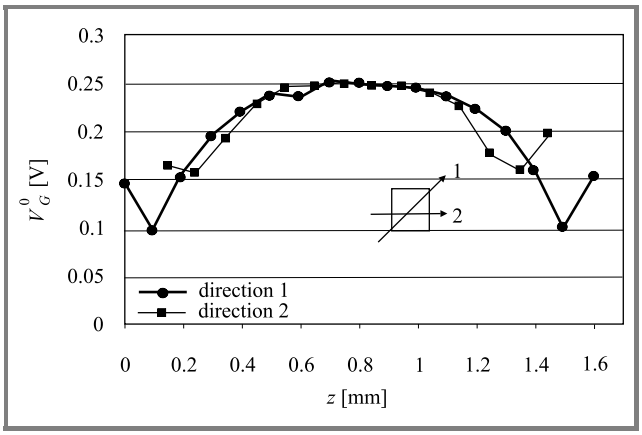


Fig. 3. The local V_G^0 values measured along two directions on a single MOS structure with a square gate.

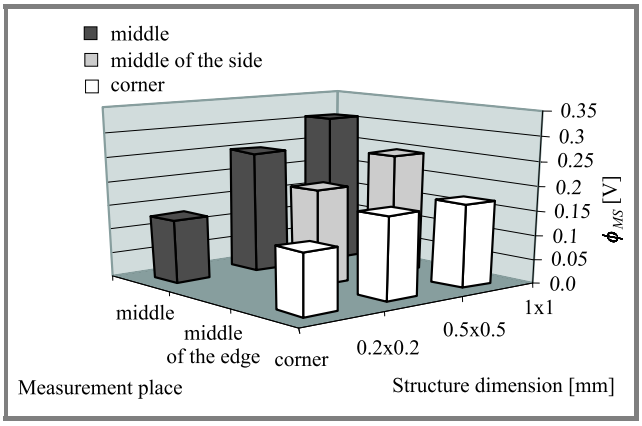


Fig. 4. The ϕ_{MS} values measured on different-size structures (fabricated on the same wafer) at the structure center, at the center of one side and in the corner.

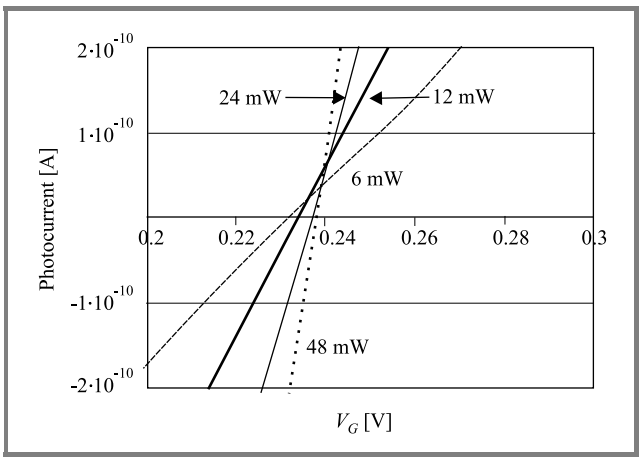


Fig. 5. *I-V* characteristics measured on a structure with $t_{\mathrm{Al}} =$ 35 nm with UV power as a parameter.

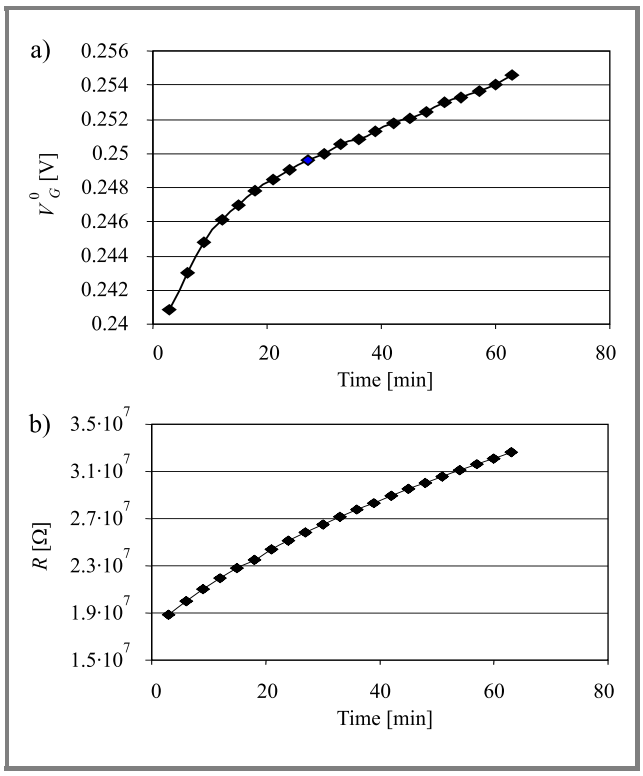


Fig. 6. Changes of the electrical properties of a MOS structure ($t_{\mathrm{Al}} = 400$ nm) during long photoelectric measurements (UV power= 30 mW): (a) V_G^0 voltage and (b) structure resistivity.

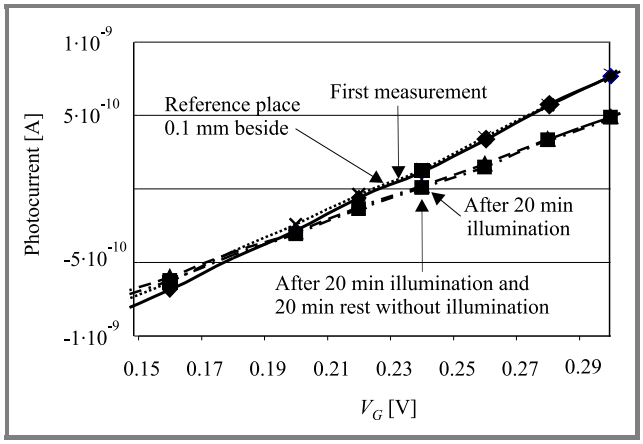


Fig. 7. *I-V* characteristics of a MOS structure with $t_{\mathrm{Al}} = 35$ nm measured before and after 20 min illumination at $P = 6$ mW and compared with the measurements taken in the same place after a 20 min period without illumination and in the reference place (its distance from the test place is 0.1 mm).

a lasting effect on the *I-V* characteristics, as shown in Fig. 7. With the light spot diameter below 20 μm and aluminum thickness of 35 nm electron trapping effects were observed down to UV power of approximately 500 μW.

6. Conclusions

Application of a laser allows local MOS structure properties to be measured by means of photoelectric techniques.

The ϕ_{MS} distribution over the square gate of a MOS structure is such, that its values are highest far away from the gate-edge regions, lower in the vicinity of gate edges, and still lower in the vicinity of gate corners.

Laser beam can cause electron trapping and emission from the conduction band of the semiconductor. These phenomena may be investigated in photoelectric measurements with a laser as a UV-radiation source.

Comparison of the physical properties of differently shaped structures could be misleading due to edge effects. Measuring local parameter values one can avoid averaging values over central and edge regions of the gate and thus enhance measurement accuracy.

References

[1] C. H. Bjorkman, J. T. Fitch, and G. Lucovsky, "Correlation between midgap interface state density and thickness-averaged oxide stress and strain at SiO_2 interfaces formed by thermal oxidation of Si", *Appl. Phys. Lett.*, vol. 56, pp. 1983–1985, 1990.

[2] S. M. Hu, "Stress related problems in silicon technology", *J. Appl. Phys.*, vol. 70, pp. R53–R80, 1991.

[3] I. De Wolf, H. E. Maes, and S. K. Jones, "Stress measurements in silicon devices through Raman spectroscopy: bridging the gap between theory and experiment", *J. Appl. Phys.*, vol. 79, pp. 7148–7156, 1996.

[4] K. F. Dombrowski, I. De Wolf, and B. Dietrich, "Stress measurements using ultraviolet micro-Raman spectroscopy", *Appl. Phys. Lett.*, vol. 75, pp. 2450–2451, 1999.

[5] H. M. Przewłocki, "The importance, the nature and the measurements methods of the ϕ_{MS} factor in MOS structures", *Electron Technol.*, vol. 27, no. 1, pp. 27–42, 1994.

Lech Kazimierz Borowicz was born in Troki near Vilnus in 1934. He received the M.Sc. and Ph.D. degrees in electron technology in 1957 and 1967, respectively. He worked in five research institutes as Assistant Professor, as well as in industry, where he was involved in the design of equipment for spectral measurements. In 2001 he received a Silver Medal in the "Medicine and biotechnology" category at the International Exhibition of Inventions and Innovations for the retinometer – a laser device for estimation of potential visual acuity in the case of patients with cataract. He is the author and co-author of several tens of papers and other publications devoted to instrumental and physiological optics and optoelectronics, as well as organization and management of research.
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Andrzej Kudła, Henryk P. Przewłocki, Danuta Brzezińska – for biography, see this issue, p. 39.

Paper

Effects of stress annealing on the electrical and the optical properties of MOS devices

Witold Rzodkiewicz, Andrzej Kudła, Zbigniew Sawicki, and Henryk M. Przewłocki

Abstract—In this paper we show the results of a study of the effects of high-temperature stress annealing in nitrogen on the refraction index of SiO_2 layers and electrical properties in metal-oxide-semiconductor (MOS) devices. We have experimentally characterized the dependence of the reduced effective contact potential difference (ECPD), the effective oxide charge density (N_{eff}), and the mid-gap interface trap density (D_{it}) on the annealing conditions. Subsequently, we have correlated such properties with the dependence of the refraction index and oxide stress on the annealing conditions and silicon dioxide thickness. Also, the dependence of mechanical stress in the Si-SiO_2 system on the oxidation and annealing conditions has been experimentally determined. We consider the contributions of the thermal-relaxation and nitrogen incorporation processes in determining changes in the SiO_2 layer refractive index and the electrical properties with annealing time. This description is consistent with other annealing studies carried out in argon, where only the thermal relaxation process is present.

Keywords— stress, MOS, Si-SiO_2 system, electrical parameters, refractive index.

1. Introduction

The annealing of metal-oxide-semiconductor (MOS) devices in nitrogen at high temperatures is broadly used to reduce fixed-charge densities at the Si-SiO_2 interface [1]. We have studied the effect of high-temperature annealing in nitrogen on the stress properties, the optical properties, and the electrical properties of MOS devices. In this paper, we report our findings of the influence of annealing in nitrogen on the basic electrical parameters of MOS devices, namely the effective charge density N_{eff}, the mid-gap interface trap density D_{it}, and the reduced effective contact potential difference ϕ^*_{MS}, as well as on the thickness-averaged stress in SiO_2 layers and on the refractive index of such layers. These investigations have been extensively described in [2, 3].

2. Experimental details

In this study 4-inch (100) n-type silicon wafers of different resistivities were used. High-resistivity wafers (3–5 Ω/cm) were doped with phosphorus, while low-resistivity (0.005–0.02 Ω/cm) ones were antimonium-doped. After an initial hydrogen-peroxide-based cleaning sequence, the wafers were subjected to a thermal oxidation process at 1000°C in order to grow silicon-dioxide layers with the thickness of approximately 20, 60, and 160 nm. The wafers were subsequently annealed in nitrogen for periods of 0, 120, and 1440 minutes at 1050°C. The front-side metallization was deposited in a thermal evaporator. The thickness of the obtained aluminum layer was approximately 35 and 400 nm. The thin front-side Al is necessary for MOS photoelectric measurements. The front-side Al was patterned using optical lithography. The backside oxide was etched prior to the deposition of backside metallization. The post-metallization annealing was carried out at 450°C for 20 minutes.

The oxide thickness t_{ox} and its refractive index n were determined using either a Gaertner-scientific 115B single-wavelength ellipsometer operating at $\lambda = 632.8$ nm or a J. A. Woollam variable angle spectroscopic ellipsometer (VASE). The analysis of spectroscopic ellipsometry data by means of a model consisting of a silicon substrate, SiO_2-Si interface layer and silicon dioxide layer was carried out. The interface layer thickness was assumed to be 1 nm. The curvature radius R of the wafer was measured using a Tencor FLX-2320 system.

The relationship between the refractive index n and the stress σ_{ox} in the oxide is given by [4–6]:

$$n(\sigma_{ox}) = n_0 + \Delta\sigma_{ox} \cdot \frac{\Delta n}{\Delta\sigma_{ox}}, \tag{1}$$

where: n_0 is the refractive index of a completely relaxed (stress-free) oxide. The value of n_0 is 1.46 [6–8]. The oxide stress σ_{ox} is calculated in this work using Stoney's formula [9–12]:

$$\sigma_{ox} = \frac{1}{6 \cdot R} \cdot \frac{E_{\mathrm{Si}}}{(1-\nu_{\mathrm{Si}})} \cdot \frac{t_{\mathrm{Si}}^2}{t_{ox}}, \tag{2}$$

where: t_{Si} is the silicon wafer thickness, E_{Si} – Young modulus for silicon, and ν_{Si} – Poisson's ratio for silicon. The ratio of $E_{\mathrm{Si}}/(1-\nu_{\mathrm{Si}})$ is 180.5 GPa for (100) silicon [13, 14].

The effective contact-potential difference ϕ_{MS} was measured with the accuracy of ±10 mV using the photoelectric method [15] implemented in the PIE MSPS photoelectric measurement system. The mid-gap interface trap density D_{it} was measured by means of the HF-LF method using a Keithley PKG 82 system. The general characterization of the electrical properties of the MOS structures was carried out using a SSM 450i system.

The effective contact potential difference (ϕ_{MS} or ECPD) is defined as [1]:

$$\phi_{MS} \equiv \phi_M - \left(\chi_{\mathrm{Si}} + \frac{E_{g,\mathrm{Si}}}{2q} + \phi_F\right), \tag{3}$$

where: ϕ_M is the barrier height at the gate-dielectric interface, χ_{Si} – the electron affinity in the silicon substrate, $E_{g,\mathrm{Si}}$ – the silicon band gap, and ϕ_F – the Fermi level in the silicon substrate measured from the mid-gap level. The effective contact potential difference is determined from [15]:

$$\phi_{MS} = V_{G0} = \phi_{S0}\,, \tag{4}$$

where: V_{G0} is the gate voltage corresponding to zero potential drop across the dielectric, which is determined by the photoelectric method, and ϕ_{S0} the silicon surface potential when $V_G = V_{G0}$.

The reduced effective contact potential difference ϕ^*_{MS} is defined as:

$$\phi^*_{MS} \equiv \phi_M - \chi_{\mathrm{Si}} \tag{5}$$

or

$$\phi^*_{MS} = \phi_{MS} + \frac{E_{g,\mathrm{Si}}}{2q} + \phi_F\,. \tag{6}$$

The electrical parameter described above depends on the barrier heights on both sides of the dielectric but is independent of the doping concentration in the substrate. On the contrary, ϕ_{MS} depends on the doping concentration in the substrate.

The density of the effective charge N_{eff} or Q_{eff}/q is calculated from:

$$Q_{eff} = C_{ox}\left(\phi_{MS} - V_{FB}\right), \tag{7}$$

where: C_{ox} – oxide capacitance per unit area. C_{ox} is determined from the measured capacitance $C_{acc}(F)$ in accumulation using $C_{ox} = C_{acc}/A$, where A is the device area.

3. Results and discussion

The results of the measurements carried out on samples with $t_{ox} \approx 20$, 60 and 160 nm have been shown in Figures 1–10. We must emphasize that the above mentioned results are based on the premise that the mechanical stress in the wafers prior to processing is not crucial. Therefore it is assumed that it is initially the same in all wafers. All results presented in our paper are fairly consistent with this assumption.

Observations and measurements carried out on the wafers with silicon dioxide thickness of about 20 nm indicate that wafer deformation due to the oxidation process doesn't occur in the form of a spherical cap. This is due to the fact that the silicon dioxide layer is too thin to effectively deform the wafer. The stress measurements were carried out both before and after the backside oxide layers were completely removed from the back side of the wafers.

The relationship between the average compressive stress σ_{ox} in the silicon dioxide layer and its thickness is shown in Fig. 1.

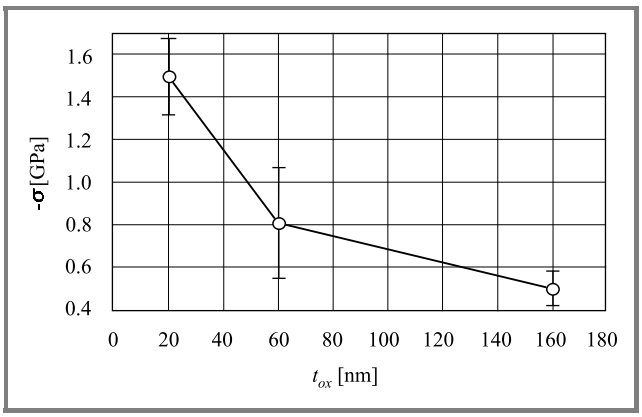


Fig. 1. The average compressive oxide stress $-\sigma_{ox}$ in the SiO_2 layer versus oxide thickness t_{ox}.

As expected, the magnitude of the compressive stress in the silicon dioxide increases as the SiO_2 thickness is decreased. The dependence of SiO_2 refractive index n_{SE} obtained by spectroscopic ellipsometry on the silicon dioxide thickness is portrayed in Fig. 2.

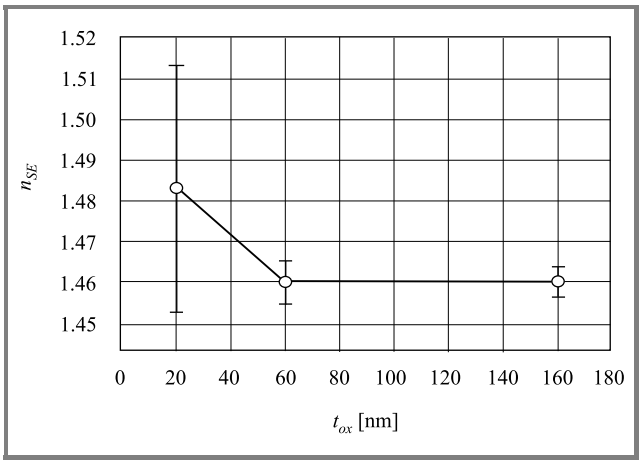


Fig. 2. SiO_2 refractive index n versus oxide thickness t_{ox}.

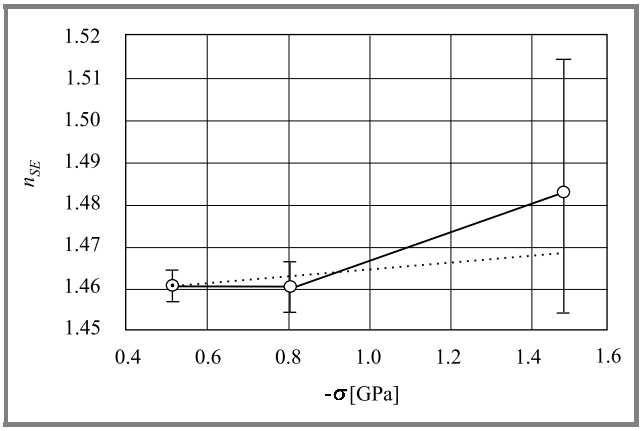


Fig. 3. The average refractive index of SiO_2 as a function of average silicon dioxide compressive stress.

By combination of the results from Figs. 1 and 2, we obtain Fig. 3, which shows the relationship between the average refractive index of SiO_2 and average compressive stress in the oxide. In this plot we point out that the refractive index increases with an increase of the compressive stress.

The effects of extended annealing in nitrogen on the average compressive stress σ_{ox} and the refractive index of SiO_2 are shown in Figs. 4 and 5. In these figures, we observe, that σ_{ox} and n have a tendency to decrease as a result of prolonged anneals due to stress relaxation in the oxide layers by viscous flow and/or decompaction. However, the increase of the oxide refractive index observed for longer annealing times results from nitrogen incorporation in the SiO_2 layer in the vicinity of the Si-SiO_2 interface.

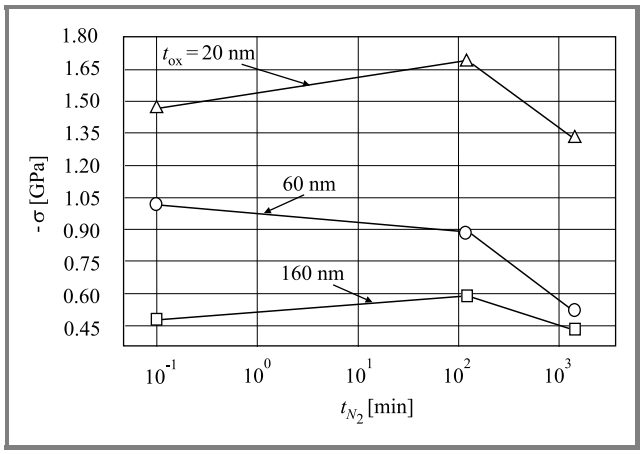


Fig. 4. The average compressive oxide stress $-\sigma_{ox}$ in the SiO_2 layer versus nitrogen-annealing time t_{N_2} for three silicon dioxide thicknesses of 20, 60, and 160 nm.

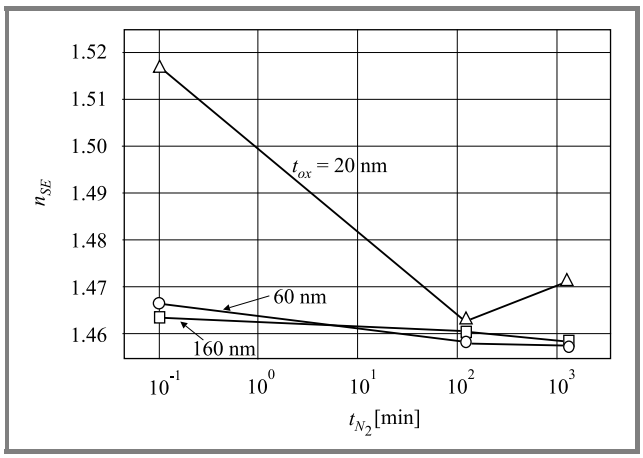


Fig. 5. SiO_2 refractive index n versus nitrogen-annealing time t_{N_2} for three silicon dioxide thicknesses of 20, 60, and 160 nm.

In Fig. 6 the dependence of the mid-gap interface trap density D_{it} on the duration of annealing in nitrogen t_{N_2} is shown. This parameter is particularly sensitive to both the nitrogen-annealing time and the oxide thickness (Fig. 7).

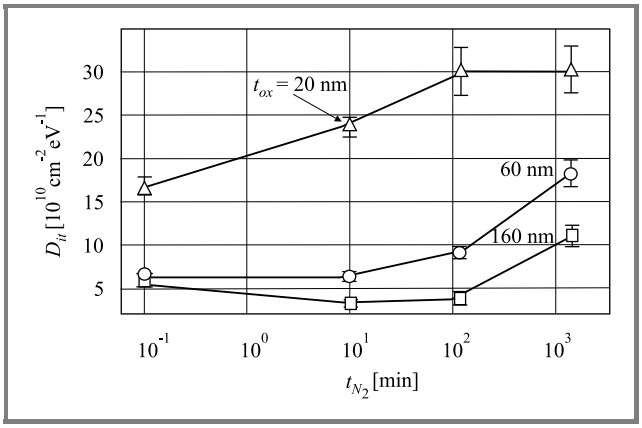


Fig. 6. Mid-gap interface trap density D_{it} versus nitrogen-annealing time t_{N_2} for three silicon dioxide thicknesses of 20, 60, and 160 nm.

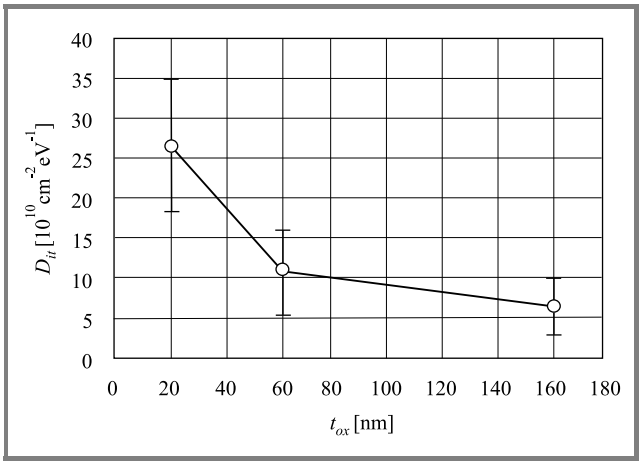


Fig. 7. Mid-gap interface trap density D_{it} versus oxide thickness.

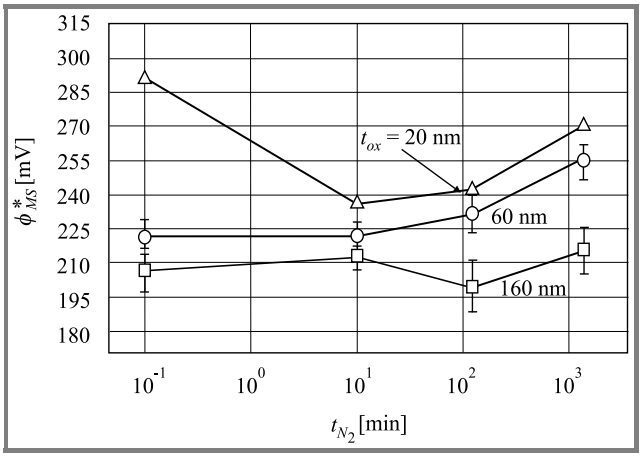


Fig. 8. The reduced effective contact potential difference ϕ^*_{MS} on nitrogen-annealing time t_{N_2} for three silicon dioxide thicknesses of 20, 60, and 160 nm.

Photoelectric measurements allowed the effective contact potential difference (the ϕ_{MS} factor or EPCD) to be determined. The reduced effective contact potential differ-

ence ϕ^*_{MS} is shown as a function of nitrogen-annealing time and oxide thickness in Figs. 8 and 9 for the three oxide thicknesses studied.

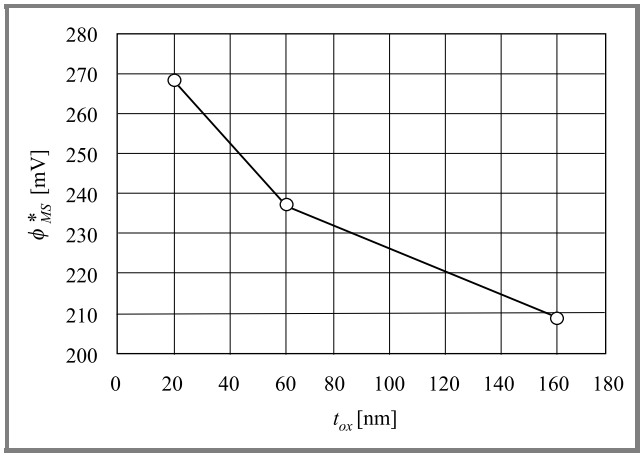


Fig. 9. The reduced effective contact potential difference ϕ^*_{MS} versus oxide thickness.

The dependence of the effective oxide charge density (N_{eff}) on the nitrogen-annealing time (t_{N_2}) is illustrated in Fig. 10.

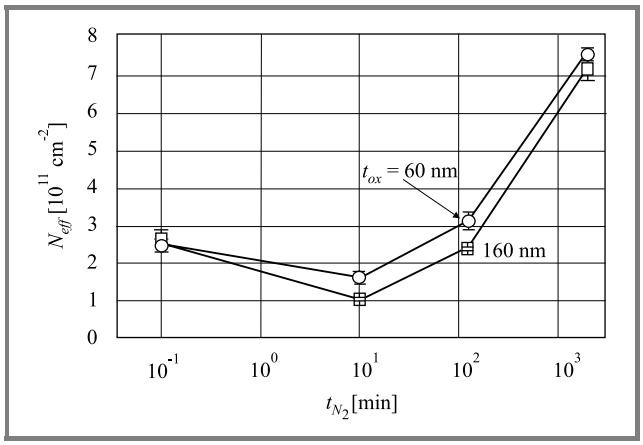


Fig. 10. Effective charge density N_{eff} in the oxide versus nitrogen-annealing time t_{N_2} for three silicon dioxide thicknesses of 60 and 160 nm.

The shape of the $\sigma_{ox}(t_{N_2})$, $n(t_{N_2})$, $D_{it}(t_{N_2})$, $\phi^*_{MS}(t_{N_2})$, and $N_{eff}(t_{N_2})$ plots results from the combination of two competing processes. The first of those is thermal-relaxation that occurs in the initial phases of annealing. The second is nitrogen incorporation that takes place in the SiO_2 layer in the vicinity of the Si-SiO_2 interface. The latter occurs at longer times of annealing in nitrogen. The trends in the above described plots are such that both the SiO_2 refractive index and the electrical parameters decrease in the case of short annealing times then start increasing when annealing times become longer.

4. Conclusions

We have investigated the effect of high-temperature stress annealing in the ambient of nitrogen on the optical and electrical properties of MOS devices. Based on the changes occurring in the refractive index of the SiO_2 layer and electrical parameters of MOS structures, we claim that two processes are present during annealing: thermal relaxation and nitrogen incorporation. The first process is responsible for the initial trends in the silicon dioxide properties. The second process is, however, responsible for the trends in the SiO_2 properties at longer annealing times.

References

[1] E. H. Nicollian and J. R. Brews, *MOS Physics and Technology*. New York: Wiley, 1982.

[2] H. M. Przewłocki and H. Z. Massoud, "Effects of stress annealing in nitrogen on the effective contact-potential difference, charges, and traps at the Si/SiO_2 interface of metal-oxide-semiconductor devices", *J. Appl. Phys*., vol. 92, pp. 2198–2201, 2002.

[3] H. Z. Massoud and H. M. Przewłocki, "Effects of stress annealing in nitrogen on the index of refraction of silicon dioxide layers in metal-oxide-semiconductor devices", *J. Appl. Phys*., vol. 92, pp. 2202–2206, 2002.

[4] W. Primak and D. Post, "Photoelastic constants of vitreous silica and its elastic coefficient of refractive index", *J. Appl. Phys*., vol. 30, pp. 779–789, 1959.

[5] G. Ghibaudo, "Modeling of silicon oxidation based on stress relaxation", *Phil. Mag*., vol. B 55, pp. 147–158, 1987.

[6] A. Szekeres, "Stress in the SiO_2/Si structures formed by thermal oxidation", in *Fundamental Aspects of Ultrathin Dielectrics on Si-Based Devices*. Dordrecht: Kluwer, 1998, pp. 65–78.

[7] B. Leroy, "Stresses and silicon interstitials during the oxidation of a silicon substrate", *Phil. Mag*., vol. B 55, pp. 159–199, 1987.

[8] C. M. Herzinger, B. Yohs, W. A. McGahan, and J. A. Woollam, "Ellipsometric determination of optical constants for silicon and thermally grown silicon dioxide via a multi-sample, multi-wavelength, multi-angle investigation ", *J. Appl. Phys*., vol. 83, pp. 3323–3336, 1998.

[9] A. Brenner and S. Senderoff, "Calculation of stress in electrodeposits from the curvature of plated strip", *J. Res. Nat. Bur. Stand*., vol. RP 1954, no. 42, pp. 105–123, 1949.

[10] B. Leroy, "Passivation-induced phenomena in silicon substrate" in *Instabilities in Silicon Devices*. Amsterdam: North Holland, 1986, pp. 155–210.

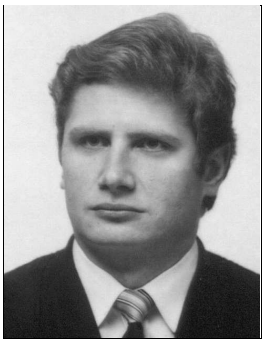
[11] P. H. Townsend, D. M. Barnett, and T. A. Brenner, "Elastic relationship in layered composite media with approximation for the case of thin films on a thick substrate", *J. Appl. Phys*., vol. 62, pp. 4438–4444, 1987.

[12] C. A. Klein, "How accurate are Stoney's equation and recent modifications", *J. Appl. Phys*., vol. 88, pp. 5487–5489, 2000.

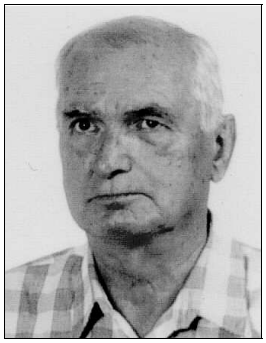
[13] W. A. Brantley, "Calculated elastic constants for stress problems associated with semiconductor devices", *J. Appl. Phys.*, vol. 44, pp. 534–535, 1973.

[14] Tencor FLX-2320 Thin Film Stress Measurement, User Manual, Rev. B, 1995.

[15] H. M. Przewłocki, "Theory and applications of internal emission in the MOS system at low electric fields", *Solid-State Electron*., vol. 45, pp. 1241–1250, 2001.



Witold Rzodkiewicz was born in Warsaw in 1971. He received the M.Sc. degree in material science and engineering from the Warsaw University of Technology in 1995. His Masters project in "Studies of the structural perfection of GaInAsSb quaternary layers grown by LPE on GaSb substrates" was carried out at the Institute of Electron Technology (IET) in Warsaw. From 1995 to 1998, he was with the Department of Materials and Semiconductor Structures Research of the Institute of Electron Technology in Warsaw as a Research Assistant, working in the area of process-induced defects in semiconductor structures. From 1998 till now, he is with the Department of MOS System Studies in IET as a Research Assistant.
e-mail: rzodki@ite.waw.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland



Zbigniew Sawicki was born in 1932. From 1951 to 1956 he studied astrophysics at the Warsaw University. Since 1958 he is engaged in the research of semiconductor technology. His special interest was oxidation and diffusion processes in silicon.

Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Andrzej Kudła, Henryk M. Przewłocki – for biography, see this issue, p. 39.

Paper

Comparison of the barrier height measurements by the Powell method with the ϕ_{MS} measurement results

Krzysztof Piskorski, Andrzej Kudła, Witold Rzodkiewicz, and Henryk M. Przewłocki

Abstract— In this work, we have compared the barrier height measurements carried out using the Powell method with the photoelectric effective contact potential difference (ϕ_{MS}) measurement results. The photoelectric measurements were performed on the samples that were previously applied in the investigation of the influence of stress on the duration of annealing in nitrogen. This paper shows that the results of barrier height measurement using the Powell method differ significantly from the ϕ_{MS} measurement results.

Keywords— barrier height, effective contact potential difference, MOS system.

1. Introduction

Significant differences between the values of barrier heights in the Al-SiO_2-Si structure found in the literature may be at least partly explained by the inaccuracies of the measurement methods and the lack of a sufficiently precise method of the verification of the obtained results.

The effective contact potential difference (ϕ_{MS}) is measured with the accuracy of $\pm$10 mV [4]. Having at our disposal such an extremely precise method of ϕ_{MS} measurement [4, 5] we have decided to use it to verify the results of the measurements of internal photoemission barrier heights in the MOS structure based on the Powell method [1, 2]. The accuracy of the barrier height (E_{BG} – metal-dielectric, E_{BS} – semiconductor-dielectric) measurements using the Powell method was estimated in [6] at $\pm$50 mV. Accordingly, we have compared the measurement results of E_{BG}, E_{BS} and ϕ_{MS}.

2. Theory

The internal photoemission phenomena may be observed in a MOS structure with a semitransparent gate, illuminated by UV radiation. The UV radiation absorbed in the electrodes (the gate or the substrate) causes the excitation of some electrons. If these electrons acquire sufficient energy to surmount the potential barrier at the electrode – insulator interface, they may pass into the insulator giving rise to a photocurrent in the external circuit. The measurement system for photoelectric measurements is shown in Fig. 1. The band diagram of the MOS system is shown in Fig. 2.

Balancing the potentials on both sides of the dielectric layer yields [6]:

$$\phi_M - U_G = \chi_{\mathrm{Si}} - \phi_I - \phi_S + \frac{E_{g,\mathrm{Si}}}{2q} + \phi_F \,, \qquad (1)$$

where: ϕ_M – the barrier height at the gate/dielectric interface, U_G – gate potential, χ_{Si} – the electron affinity of the silicon substrate at the interface, ϕ_I, ϕ_S – the potential drop in the dielectric and at the semiconductor surface, $E_{g,\mathrm{Si}}/2q$ – the voltage equivalent of half energy bandgap in the semiconductor, q – the electron charge, ϕ_F – the Fermi level.

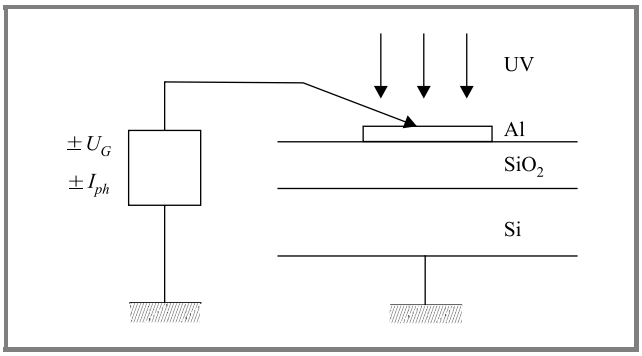


Fig. 1. The measurement system: a MOS structure with a semitransparent gate is illuminated by UV light. Photocurrent is measured in the external circuit.

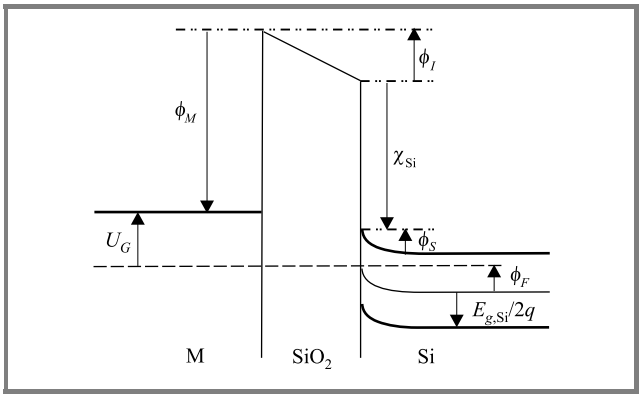


Fig. 2. Band diagram of a MOS system, at an arbitrary gate potential U_G.

The definition of the effective contact potential difference (ϕ_{MS}) offers the possibility of a comparison between the difference of internal photoemission barrier heights from both sides of the dielectric and the value of ϕ_{MS}.

The effective contact potential difference (ϕ_{MS}) is defined as:

$$\phi_{MS} = \phi_M - \left(\chi_{\mathrm{Si}} + \frac{E_{g,\mathrm{Si}}}{2q} + \phi_F\right). \tag{2}$$

The reduced effective contact potential difference (ϕ^*_{MS}) is defined as:

$$\phi^*_{MS} = \phi_M - \chi_{\mathrm{Si}} \tag{3}$$

or

$$\phi^*_{MS} = \phi_{MS} + \frac{E_{g,\mathrm{Si}}}{2q} + \phi_F\,. \tag{4}$$

The ϕ^*_{MS} value depends on the barrier height on both sides of the dielectric and does not depend on the doping concentration in the substrate.

To make our paper easier to read we propose the following symbols to denote the values of the reduced effective contact potential difference obtained using different methods:

- $\phi^*_{MS}(1)$ – the reduced effective contact potential difference determined on the basis of direct photoelectric measurements (4),
- $\phi^*_{MS}(2)$ – the reduced effective contact potential difference calculated (3) using the E_{BG} and E_{BS} values measured by the Powell method.

Subtracting (3) from (4) we have:

$$\phi^*_{MS}(1) - \phi^*_{MS}(2) = \phi_{MS} - \phi_M + \chi_{\mathrm{Si}} + \frac{E_{g,\mathrm{Si}}}{2q} + \phi_F = R. \tag{5}$$

The difference R is equal to 0 for an ideal measurement. Otherwise R stands for the error of the barrier-height measurement.

The value of R may be used to evaluate the accuracy of barrier height measurements. The ϕ_{MS} factor is determined from photocurrent measurement, while ϕ_F is determined from capacitance – voltage (C–V) measurements with the total accuracy better than 10 mV. In this case the value of R higher than 10 mV means that at least one of the considered barrier heights was measured inaccurately.

3. Experimental characterization

N-type (100) silicon wafers were used in this work. The wafers were doped with phosphorus to obtain the resistivity of 3–5 Ωcm. After the initial hydrogen-peroxide-based cleaning sequence, the wafers were thermally oxidized at 1000°C in oxygen to grow silicon-dioxide layers with the thickness of approximately 20, 60, and 160 nm. The wafers were subsequently annealed in nitrogen at 1050°C for periods of 0, 10, 120, and 1440 min. The gate metallization was carried out in a thermal evaporator so that the obtained Al thickness was 35 nm. Thin gate Al is necessary in MOS photoelectric measurements. The metallization was then patterned with optical lithography and the backside oxide was etched prior to the backside metalization. Postmetallization annealing was carried out in forming gas for 20 min at 450°C.

4. Experimental results and discussion

The wafers, used for the photoelectric measurements were previously applied in the investigation of the influence of stress on the time of annealing in nitrogen [3]. Accordingly, the photoelectric parameters (ϕ_{MS}, E_{BG}, E_{BS}) will be shown in this work as a function of the duration of annealing in nitrogen.

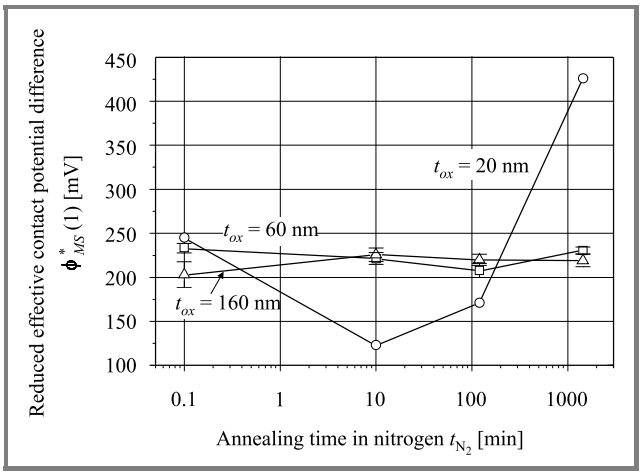


Fig. 3. The determined reduced effective contact potential difference $\phi^*_{MS}(1)$ versus the time t_{N_2} of annealing in nitrogen for oxide thickness of 20, 60, and 160 nm.

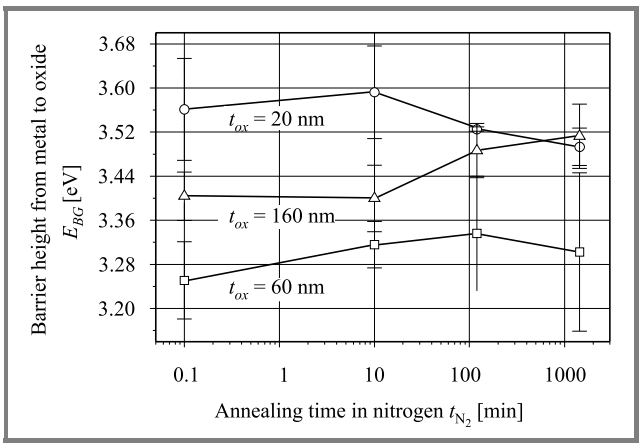


Fig. 4. Results of the E_{BG} measurements versus the time t_{N_2} of annealing in nitrogen for oxide thickness of 20, 60, and 160 nm.

The dependence of the determined reduced effective contact potential difference $\phi^*_{MS}(1)$ on the annealing time is shown in Fig. 3.

The dependence of the measured E_{BG} barrier on the annealing time for oxide thickness of 20, 60, and 160 nm is shown in Fig. 4.

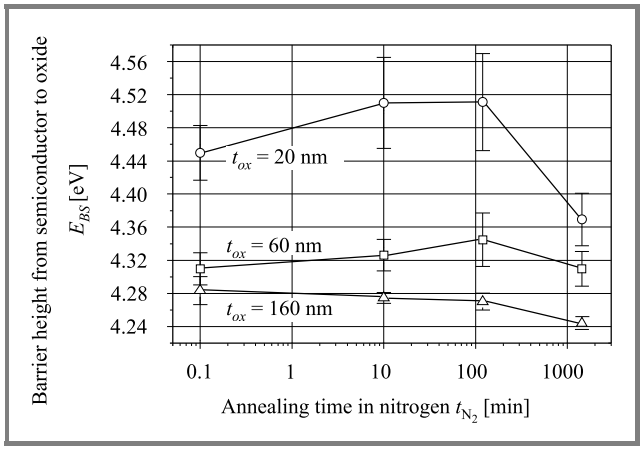


Fig. 5. Results of the measured barrier height between semiconductor and oxide E_{BS} versus the time t_{N_2} of annealing in nitrogen for oxide thickness of 20, 60, and 160 nm.

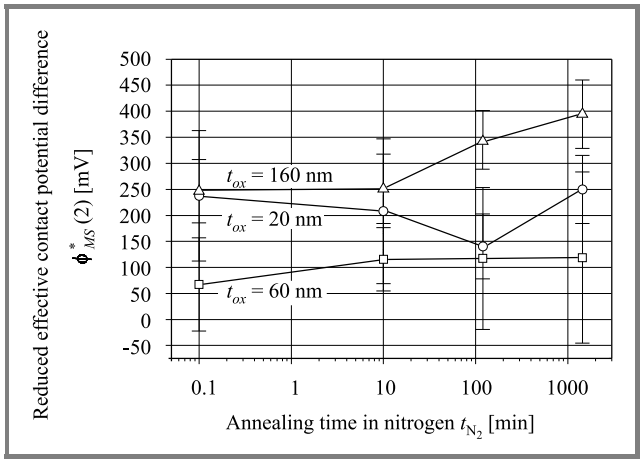


Fig. 6. The calculated effective contact potential difference $\phi^*_{MS}(2)$ versus the time t_{N_2} of annealing in nitrogen for oxide thickness of 20, 60, and 160 nm.

Table 1
Values of measurement error R

t_{N_2} [min]	R [mV]		
	t_{ox} [nm]		
	20	60	160
0	7.64	165.74	–43.26
10	–85.08	105.65	–25.3
120	30.78	90.33	–122.32
1440	176.08	111.84	–175.37

Figure 5 shows the dependence of the E_{BS} measurement results on the annealing time for oxide thickness of 20, 60, and 160 nm.

The dependence of the reduced effective contact potential difference $\phi^*_{MS}(2)$ (calculated on the basis of the measurements of the barrier heights E_{BG} and E_{BS}) on the annealing time is shown in Fig. 6.

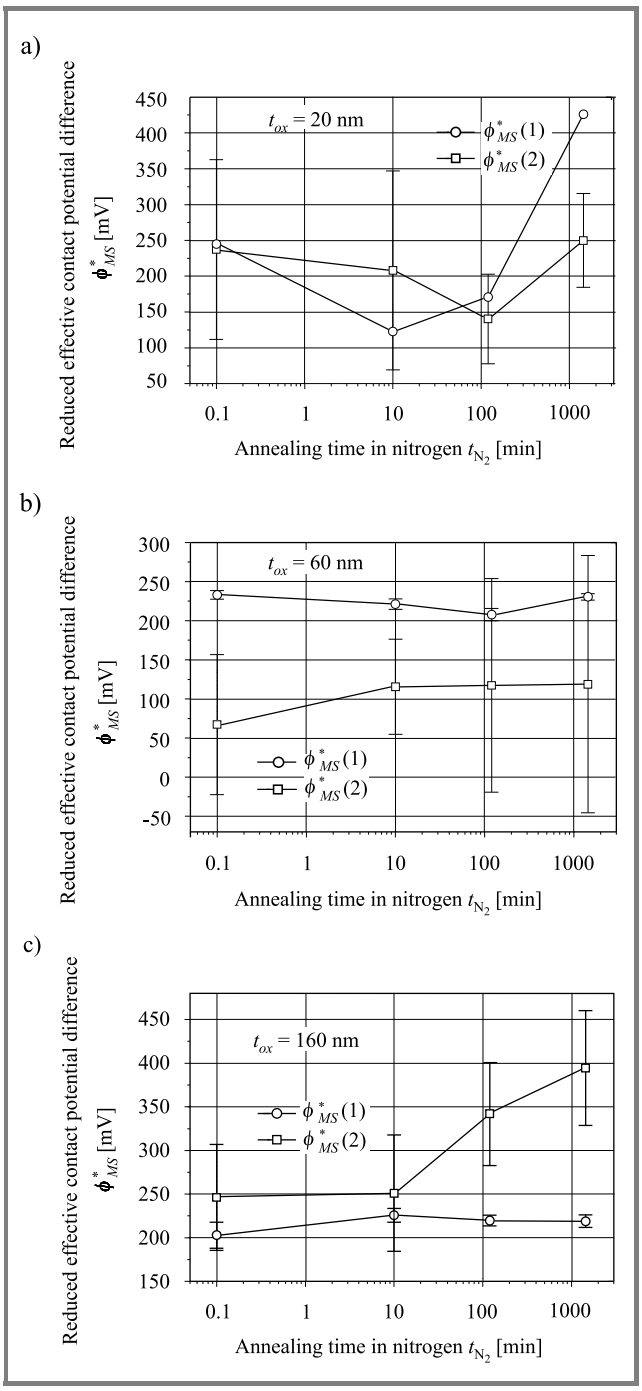


Fig. 7. Measured $\phi^*_{MS}(1)$ and $\phi^*_{MS}(2)$ calculated reduced effective contact potential difference versus the time t_{N_2} of annealing in nitrogen for oxide thickness of: 20 nm (a); 60 nm (b); 160 nm (c).

Subsequently, we have compared the values of $\phi^*_{MS}(1)$ and $\phi^*_{MS}(2)$ shown as a function of the annealing time in Fig. 7 for different oxide thicknesses.
The values of the *R* factor in [mV] determined from (5) are given in Table 1. The table indicates that the measurement error *R* can be both positive and negative.

5. Conclusions

In this work, we have compared the reduced effective contact potential difference $\phi^*_{MS}(1)$ (determined on the basis of the ϕ_{MS} measurement) with the reduced effective contact potential difference $\phi^*_{MS}(2)$ (calculated on the basis of the barrier height measurements using the Powell method).
This research shows (Fig. 7 and Table 1) that the barrier heights measured using the Powell method are significantly different from the results of ϕ_{MS} measurements. We attribute these differences to the poor accuracy of the Powell method.
It is believed that the main causes of this inaccuracy are:

- errors made in the extrapolation of *I–V* characteristics;
- improper values of the p-factor used for calculations of the barrier heights.

The positive value of the measurement error *R* may be explained by too low a value of the barrier height measured at the gate – SiO_2 interface or too high a value of the barrier height measured at the SiO_2 – semiconductor interface. The negative value of the error may be explained by too low a value of the barrier height measured at the SiO_2 – semiconductor interface or too high a value of the barrier height measured at the metal – SiO_2 interface. The non-zero value of *R* is in our view primarily caused by inappropriate values of the p-factor used in the barrier height determination.
In our further research we will focus on the main factors affecting the accuracy of the barrier height determination methods. In particular, the physical nature and the ways to choose the appropriate values of the p-factor will be studied.

References

[1] R. J. Powell, "Photoinjection into SiO_2: use of optical interference to determine electron and hole contributions", *J. Appl. Phys.*, vol. 40, p. 5093, 1969.

[2] R. J. Powell, "Interface barrier energy determination from voltage dependence of photoinjected currents", *J. Appl. Phys.*, vol. 41, p. 2424, 1970.

[3] H. M. Przewłocki and H. Z. Massoud, "Effects of stress annealing in nitrogen on the effective contact-potential difference, charges, and traps at the Si/SiO_2 interface of metal-oxide-semiconductor devices", *J. Appl. Phys.*, vol. 92, pp. 2198–2201, 2002.

[4] H. M. Przewłocki, "The importance, the nature and the measurement methods of the ϕ_{MS} factor in MOS structures", *Electron Technol.*, vol. 27, no. 1, pp. 7–42, 1994.

[5] H. M. Przewłocki, "Theory and applications of internal photoemission in the MOS system at low electric field", *Solid-State Electron.*, vol. 45, p. 1241, 2000.

[6] A. Kudła, "Fotoelektryczne metody wyznaczania wysokości barier potencjału w strukturze MOS", *Prace Instytutu Technologii Elektronowej*, no. 5–7, pp. 7–115, 1998 (in Polish).

Krzysztof Piskorski was born in Zgierz, Poland, in 1976. He received the M.Sc. degree from the Technical University of Łódź, Poland, in 2002. His Masters project in "Dry etching of A_{III}-B_V nitrides" was carried out at the Ecole Centrale de Lyon, France, in 2001. He is currently working as a Research Assistant at the Institute of Electron Technology (IET) in Warsaw, Poland, with the Department of MOS System Studies. His research interests include photoelectric measurements for MOS structures.
e-mail: kpisk@ite.waw.pl
Institute of Electron Technology
Lotników av. 32/46
02-668 Warsaw, Poland

Andrzej Kudła, Henryk M. Przewłocki – for biography, see this issue, p. 39.

Witold Rzodkiewicz – for biography, see this issue, p. 119.

Paper

Semi-automatic test system for characterization of ASIC/MPWS

Jerzy Zając, Janusz Wójcik, Andrzej Kociubiński, and Daniel Tomaszewski

Abstract—A measurement system for integrated circuit testing has been developed. It consists of a semi-automatic probe station and a set of measurement equipment controlled by commercially available measurement software. The probe station is controlled by dedicated software. Both the measurement and station-control software communicate using the DDE protocol. The measurement system is flexible. It is particularly suitable for semi-automatic testing of multi-project wafers. Output data generated by the system is used for the characterization of the CMOS technologies.

Keywords— automatic testing, diagnostics of technology, multi-project wafers.

1. Introduction

Semi-automatic test system developed by the Industrial Institute of Electronics (IIE) and Institute of Electron Technology (IET) is used to test wafers containing different types of structures and to characterize diagnostics structures irregularly distributed across the wafer. Testing of standard wafers with uniformly distributed structures is also possible.
The system contains probing and measurement subsystems and is controlled by a PC working under the MS Windows operating system. Both subsystems consist of control application (software) and related hardware.
In this paper the system is described and an application example is shown. The intentions of the authors concerning future work are presented, too.

2. Motivation

Numerous application-specific integrated circuits (ASICs) are designed and manufactured in the Institute of Electron Technology. Dedicated test structures, which accompany the ASICs, are used for characterization of technologies and verification of ICs designs. Testing of both standard wafers with uniformly distributed structures and multi-project wafers (the so-called MPWs) containing different types of structures is of great importance.
In order to meet this requirement a semi-automatic IC testing system has been developed in the Industrial Institute of Electronics and implemented in the IET. The system consists of several hardware and software components, designed and manufactured in the IET. The components are now available commercially.
The testing system and all its hardware and software components are described in detail below. Communication between the software packages is presented, too. The operation of the system is described and illustrated by a practical example. Finally, possible future improvements of the system are discussed.

3. System structure

The testing system consists of the following components (Fig. 1):

- probe station (redesigned version of a standard device) controlled by the ELECT program;
- Keithley SMU236/7 units (connected with a device-under-test (DUT) via Keithley switching matrix 707A), controlled by the METRICS program, available commercially.

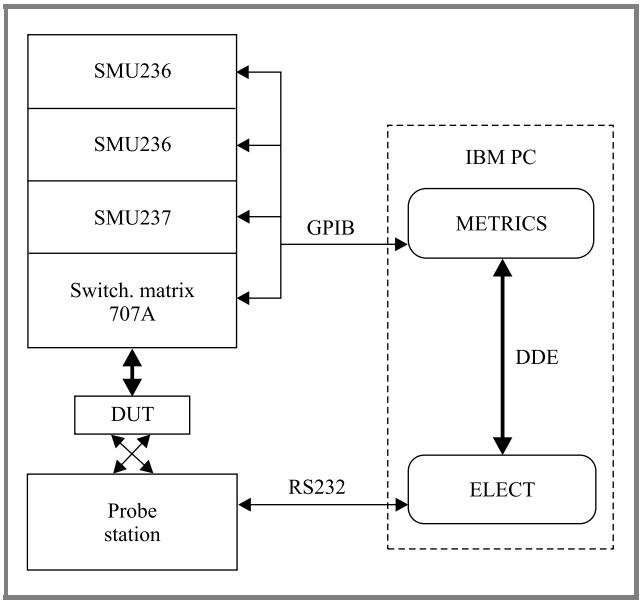


Fig. 1. Block diagram of the testing system.

3.1. Probe station and ELECT program

The probe station is equipped with a set of probe cards dedicated to testing of specific DUTs. At present individual micromanipulators are not available.
The ELECT program is fully responsible for the operation of the probe station. The following operations are available:

- navigation of the chuck; the resolution of chuck positioning is 5 μm; this enables measurement of devices with small pads;

– adjustment of the chuck position and orientation; these options are necessary for proper movement of the chuck during automatic testing;
– testing of the chuck positioning repeatability;
– start of measurement;
– acquisition of measurement results obtained by METRICS.

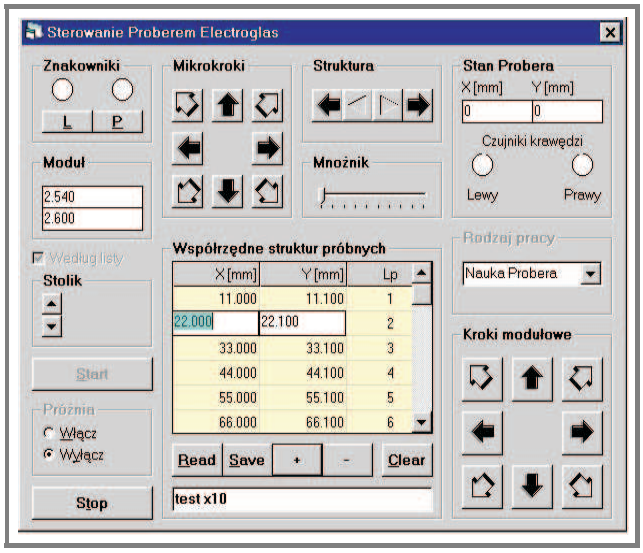


Fig. 2. The main form of ELECT in the session preparation mode with a grid containing the list of co-ordinates.

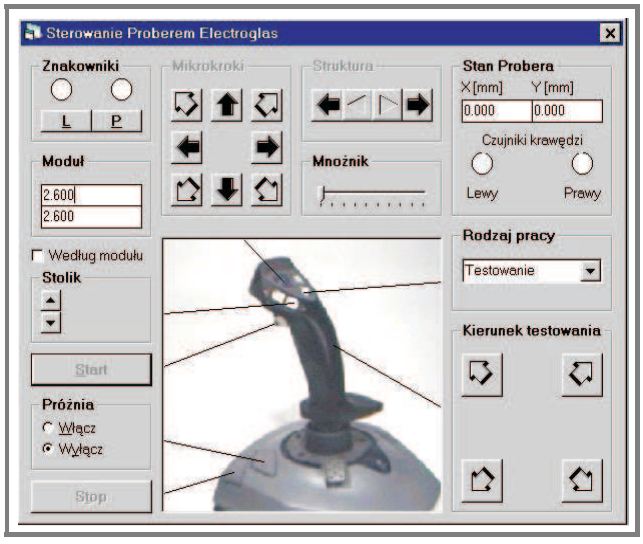


Fig. 3. The main form of ELECT in the session runtime mode.

The ELECT application co-operates with the probing station via a RS232 interface. The application can operate in two ways: session preparation (control form shown in Fig. 2) and session runtime (Fig. 3).

3.2. Measurement equipment and METRICS program

The measurement subsystem is the standard System 93 I-V delivered by Keithley. At the moment the subsystem contains three DC source-measure units (SMUs), a switching matrix, a synchronisation unit (not shown in Fig. 1) and METRICS, that is the control application. Other equipment may be added in the future. The application and measurement units are interfaced by GPIB.
The METRICS program is used to :

– arrange internal data-base consisting of the so-called projects; the project stores measurement configurations as well as measurement data;
– arrange the measurement projects; the project is a collection of the so-called setups;
– arrange the measurement setups; the setup defines a single measurement configuration and stores data obtained recently using this configuration;
– specify measurement configuration using a graphical editor (definition of connections of the SMUs with DUT nodes, configuration of the switching matrix, if used; definition of the SMUs input data);
– define additional calculations performed upon measured data; simple extraction of parameters of the measured devices may be done; post-processing formulae are stored in the measurement setups;
– define internal tests upon measured data;
– program measurement devices, switching matrix and trigger controller;
– arrange a sequence of measurements (e.g., selected setups may be executed once or repeatedly);
– trigger the measurements;
– receive measured data from the measurement equipment and store it in the internal memory of the corresponding setup;
– visualise measured data;
– export measured data; data can be exported as text files and as MS Excel worksheets; these files may be used for further processing.

Measurement projects may be created using the METRICS independently from the other parts of the system. Thus METRICS may be used either as a stand-alone application or may be invoked by ELECT. The communication between these two programs is described in the next chapter.

3.3. Communication between ELECT and METRICS

Both software packages of the system co-operate using the dynamic data exchange (DDE) software interface [1–3]. METRICS works as a DDE server and ELECT as a DDE client. ELECT sends commands to METRICS. Then it receives measurement results obtained after the execution of all setups in the project selected in the appropriate text

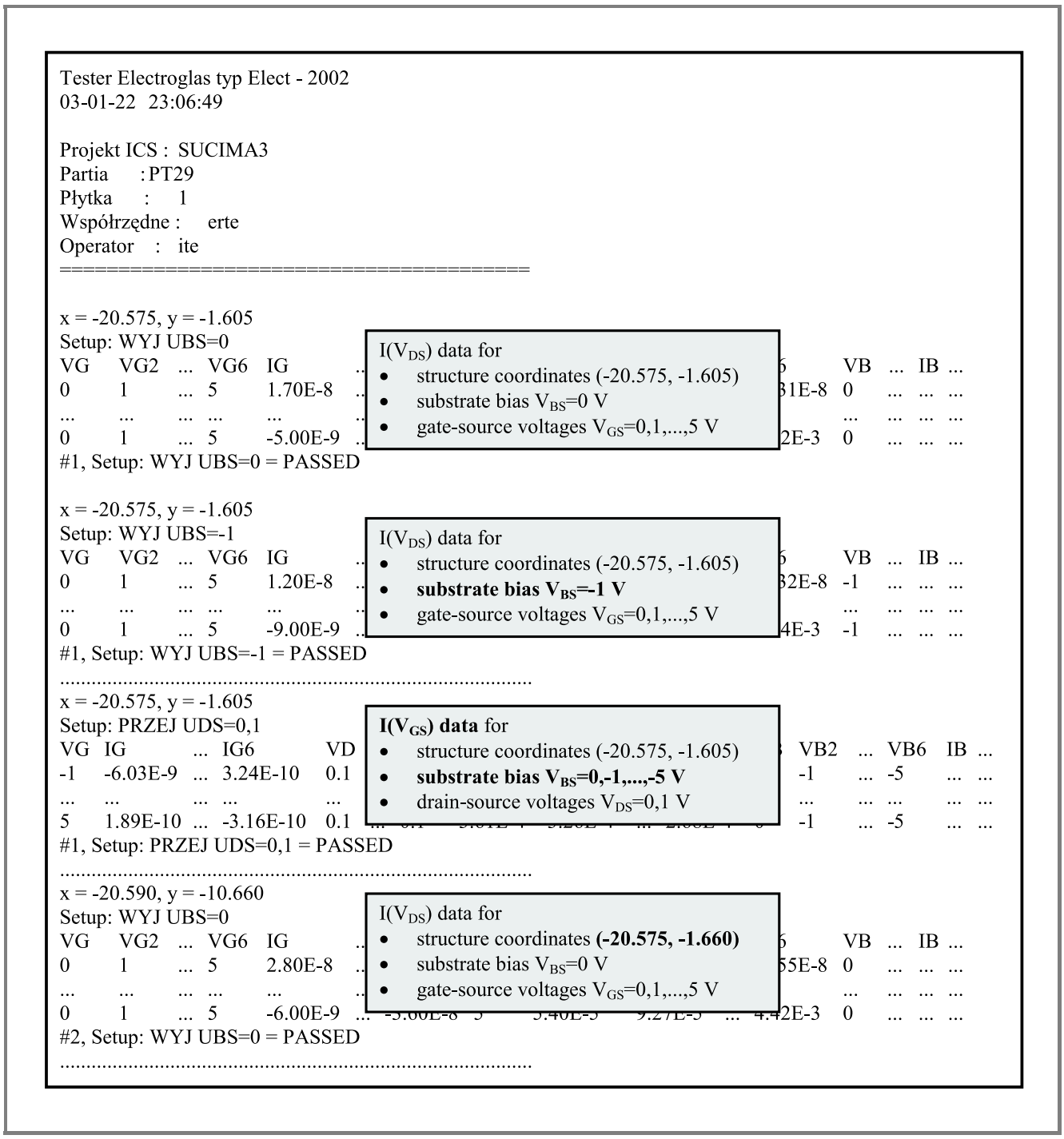


Fig. 4. Contents of the text file written by ELECT program; each section corresponds to a single setup of METRICS project SUCIMA3.

box of the main ELECT form. After the measurements ELECT stores all results in a text file. At the top of the file there is a header containing:

- date/time of measurements;
- user, lot and wafer identifiers;
- project file name;
- name of the file, which contains the co-ordinates list.

These data are saved as a file of the comma separated values (CSV) format. The data obtained for the given setup and for the given structure are also marked with a setup identifier, structure (die) co-ordinates and measurement results. The detailed structure of the result file (e.g., names of setups, number of results, etc.) is adjusted automatically, according to the project executed by METRICS (Fig. 4). This file can be easily imported by other standard applications, e.g., MS Excel, in order to extract physical parameters of structures.

4. System operation

The preparation of the measurement session can be done separately for ELECT and METRICS. The collections of measurement set-ups are prepared in the METRICS environment in a standard way. Next they are stored in a native format of the METRICS application.
Probing station requires a list of co-ordinates of the structures to be tested. This list can be created using the ELECT during the learning session. Then the user moves the chuck of the probing station to appropriate position, using a joystick or a mouse and stores this position in the list shown in Fig. 2. The complete list can be saved next in a text file. The list of co-ordinates can be also prepared using an external text editor or directly in the grid component in the main form of the ELECT program. The list of co-ordinates is necessary if the test structures irregularly distributed on the wafer are to be measured. In other case, module dimensions of the structure must be used.
The probing station is tested while the ELECT application is loaded. Afterwards, the user can start operations. First a wafer should be placed on the chuck and carefully aligned. Next the chuck moves to the central position. The user may then adjust precisely the position of the central structure. After confirming adjustment, (0,0) co-ordinates are assigned to the central structure and the user can choose another structure as a starting point. The testing may start under the condition that the initial direction is specified. The co-ordinates of the tested structure are displayed in a control window. The progress of wafer testing may be also shown if the co-ordinate list is used. In the case of a wafer with uniformly distributed structures one or two edge sensor must be used.
ELECT and METRICS work concurrently during wafer testing. The necessary data are exchanged between them through a DDE channel as follow:

1. ELECT requests test information from METRICS.
2. METRICS sends test names to ELECT.
3. ELECT sends a DDE message *Start measure*.
4. METRICS starts execution of the project.
5. When the execution is finished, METRICS sends message *Pass* or *Fail* accordingly to the test result.
6. If ELECT receives the message *Ready*, the probing station moves the chuck to the next position (from the list or using structure module dimensions) and requests test results from METRICS.
7. METRICS sends a DDE message containing test results.
8. When result transfer and chuck movement are completed, ELECT appends the test results of the structure to the result file and sends a DDE message *Start measure*.

Execution of Step 4 through Step 8 continues until the co-ordinates list is exhausted or all structures on the wafer are measured.
Prior to starting measurements, the user must write the name of the METRICS project in the ELECT window.
If MERTICS is not running during the execution of Step 1 ELECT invokes it automatically and the project file name is used to load the project.
Measurements of the wafer can be terminated or stopped temporarily after finishing tests of a given structure.
The header of the output file contains:

- measurement date;
- user; lot and wafer identifiers;
- METRICS project (set-up collection) specification and co-ordinate file name.

The results of structure testing are organized in blocks containing:

- structure co-ordinates;
- sub-blocks with set-up name and variables, values, and PASS/FAILED classification related to the set-up.

The probing station can use up to two inkers (2 inkers are used for delayed inking). If the station uses a structure module at least one edge sensor must be connected. In this case the station can be also controlled by a test system through the standard hardware interface, but structure co-ordinates are lost then.

4.1. Example of data

The ELECT/METRICS measurement system has been used for characterization of CMOS devices fabricated on thick-film SOI wafers. The process sequence to fabricate such devices is currently being developed by IET to be applied in the fabrication of monolithic pixel detectors. These unique devices are to detect ionizing radiation in medical applications [3]. The TSSOI test structure is the main tool for the characterization of the new technology [4]. It contains not only numerous devices like transistors, diodes, capacitors, resistive paths but also digital cells. I-V characteristics are measured using the ELECT/METRICS system. The results are used subsequently to extract process and device parameters. An example of the measured output chcracteristics of a n-channel MOSFET are shown in Fig. 5 together with SPICE simulations. The parameters of the MOSFETs models required by the simulations were obtained using the MOSTXX extraction tool developed recently in the IET [5].
It should be mentioned that TSSOI structures are not the only test structures fabricated in IET. Thus the flexibility of the ELECT program (I-V data were measured on structures selected according to a list of co-ordinates) appears to be very important.

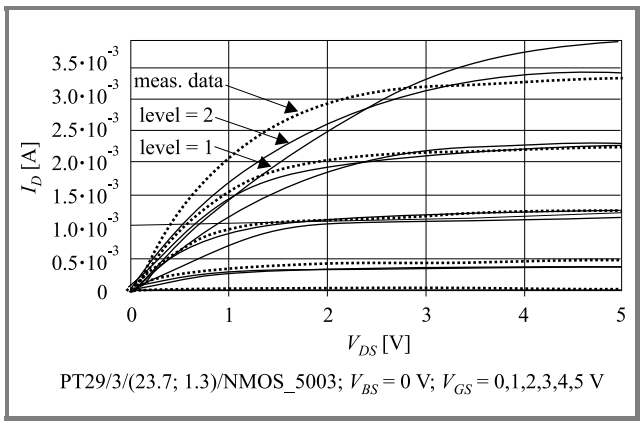


Fig. 5. Measured and simulated output characteristics of an n-channel MOSFET fabricated on a thick-film SOI wafer (W/L = 50 μm/3 μm).

In the case of measurements of multi-project wafers each test structure requires a separate METRICS project file. Also separate probe cards and co-ordinates lists files may be necessary.

5. Perspectives, future work

The presented system will be used for measurements required by detailed characterization of the fabrication process, as well as for measurements necessary for faster, parametric testing. In this case wafers mapping seems to be very relevant. Thus future efforts to improve the automatic testing system will be concentrated on the implementation of mapping. An additional application will be developed for this purpose.

METRICS application is a general-purpose measurement tool, it is, however, subject to certain limitations. Thus it will be necessary to develop dedicated measurement programs using high-level languages, such as C++ or Delphi/Pascal. In order to use them for the purposes of automatic testing it will be necessary to combine them with ELECT via the DDE protocol. Thus development of ELECT code seems to be the next important task in the presented system.

6. Conclusion

A measurement system for automatic testing of integrated circuits has been developed. It consists of a semi-automatic probe station and measurement equipment controlled by commercial measurement software. The probe station is controlled by dedicated software developed in IET. Both programs communicate using a DDE protocol but may also operate independently. The measurement system is flexible. It enables testing of both standard wafers and multi-project wafers to be tested using edge sensors and lists of co-ordinates, respectively.

Output data generated by the system is accepted by the extraction software used in IET for the characterization of the CMOS fabrication process.

The system is under development. Future work will concentrate on wafer mapping and establishment of links to other dedicated measurement programs in order to use them in automatic testing.

References

[1] H. Rodent, "Supportingin the clipboard, ODE and OLE in applications", MSDN Library Visual Studio 6.0, 1992.

[2] H. Rodent, "Quick and easy DDE server", MSDN Library Visual Studio 6.0, 1992.

[3] "METRICS documentation", http://www.metricstech.com

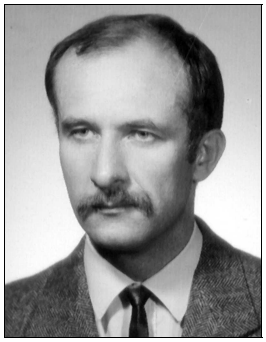
[4] J. Marczewski *et al*., "Monolithic silicon pixel detectors in SOI technology", in *Linear Coll. Worksh*., Prague, Czech Republic, 2002, http://www-hep2.fzu.cz/ecfadesy/Talks/Vertex_Detector/Marczewski_Jacek_Prague_ECFA_DESY.pps

[5] M. Barański *et al*., "TSSOI as an efficient tool for diagnostics of SOI technology in the IET", in *6th Symp. Diagn.&Yield*, Warsaw, Poland, 2003.

[6] D. Tomaszewski *et al*., "A versatile tool for MOSFETs parameters extraction", in *6th Symp. Diagn.&Yield*, Warsaw, Poland, 2003.

Jerzy Zając received the M.Sc. and Ph.D. degree in electrical engineering from Warsaw University of Technology in 1976 and Gdańsk University of Technology in 2000, respectively. His research interests include development of complex measurement systems, machine vision, and industry automation. His recent activities are related to integration of different measurement systems with databases.

e-mail: jzajac@pie.edu.pl
Industrial Institute of Electronics
Długa st 44/50
00-241 Warsaw, Poland



Janusz Wójcik received the M.Sc. degree in electrical engineering from Warsaw University of Technology in 1975. His research interests include development of complex measurement systems for semiconductor industry. His recent activities are related to testers and probing stations for HF measurement of semiconductor structures.

e-mail: jwojcik@pie.edu.pl
Industrial Institute of Electronics
Długa st 44/50
00-241 Warsaw, Poland

Andrzej Kociubiński – for biography, see this issue, p. 92.

Daniel Tomaszewski – for biography, see this issue, p. 93.

Paper

A versatile tool for extraction of MOSFETs parameters

Daniel Tomaszewski, Andrzej Kociubiński, Jacek Marczewski, Krzysztof Kucharski, Krzysztof Domański, and Piotr Grabiec

Abstract— Extraction of MOSFET parameters is a very important task for the purposes of MOS integrated circuits characterization and design. A versatile tool for the MOSFET parameter extraction has been developed in the Institute of Electron Technology (IET). It is used to monitor the technologies applied for fabrication of several groups of devices, e.g., CMOS ASICs, SOI pixel detectors. At present two SPICE MOSFET models (LEVEL = 1, 2) have been implemented in the extraction tool. The LEVEL = 3 model is currently being implemented. The tool combines different methods of parameter extraction based on local as well as global fitting of models to experimental data.

Keywords— MOSFETs parameters, SPICE, least squares method.

1. Introduction

Simulations of MOS ICs (integrated circuits) require sufficiently accurate MOSFET models. Moreover, the parameters of these models should be determined in a reliable way. Parameter extraction is an important task for characterization groups in MOS ICs fabrication departments.

The complexity of parameter extraction is usually closely related to the complexity of the model itself. The complexity of the model typically grows with improving accuracy. The desired accuracy depends, of course, on the application. Simulations of analog ICs usually require more accurate models than simulations of digital ones. Moreover, analysis of sub-micron devices also requires complex models that account for numerous phenomena affecting small-geometry MOSFET operation [1].

Improvement of accuracy or increase of complexity usually results in a large number of model parameters (e.g., [2]). The parameters may be divide into several groups, such as primary parameters responsible for the main phenomena in the device (e.g., threshold voltage, mobility), secondary parameters describing less important phenomena (e.g., mobility degradation coefficients), parameters describing parasitic components (e.g., serial resistances of the leads), and fitting parameters that often have no explicit physical meaning [2] (e.g., dependence of channel size variation on channel length or width). In the case of complex models with a large number of parameters it is often difficult to establish relationships between them. This complicates the extraction task even more.

The extraction of MOSFET parameters may also be useful for the characterization of technology used for ICs fabrication. Hence reliable methodologies and tools for the extraction of MOSFET model parameters are necessary. Several solutions are available commercially. The large IC-CAP package [3] seems to be one of the most widely used. It combines data acquisition with parameter extraction for a variety of most relevant models of semiconductor devices. More compact extractors may, however, be useful too, since they may be better suited to local needs. This paper presents the MOSTXX program (MOS transistor parameter extraction in Excel), which is an efficient tool for MOSFET parameter extraction. The program has been developed in the IET and is used for characterization of MOS technologies.

2. Parameter extraction – an overview

The extraction of MOSFET model parameters is based on the fitting of the simulated *I-V* characteristics to the corresponding experimental *I-V* data. The misfit (error) between the model and the data is defined as:

$$F_{err}(\boldsymbol{P}) = \sum_{i=1}^{N} \left[I^{mod}(\boldsymbol{V}_i, \boldsymbol{P}) - I^{mea}(\boldsymbol{V}_i) \right]^2, \qquad (1)$$

where $\boldsymbol{P}$ denotes a vector of parameters, $\boldsymbol{V}_i$ is a vector of bias voltages for a given ith operation point (usually it consists of three voltages: substrate-source voltage V_{BS}, gate-source voltage V_{GS} and drain-source voltage V_{DS}), I^{mod} and I^{mea} denote simulated and measured currents at a given operation point, N is the number of operation points taken into account for extraction purposes.

The misfit definition (1) has been widely used also in sophisticated modeling and characterization tools, such as IC-CAP. The parameter vector $\boldsymbol{P}$ can be determined via minimization of the error function F_{err}. Thus parameter extraction can be considered as a nonlinear least-squares problem.

The parameters may be determined using two methods, namely:

- "locally", i.e., in a well-defined bias range (e.g., strong inversion at low drain-source voltage); only selected parameters may be determined in this way;
- "globally", i.e., taking into account different bias conditions (e.g., a complete family of $I(V_{BS}, V_{GS}, V_{DS})$ data); all model parameters may be determined simultaneously.

The computational effort of the extraction algorithm strongly depends on model linearity, number of operating points

and number of parameters. In the case of a linear model the least-squares problem is significantly easier to solve. Moreover the solution is fast and reliable. The interpretation of the parameters obtained in this way is usually straightforward. This is the case, e.g., of local extraction of the threshold voltage V_T using I–V_{GS} data in the strong inversion at low drain-source voltage. Potential difficulties may arise from the determination of the bias range where the model may be considered as linear. The "quality" of the measured data seems to be relevant at this point.
In the case of a nonlinear model the least-squares problem becomes a nonlinear programming task. This approach usually allows a set of parameters to be extracted at the same time. However, the solution is more time-consuming and strongly depends on the initial approximation of parameter values. An inappropriate starting point may lead to a local minimum of the error function. Problems may also arise from mutual relationships between parameters being searched. Threshold voltage parameters and the parameters describing channel-length modulation, for example, depend on the substrate doping concentration. This may lead to convergence problems or to non-physical values of the extracted parameters.
The local and global parameter extraction may be used independently. In practice, however, the following solution may be used. A set of local extraction algorithms is used to evaluate a set of parameters. These are next used as a starting point in global extraction.

3. MOSTXX program

MOSTXX is a tool for MOSFET parameter extraction. It has been developed in the MS Excel environment and implemented using the VBA-code. It is currently used for characterization of CMOS ICs. Its features are briefly described below.

3.1. Models

At present two MOSFET models are implemented in the MOSTXX software. These are:

- SPICE (LEVEL = 1) DC model,
- SPICE (LEVEL = 2) DC model,
- SPICE (LEVEL = 3) DC model.

Several other "sub-models" (extracted from the models mentioned above) are implemented for the purposes of local identification of parameters. These are:

- linear and parabolic functions useful for local extraction of several parameters;
- a model of the dependence of threshold voltage on substrate bias;
- a model of output conductance in saturation corresponding to the SPICE (LEVEL = 1) DC model;
- models of the effect of vertical electric field on carrier mobility corresponding to the SPICE (LEVEL = 2, 3) DC models;
- models of output conductance in saturation corresponding to the SPICE (LEVEL = 2, 3) DC models.

The models mentioned above are coded in the program as separate modules. They have a common, flexible interface to the routines that call them. This interface is defined by the vector of model parameters to be extracted, vector of auxiliary parameters that enable additional data to be transferred to the model, vector of bias voltages defining the operating point of the device, and finally the vector of first-order derivatives and the matrix of second-order derivatives of the model output variable (current, threshold voltage, conductance, etc.) versus parameters. The last two items are required by the procedures minimizing the error function.
The routines containing the VBA-codes of the model also have a common structure. Any procedure calculates the output variable and the corresponding derivatives of the output variable versus model parameters for a given operating point.
Standardization of the model-defining procedures enables straightforward implementation of a new model in the MOSTXX program.

3.2. Methods and computational tools

A number of algorithms of local and global parameter extraction are implemented in the MOSTXX program.
The linear least-squares method is used for the estimation of slopes and threshold voltages of $I(V_{GS})$ curves in strong inversion at low drain bias.
The DFP (Davidon-Fletcher-Powell) and BFGS (Broyden-Fletcher-Goldfarb-Shanno) methods [4], i.e., the quasi-Newton method of the minimization of an objective (error) function defined as a global deviation of the model from the experimental data (1), is used to extract complete sets of parameters of several models, namely the models of mobilities and complete models of *I-V* characteristics. This method seems to be very efficient, fast and reliable under the assumption that a starting point for the set of the parameters to be found is reasonably close to the final solution. This can be easily achieved using data generated with local extraction method or using parameter values obtained previously for another similar device.
The DFP, BFGS algorithms require procedures of the minimization in direction [4]. An appropriate routine, efficiently combining the Newton and bi-section methods has been implemented in MOSTXX. The DFP algorithm extensively uses a variety of vector and matrix operations. MOSTXX has been equipped with a library of procedures performing these operations. Thus the implementation of any other optimization method may be done easily.

3.3. Operation

The general flowchart of MOSTXX program is shown in Fig. 1.

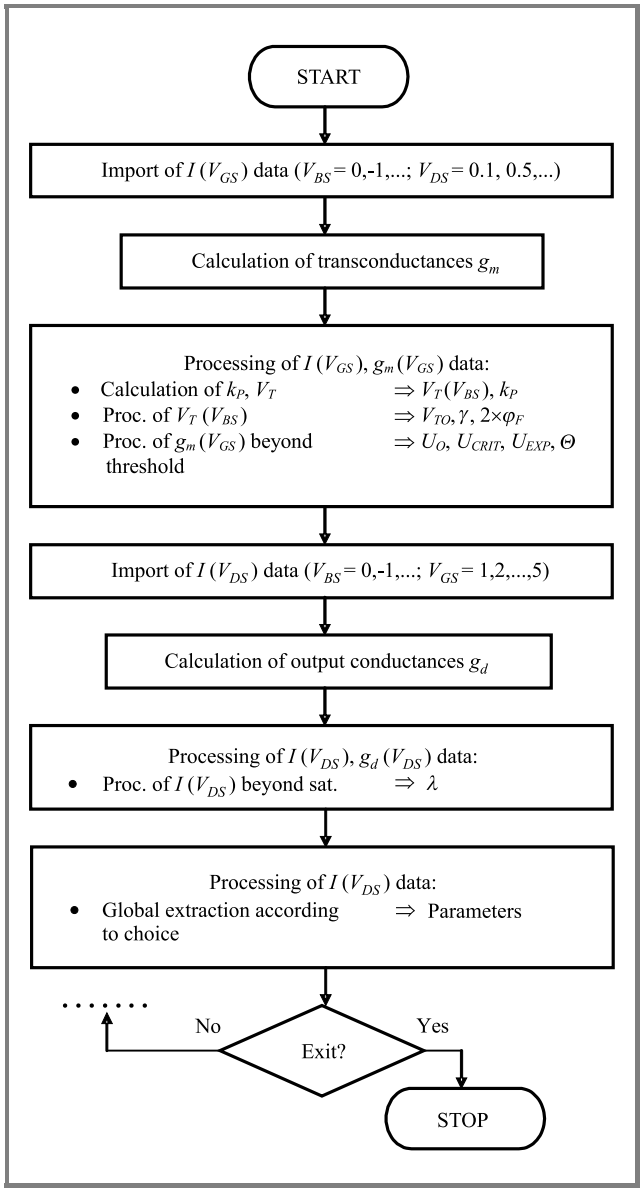


Fig. 1. Flowchart of MOSTXX; a typical sequence of extraction operations is shown.

MOSTXX operation starts with the arrangement of a workspace for the input and output data. It has been assumed that a single Excel worksheet corresponds to one family of the measured *I-V* characteristics only. It can be used for further processing of these data. Temporary as well as final results of the extraction from the measurement data are stored in this worksheet. There is also a special "summary" worksheet, where the final results of the extraction are stored in a manner enabling further statistical processing of the data. This worksheet plays a role of a local database of parameters.

Next the program displays the main form, giving the user an access to all program options. The availability of different form controls is related to the current status of data processing. This is recognized using contents of the current worksheet.
A set of operations is available after pressing appropriate controls on the form. These are briefly discussed below.

Import of measurement data. The program accepts *I-V* data in a simple ASCII format as well as data in the format used by the stand-alone METRICS measurement program [5] or semi-automatic METRICS/ELECT system [6]. The latter is particularly useful for wafer mapping purposes. In this case the program prompts the user to select the device data to be used for extraction (Fig. 2). The application automatically recognizes the structure of the *I-V* data, i.e., bias loops, number of points in loops, etc. The measurement data can be plotted. This option allows the user to accept the measured data before further processing or eventually reject them.

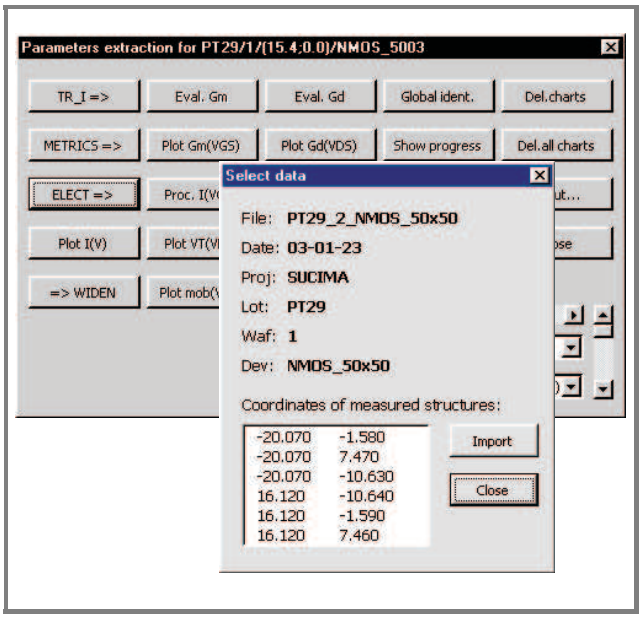


Fig. 2. Import of *I-V* data from the file written by the METRICS/ELECT system; additional form allows to select the device under test (DUT).

Calculation and visualization of transconductances. This option is used, if the local extraction of the threshold voltage or extraction of mobility and parameters of its degradation due to the strong electric vertical field are to be done. Samples of $g_m(V_{GS})$ curves are shown in Fig. 3.

Extraction of the threshold voltage and its parameters. The threshold voltage V_T is calculated using a linear fit of $I(V_{GS})$ data in the point where the corresponding transconductance curve reaches the maximum value. This processing is done for all the substrate bias V_{BS} values. Next the $V_T(V_{BS})$ characteristics are processed using the least-squares method and threshold voltage parameters are calculated (Fig. 4).

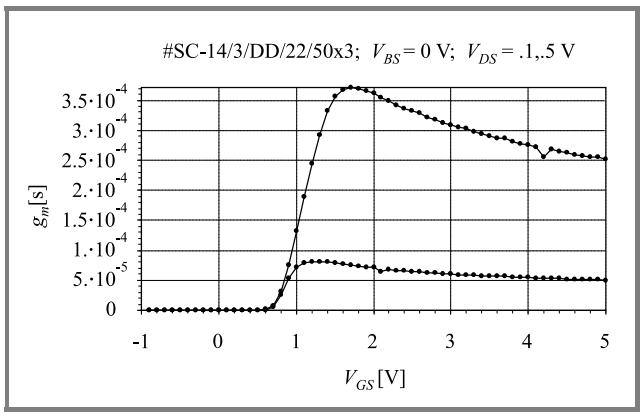


Fig. 3. Transconductances of an n-channel MOSFET ($W = 50\ \mu$m, $L = 3\ \mu$m).

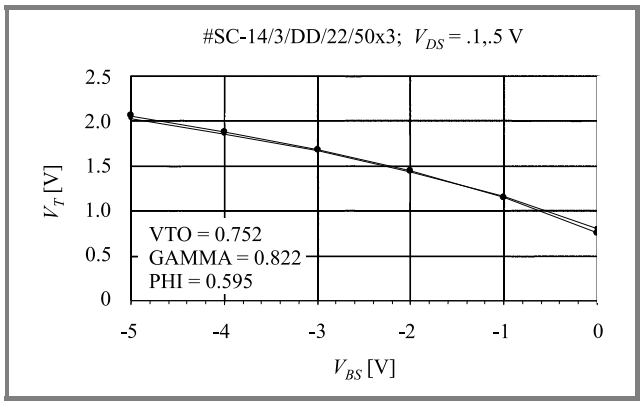


Fig. 4. Extraction of the parameters of the threshold voltage of an n-channel MOSFET ($W = 50\ \mu$m, $L = 3\ \mu$m).

Extraction of the mobility and its parameters. The low-field value of the mobility is calculated using the maximum value of the $g_m(V_{GS})$ data. Next, a range where the transconductance decreases (Fig. 3) is taken into account. The parameters of the mobility degradation are calculated using the least-squares method. Sample results are shown in Fig. 5. It is clear that the fit in the case of the SPICE (LEVEL = 2) mobility model containing 3 parameters is better than in the case of the SPICE (LEVEL = 3) one containing 2 parameters only.

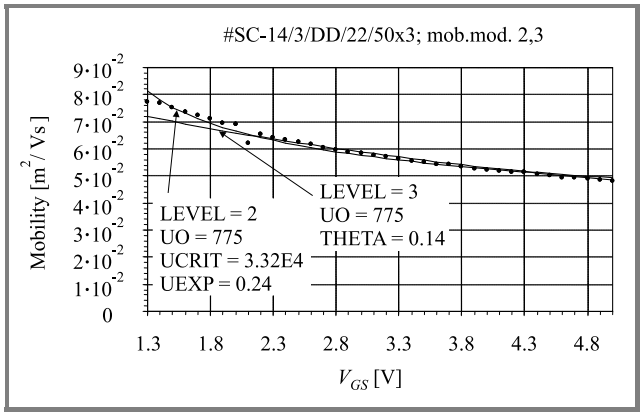


Fig. 5. Extraction of mobility parameters of an n-channel MOSFET ($W = 50\ \mu$m, $L = 3\ \mu$m) using two SPICE models.

Calculation and visualization of output conductances g_d. This option must be used, if the local extraction of MOSFET parameters in the saturation range (e.g., λ in SPICE LEVEL = 1 model) is to be done. Samples of $g_d(V_{DS})$ curves are shown in Fig. 6.

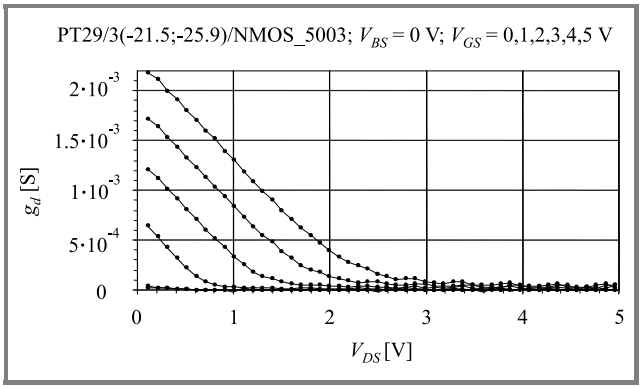


Fig. 6. Output conductances of an n-channel MOSFET.

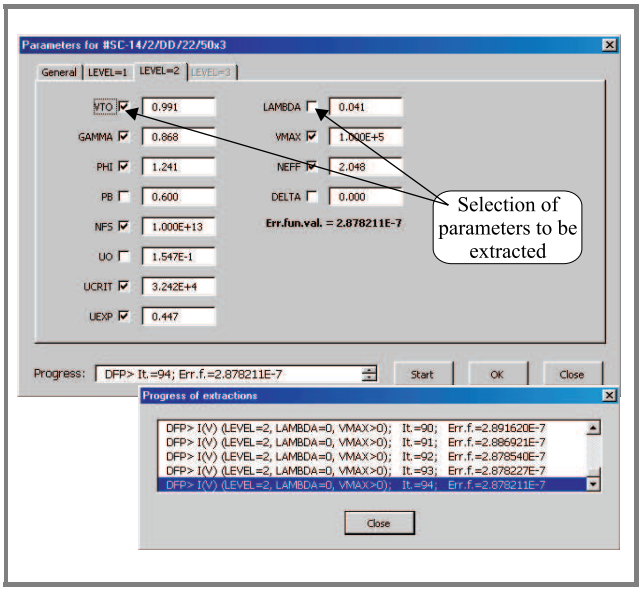


Fig. 7. Global extraction of n-channel MOSFET parameters.

Global extraction of parameters and visualization of modelled *I-V* characteristics. This is the most important option of the MOSTXX tool. It enables a set of parameter values to be determined that ensures the best fit of the model to the experimental data.

When the user selects this option an appropriate multi-tab form appears (Fig. 7). The user selects a model, choosing the appropriate tab. Then parameters to be extracted automatically should be checked. Otherwise they will be assumed constant during the minimization of the error function (1). The whole set of parameters may also be determined "manually". After pressing the "OK" button a misfit (error function) value is displayed. Simultaneously,

if model characteristics charts are embedded in the worksheet, they are immediately updated. This method can be used to extract parameters "manually", as well as to quickly investigate the effect of parameter variations on the model characteristics. Pressing the "OK" button also updates the appropriate cells in the current worksheet and in the "summary" worksheet, where the final extraction results are stored.

After pressing the "Start" button the minimization of the error function is initiated. The time required for calculations depends on the initial parameter values, the number of measured points, the number of parameters to be extracted and the complexity of the model formulae. Thus the "quality" and the efficiency of the VBA code of the model is very relevant to the extraction. The user may monitor the progress of the calculations watching successive error function values displayed in a list-box control. The contents of this field are available also after calculations are finished and may be displayed on demand (Fig. 7). This option may be particularly useful in the case of a wrong starting point, when the minimization does not converge. The user can then easily set a better initial approximation of the parameter values. The calculations are usually fast (several seconds for a whole family of *I-V* data) and reliable. The final misfit value is also displayed in the form.

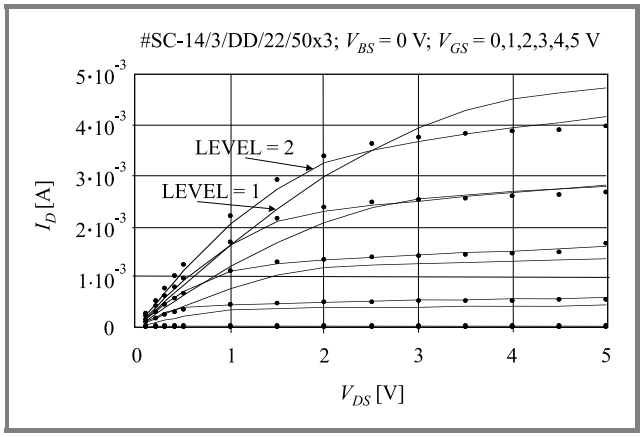


Fig. 8. Results of global extraction of a n-channel MOSFET ($W = 50\ \mu$m, $L = 3\ \mu$m) parameters.

As has been mentioned above, the *I-V* characteristics of the model may be plotted together with the experimental data. A sample of extraction results is shown in Fig. 8 for a n-channel MOSFET with channel dimensions: $W = 50\ \mu$m, $L = 3\ \mu$m. One can see that the results of the extraction based on the LEVEL = 2 model is quite satisfactory, while the fitting in the case of the LEVEL = 1 model is much worse. The results of parameter extraction for a p-channel MOSFET ($W = 50\ \mu$m, $L = 50\ \mu$m) parameters are shown in Fig. 9. In this case a good fit has been obtained using both models.

The most efficient way to use the MOSTXX software for parameter extraction is to perform first a series of local extractions of the threshold voltage and its parameters, mobility and its parameters and output conductance in the saturation range. The results of these calculations can be then used as the starting point for global parameter extraction.

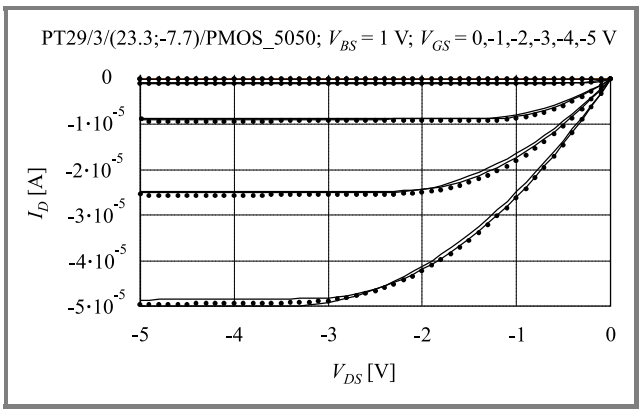


Fig. 9. Results of the global extraction of a p-channel MOSFET ($W = 50\ \mu$m, $L = 50\ \mu$m) parameters.

3.4. Other features

The MOSTXX program operates in the MS Excel environment. Thus all Excel options are available. Additional non-standard calculations may be done easily.

The following MOSTXX advantages can be mentioned: good convergence, computational efficiency, the ability to perform calculations using incomplete data, for example only global parameter extraction may be done if $I(V_{DS})$ data is available for a single V_{BS} voltage only. Thus the manual mode of operation is very useful.

4. Future work

The presented tool is undergoing further development. The following goals are to be reached:

- Implementation of the BSIM3 model, which became a standard solution in the area of simulation of CMOS ICs.
- Implementation of BSIM3 SOI model; this task may be relevant if CMOS SOI technology is be considered as a standard technology in the IET laboratory.
- Establishment of a link to the statistical process control (SPC) module.
- Translation of the VBA code into Delphi or C (a database project would then be required).

It is expected that MOSTXX will be a very useful tool for the characterization of MOS technologies in the IET. Due to its efficiency it could even be used as an on-line postprocessor in the measurement programs.

5. Conclusion

An efficient and versatile tool for MOSFET parameter extraction combining a set of identification methods has been developed in IET. Two standard MOSFET models are implemented in MOSTXX in the full form. The third one is implemented partially. Samples of the MOSTXX operation are shown. Further development of the program is planned.

Acknowledgment

Partially supported by E.C. Contract GRID-CT-2001-000561.

References

[1] L. Łukasiak, *Modele i elektryczne metody charakteryzacji przyrządów MOS i MOS SOI*. Warszawa: WPW, 2002 (in Polish).

[2] BSIM4 model specification, http://www device.eecs.berkeley.edu/ bsim3/ bsim4.html

[3] "IC-CAP parameter extraction and device modeling software", http://eesof.tm.agilent.com/ products/85190a-a.html

[4] W. Findeisen, J. Szymanowski, and A. Wierzbicki, *Teoria i metody obliczeniowe optymalizacji*. 2nd ed., Warszawa: PWN, 1980 (in Polish).

[5] "METRICS documentation, METRICS technology", http://www.metricstech.com

[6] J. Zając, J. Wójcik, A. Kociubiński, and D. Tomaszewski, "Semi-automatic test system for characterization of ASIC/MPWs", in *6th Symp. Diagn. & Yield 2003*, Warsaw, Poland, 2003.

Daniel Tomaszewski – for biogaphy, see this issue, p. 93.

Andrzej Kociubiński, Jacek Marczewski, Krzysztof Kucharski – for biogaphy, see this issue, p. 92.

Krzysztof Domański – for biogaphy, see this issue, p. 91.

Piotr B. Grabiec – for biogaphy, see this issue, p. 45.

Paper

Standardization of the compact model coding: non-fully depleted SOI MOSFET example

Władysław Grabiński, Daniel Tomaszewski, Laurent Lemaitre, and Andrzej Jakubowski

Abstract— The initiative to standardize compact (SPICE-like) modelling has recently gained momentum in the semiconductor industry. Some of the important issues of the compact modelling must be addressed, such as accuracy, testing, availability, version control, verification and validation. Most compact models developed in the past did not account for these key issues which are of highest importance when introducing a new compact model to the semiconductor industry in particular going beyond the ITRS roadmap technological 100 nm node. An important application for non-fully depleted SOI technology is high performance microprocessors, other high speed logic chips, as well as analogue RF circuits. The IC design process requires a compact model that describes in detail the electrical characteristics of SOI MOSFET transistors. In this paper a non-fully depleted SOI MOSFET model and its Verilog-AMS description will be presented.

Keywords—Verilog-AMS, compact model coding, SOI MOSFET.

1. Introduction

Most compact models developed in the past did not account for several key issues that are of highest importance when introducing a new compact model to the semiconductor industry in particular going beyond the ITRS roadmap technological 100 nm node [1]. Compact models were developed in several different ways by dedicated modelling groups including university research, internal company research and developments. Individual groups were targeting different domains and providing application specific solutions. Each modelling group uses individual tools and techniques for these modelling tasks, which are in many cases strongly dependent on the selected target simulation tool. At this time there is no standardized language to formulate compact model equations. Most developers write first drafts of their device models in interpreted languages. Most currently used interpreted modelling languages are those offered by commercial interactive mathematical tools. Other researchers prefer to use open-source interpreters and compilers.

Generally speaking, compact device modelling suffers from a lack of well-established, standardized methodology. Moreover, neither approach provides the option to share the code between different modelling teams and industrial partners. Nevertheless, to provide reliable industrial solution a compact model and its description have to be standardized. Development of compact device models involves different tasks which could be summarized as follows: building physics based constitutive equations; encoding constitutive equations in computer language; implementing the code into electrical simulators; validating compact device model implementation. In this paper, we exploit hardware behaviour languages (HDL) as the basis for device model developments. More specifically, we demonstrate how compact modelling can be standardized using Verilog-AMS [2], one of the most popular HDL. Already at the model coding phase, the developers can evaluate the validity and robustness of their model using any electrical circuit simulator that supports Verilog-AMS.

The current standardization efforts include models for bulk and SOI MOSFETs, and Si/SiGe BJT technologies. The silicon-on-insulator CMOS integrated circuits (ICs) offer many advantages as compared to their conventional bulk silicon counterparts. An important application for non-fully depleted SOI technology is high performance microprocessors, other high speed logic chips as well as analogue RF circuits. The IC design process requires a compact model that describes in detail the electrical characteristics of SOI MOSFET transistors. In this paper the non-fully depleted SOI MOSFET model [3] and its Verilog-AMS description will be presented.

2. Compact modelling with Verilog-AMS

Verilog-AMS provides a new dimension of design, simulation, and testing capability for electronic systems. The powerful features of Verilog-AMS language can be used for fast testing of analog and mixed-signal systems. The analog and mixed signal languages, such as Verilog-AMS, VHDL-AMS and Verilog-A, are primarily viewed as simulation tools for mixed-signal and mixed-domain systems. However, the versatility of the behaviour languages at the device-level abstraction allows to apply them to compact modelling as well.

The Verilog-AMS is a superset of Verilog and Verilog A, respectively. Besides the digital components, the language also supports analog behavioural constructs. The behavioural languages for digital system modelling, i.e., Verilog-Digital or Verilog-D proceed with the discrete change in signals at discrete points in time. However, Verilog-AMS simulation takes place in the analog

domain and all the differential algebraic equations (DAEs) are solved at every point in time. The generalized form of Kirchhoff's potential (KPL) and flow laws (KFL) formulates the DAEs.
Each model may require a different approach, yet some of the general guidelines can be followed to make the process systematic.

2.1. Analog behavioural modelling

The behavioural description with the Verilog A language can be used to represent different types of behaviours, including linear, nonlinear, integro-differential and analog event driven or any combination thereof. It is important to note that it is valid in the time domain and encapsulates a large signal behaviour. Performing linearization of the large signal model around its operation point allows small signal AC analysis to be performed as well. All behavioural models strongly depend on the understanding of the model, its formulation for the well defined regions of valid operation. The model must be stable as well as continuous in the description of its main region of operation. The model developer is responsible for the model stability. The Verilog-A language provides capabilities to effectively handle model non-continuity but still relies on the developer for recognizing and utilizing these capabilities.
In the Verilog-A language, all analog behaviour description is defined within the analog statement. That statement encompasses all necessary statements used to describe interrelations of input and output signals of a given module. The analog statement is used to define the behaviour model in terms of contributions statements, control flow and/or analog events statements. Valid syntax is shown below:

analog begin
    <statements>
end

The groups of the statements within the analog block are processed sequentially in the given order and at each time step during a transient simulation. This formulation allows the model developer to define the flow of control within the block module which has implications in the formulations of the analog behaviours for stability and robustness.
Parameter declarations are extensions of the basic variable type definitions supported by Verilog-A and have similar meaning to other programming languages. In addition, parameter definition supports an extended declaration syntax for range checking which allow developers to control acceptable parameter ranges or parameter values. Specifying the default value and valid range of its values assigned during model initialization, the user is able to restrict the parameter values to guarantee proper use of the model. The basic parameter syntax is shown below:

parameter real W= 1E-6 **from**(0.0 **: inf**);

The parameter W has the value of 1 μm, which is considered constant during model evaluations. An optional range specification limits any inputs of W between zero (0.0) and infinity (*inf*) but excluding upper and lower limits.

```
'include "disciplines.h"
//
// Module declaration
module soi(d, fg, s, bg)
//
// Terminal declaration
input      d, fg, s, bg;
electrical d, fg, s, bg;
//
// Parameter declaration
parameter real L= 1E-6
from [0.0:inf];
parameter real W= 1E-6
from [0.0:inf];
//
// Internal variables declaration
real Gammaf, EpsSi, Qdopm, Coxf;
//
// Analog block
analog begin
//
Vds=V(d,s); //Access function
//
//
Gammaf = sqrt(2.0*EpsSi*QDop)/Coxf;
//
Idc(s,d) < + TotIS + TotID;
end //analog
endmodule
```

Fig. 1. A Verilog-A template for a compact MOSFET model.

The flows and potentials on nets, ports, and branches in Verilog-AMS can be accessed using *access functions*. The name of the access function is taken from the discipline of the net, port, or branch associated with the signal (Fig. 1), i.e.,

Vds = **V(**d, s**)**;

The function above creates an unnamed branch from d to s (if it does not already exist) and then accesses the branch flow. In the following example $I(d)$ does the same from d to the global reference node (*ground*).
In Digital Verilog, the behavioural description always uses blocks, integer variables and bit signals. With Verilog-A, the behavioural description uses analog block, integer and real variables and analog signals. Analog signal is declared following a discipline. A discipline is an abstract data type that describes a continuous domain by associating the potential and flow nature. To assign an expression to a signal, Verilog-A introduces the *contribution operator* $<+$. It is used in a contribution statement where a behaviour relation is described between input and output signal:

output_signal $<+$ f(input_signal);

The generalized form of the contribution statement above consists of two parts, a left-hand side, and a right-hand

side, separated by the *contribution operator* $<+$. Any real-number expression may be used as the right-hand side. The left-hand side specifies the source branch signal that the right-hand side is to be assigned to. It must consist of an access function applied to a branch. The access function within the expression function may be used in any mode of operation: linear, nonlinear, algebraic or dynamic, as well as constants and parameters. It can be applied to any nodes or ports following the corresponding discipline definition.
The example below illustrates the assignment of the gain of the amplifier without including frequency effects:

```
V(out) <+ -gain * V(in)
```

This type of functional models is used for top level system architecture design and analysis but can be easily applied to the transistor level modelling.

2.2. Multi-discipline description

The Verilog-A language can support the description of systems used in many domains, such as electrical, mechanical, fluid dynamics and thermodynamics. For this purpose, the language provides a standard definition file where the main disciplines are defined. In the same module, nodes from different disciplines can be mixed.
Several semiconductor companies provide Verilog-AMS models of their products for simulation purposes. In addition, most of the common digital as well as analog blocks are freely available on the world wide web in the form of HDL models. If these HDL codes are converted to Verilog-AMS, they can be extremely useful for testing the analog ICs, particularly the ICs under development. Following sections show Verilog-AMS applications to compact modelling of the non-fully depleted SOI MOSFET modelling.

3. Non-fully depleted SOI MOSFET

The non-fully depleted SOI MOSFETs model presented in the paper is based on the general formula (1):

$$i_S(t) + i_{Gf}(t) + i_D(t) + i_{Gb}(t) = 0\,, \tag{1}$$

where $i_S(t)$, $i_{Gf}(t)$, $i_D(t)$, $i_{Gb}(t)$ denote currents flowing into the device at any time instant t. The equation (1) is to be solved for a floating body-source voltage $v_{BS}(t)$, which is the main variable of the model. The condition (1) expresses the overall electrical neutrality of the MOSFET. The formula (1) is illustrated in Fig. 2, which shows the approximate spatial distribution of paths of carriers in the device. This diagram also shows the main phenomena relevant to the non-fully depleted SOI MOSFET operation and accounted for in the model:

– transport at the Si-SiO_2 interface ($I_{c,f/b}$ – diffusion/drift at the front/back Si-SiO_2 interface);

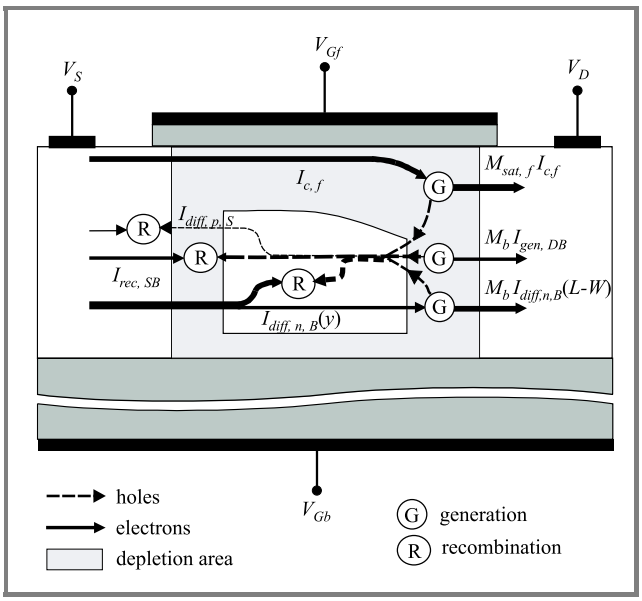


Fig. 2. Approximation of the distribution of currents components in the non-fully depleted SOI MOSFET.

– diffusion of electrons in the floating body ($I_{diff,n,B}$);

– thermal generation/recombination in the space-charge areas ($I_{rec,SB}$, $I_{gen,DB}$);

– avalanche ionization in the regions with strong electric field (M_b in the case of the drain-body junction, $M_{sat,f/b}$ in the case of the pinch-off regions);

– displacement currents (in the case of AC analysis only).

The DC model has been developed as a solution of (1) using the sinusoidal steady-state analysis (S^3A) method [4]. According to this method any time-dependent variable in the device may be expressed as:

$$u(t) = U + u^* \cdot \mathrm{e}^{\mathrm{j}\omega t} \tag{2a}$$

or

$$v(y,t) = V(y) + v^*(y) \cdot \mathrm{e}^{\mathrm{j}\omega t}\,, \tag{2b}$$

where: y – spatial coordinate along the channel; U, $V(y)$ – steady-state components; u^*, $v^*(y)$ – complex amplitudes.

This approach allows for efficient formulation of consistent DC and AC models of the corresponding device. These models account for identical sets of phenomena. Moreover, the AC models do not suffer from the requirement to partition the channel charge into source- and drain-related components. The AC models of nodal currents may be used for straightforward calculation of device inter-nodal admittances (conductances and capacitances), which are in general bias- and frequency-dependent.
Separation of the steady-state and non-steady-state components of (1) leads to the general expressions describing DC

Eq. (3a) and AC Eq. (3b) models of the non-fully depleted SOI MOSFET characteristics:

$$I_S + I_D = 0\,, \tag{3a}$$

$$i_s^* + i_{gf}^* + i_d^* + i_{gb}^* = 0\,. \tag{3b}$$

The solution of these two equations allows the DC and AC components of the floating body potential $v_{BS}(t)$ to be determined. The accuracy of the obtained result depends, of course, on the accuracy of DC current modelling. The model used in this paper will be presented in the next section.

3.1. Current formulation

The total DC source current I_S may be viewed as a superposition of the source-body junction current I_{SB} and the channel current $I_{c,f/b}$:

$$I_s = I_{SB} - I_{c,f/b}\,. \tag{4}$$

Similarly, the total DC drain current I_D may be considered as the difference between the drain-body junction current I_{DB} and the channel current $I_{c,f/b}$ multiplied by the appropriate factor of avalanche multiplication within the "pinch-off" region:

$$I_D = -I_{DB} + M_{sat,f/b} I_{c,f/b}\,. \tag{5}$$

A simple model of the inversion layer current has been used. For simplicity only front inversion layer currents are described below. In general similar expressions are valid also for back interface. The non-saturation region is described with the simple model:

$$I_{c,f} = \frac{W\mu_{c,f}C_{ox,f}}{L}\left(V_{GfS} - V_{TH,f} - \frac{V_{DS}}{2}\right)V_{DS}\,, \tag{6}$$

where: L, W – channel length and width, respectively, $\mu_{c,f}$ – channel mobility at the front interface, $V_{TH,f}$ – front gate threshold voltage.

In saturation the model is as follows:

$$I_{c,f} = \frac{W\mu_{c,f}C_{ox,f}}{L - \Delta L_{sat,f}}\left(V_{GfS} - V_{TH,f} - \frac{V_{DSsat,f}}{2}\right)V_{DSsat,f}\,, \tag{7}$$

where: $V_{Dsat,f}$ – saturation voltage for the front MOS structure, $\Delta L_{sat,f}$ – front channel length reduction in saturation. These variables are defined as follows:

$$V_{DSsat,f} = V_{GfS} - V_{TH,f} + L\,E_{sat,f} - \sqrt{\left(V_{GfS} - V_{TH,f}\right)^2 + \left(LE_{sat,f}\right)^2}\,, \tag{8}$$

$$\Delta L_{sat,f} = \sqrt{\frac{2\varepsilon_{\mathrm{Si}}}{qN_B}\left(V_{DS} - V_{DSsat,f}\right) + \left(\frac{\varepsilon_{\mathrm{Si}}E_{sat,f}}{qN_B}\right)^2} - \frac{\varepsilon_{\mathrm{Si}}E_{sat,f}}{qN_B}\,. \tag{9}$$

In the above equations $E_{sat,f}$ denotes the critical field resulting in the saturation of carrier velocity at the front interface.

The source-body junction current consists of two components: the diffusion current at the "source" border of the quasi-neutral region of density $J_{diff,n,B}$, and the recombination current originating from the source-body junction space-charge region. The area of the source-body junction may be calculated as the product of the channel width W and the difference between body thickness t_{Si} and depletion widths formed by the front and back gates (W_{GfB} and W_{GbB}, respectively). The depletion widths depend on the gate-body voltage (in the depletion and weak inversion range) or body-source voltage (in the strong inversion range). Thus the following formula for the source-body junction current can be formulated:

$$I_{SB} = W\left(t_{\mathrm{Si}} - W_{GfB} - W_{GbB}\right) \times \left[J_{diff,n,B}(W_{SB}) - qW_{SB}\langle R_{th,SB}\rangle\right]. \tag{10}$$

In the above formula W_{SB} is the source-body junction width whereas $\langle R_{th,SB}\rangle$ denotes a mean value of the recombination rate in the source-body junction. It is described with formula (11) derived elsewhere [6]:

$$\langle R_{th,SB}\rangle = \frac{n_i}{2\tau_j}\cdot\frac{\sqrt{\mathrm{e}^{-u_{SB}} - 1}}{u_{bi} + u_{SB}} \times 2\,\mathrm{arctg}\left|\frac{\dfrac{N_{SD}}{n_i} - \dfrac{n_i}{N_B}\mathrm{e}^{-u_{SB}}\sqrt{\mathrm{e}^{-u_{SB}} - 1}}{\mathrm{e}^{-u_{SB}} - 1 + \left(\dfrac{N_{SD}}{n_i} + 1\right)\left(\dfrac{n_i}{N_B}\mathrm{e}^{-u_{SB}} + 1\right)}\right|, \tag{11}$$

where: n_i – intrinsic carrier concentration, N_{SD} – source/drain doping, τ_j – junction lifetime, u_{SB}, u_{bi} – source-body and built-in voltage, respectively, normalized to thermal voltage $V_t = \mathrm{k}T/q$.

The diffusion current in the quasi-neutral region is described in the following way:

$$J_{diff,n,B}(y) = q\frac{D_{n,B}}{L_{n,B}}\frac{n_i^2}{N_B} \times \frac{\mathrm{e}^{-u_{DS}-u_{SB}}\cdot\mathrm{ch}\dfrac{y - W_{SB}}{L_{n,B}} - \mathrm{e}^{-u_{SB}}\mathrm{ch}\dfrac{L - W_{DB} - y}{L_{n,B}}}{\mathrm{sh}\dfrac{L - W_{SB} - W_{DB}}{L_{n,B}}}\,, \tag{12}$$

where: $D_{n,B}$ – diffusion coefficient, $L_{n,B}$ – diffusion length, W_{DB} – drain-body junction width.

The drain-body junction current consists of two components: the diffusion current at the "drain" border of the quasi-neutral region and the recombination current originating from the drain-body junction space-charge region.

Moreover, these two currents are multiplied within the drain-body junction depletion area. Thus the drain-body junction current can be described in the following way:

$$I_{DB} = W(t_{Si} - W_{GfB} - W_{GbB})M_{DB} \times \left[J_{diff,n,B}(L - W_{DB}) - qW_{DB}\langle G_{th,DB}\rangle\right]. \quad (13)$$

In the above formula $\langle G_{th,DB}\rangle$ denotes the mean value of the generation rate in the drain-body junction. Formula $\langle G_{th,DB}\rangle$ is described with formula (14) derived in [6]:

$$\langle G_{th,DB}\rangle = \frac{n_i}{2\tau_j}\frac{\sqrt{1-e^{-u_{DS}-u_{SB}}}}{u_{bi}-u_{DS}-u_{SB}} \times \left(\ln\left|\frac{\frac{N_{SD}}{n_i}+1-\sqrt{1-e^{-u_{DS}-u_{SB}}}}{\frac{N_{SD}}{n_i}+1+\sqrt{1-e^{-u_{DS}-u_{SB}}}}\right| + \ln\left|\frac{\frac{n_i}{N_B}e^{-u_{DS}-u_{SB}}+1+\sqrt{1-e^{-u_{DS}-u_{SB}}}}{\frac{n_i}{N_B}e^{-u_{DS}-u_{SB}}+1-\sqrt{1-e^{-u_{DS}-u_{SB}}}}\right|\right), \quad (14)$$

where: u_{DS} – drain-source and voltage, normalized to thermal voltage $V_t = kT/q$; other variables have the same meanings as in the case of $\langle R_{th,SB}\rangle$ (see Eq. (11)).

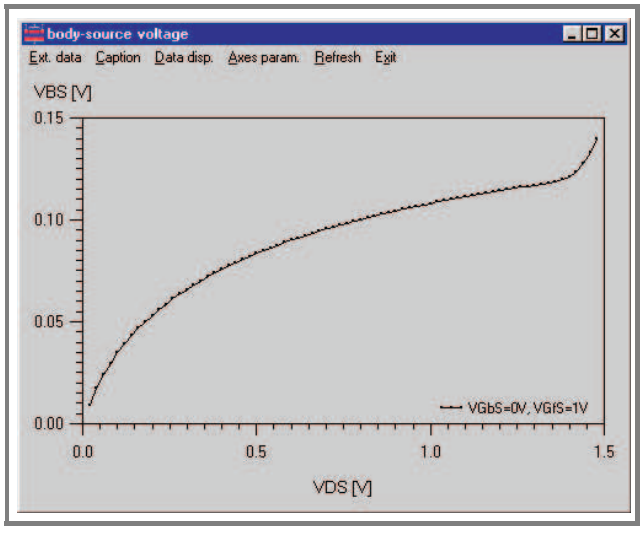


Fig. 3. Body-source voltage versus drain bias in the model of a non-fully depleted SOI MOSFET.

Rather complex formulae describing thermal generation/recombination currents in the space-charge regions have been derived using Shockley-Read-Hall generation/recombination model under the assumption of quasi-linear electric field distribution inside the space charge area. The main reason for this approach was to develop a closed-form model valid particularly for low bias conditions. In the case of a small V_{DS} voltage generation/recombination model is very important for obtaining the solution of Eq. (3a).

Figure 3 shows that in the presented model the iterative solution of this equation leads to a proper source-body voltage behaviour regardless of the drain bias. This is not the case with several other models of non-fully depleted devices.

3.2. Charge definition

After the solution of Eq. (3a) (V_{BS} voltage) was found all other variables describing the device operation can be calculated. Finally charges associated with the source (Q_S), front gate (Q_{Gf}), drain (Q_D) and back gate (Q_{Gb}) electrodes can be obtained in the presented model in a typical way:

$$Q_S = W\left[\int_0^L\left(1-\frac{y}{L}\right)q'_{c,f}(y)dy + \int_0^L\left(1-\frac{y}{L}\right)q'_{c,b}(y)dy\right], \quad (15)$$

$$Q_{Gf} = W\int_0^L q'_{Gf}(y)dy, \quad (16)$$

$$Q_D = W\left[\int_0^L\frac{y}{L}q'_{c,f}(y)dy + \int_0^L\frac{y}{L}q'_{c,b}(y)dy\right], \quad (17)$$

$$Q_{Gb} = W\int_0^L q'_{Gb}(y)dy, \quad (18)$$

where $q'_{Gf}(y)$, $q'_{Gb}(y)$ denote densities (per unit area) of the gates charges, whereas $q'_{c,f}(y)$, $q'_{c,b}(y)$ are densities of the inversion layer charges at the front and back Si-SiO_2 interfaces. The densities of the charges corresponding to the front interface are given with the following formulae:

$$q'_{c,f}(y) = -C_{ox,f}\left[V_{Gf} - V_{FB,f} - \psi_{s,f}(y) + \frac{q'_{B,f}(y)}{C_{ox,f}}\right], \quad (19)$$

$$q'_{G,f}(y) = C_{ox,f}\left[V_{Gf} - \Phi_{MS,f} - \psi_{s,f}(y)\right]. \quad (20)$$

Both expressions mentioned above require surface potential distribution $\Psi(y)$. It depends on gate bias. In the case of accumulation, depletion and weak inversion conditions it is almost constant along the gate, whereas in the case of strong inversion conditions it follows the Fermi

quasi-level. Analogous expressions are of course valid for the back interface.

4. Automatic device model synthesizer modelling example

This section concludes the paper by describing the simulator implementation of the non fully-depleted SOI MOSFET model using the automatic device model synthesizer (ADMS) tool [4].

Creating a new SPICE model consists of two main steps: defining the parameters that the user will enter from the net list or schematic level, writing the model *c*-code itself. Following these steps a new SPICE model can be used in linear, nonlinear (i.e., harmonic balance), transient and circuit envelope simulation modes depending on the simulator capabilities. The ADMS processes both steps. Based on the Verilog-AMS compact model description [2, 3], the ADMS tool generated all necessary *c*-code to handle the model and its instance parameters, model codes including all additional functions, as well as the required derivatives with respect to terminal voltages. The ADMS tool also generates the needed make files to simplify the procedure of compiling and linking

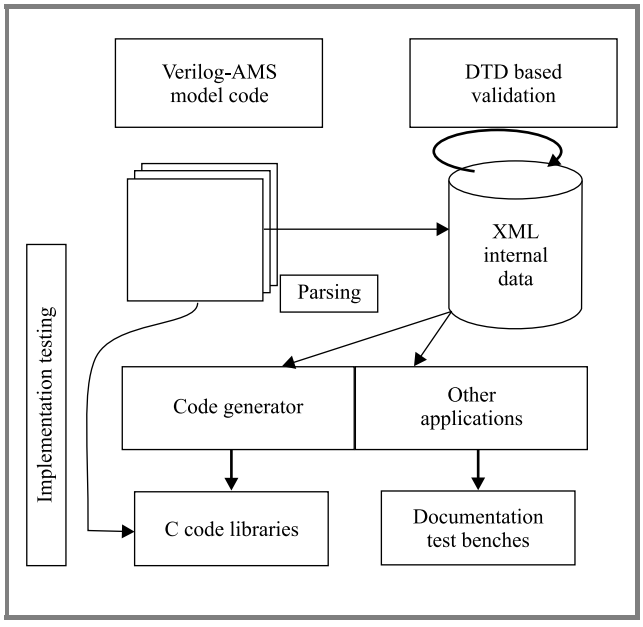


Fig. 4. Basic component of the ADMS system and internal data flows.

the new model with the simulator. ADMS reduces the implementation efforts of compact device definition using Verilog-AMS model description. At the same time, it offers a way to substantially improve the robustness of new compact device models. Implementation of the same model across different electrical circuit simulators is automated [7, 8].

Figure 4 shows basic component of the ADMS tool and illustrates the internal data flow. The main data representations used in ADMS are based on XML, which is the universal format for structured documents and data on the world wide web. XML provides a large set of technologies that simplify the design of robust, re-usable code. It offers a simple way to validate internal data. Rules that organize valid data are written in a subset of XML. The set of rules forms the so-called document type definition (DTD). Embedding of an external DTD is possible. This feature favors the re-use of well established formats. Technologies built around XML give a powerful means to manipulate XML data.

5. Conclusion

The behavioural modelling futures of the Verilog-AMS merged with the ADMS tool [4] offers an excellent compact modelling environment. It facilitates faster development of advanced models and allows faster implementation into commercial IC design tools. Existing models could be smoothly extended to include important effects such as RF and thermo-dynamical effects. Furthermore, developers of new compact models now have access to the a coherent and highly reliable modelling framework simplifying model evaluation procedures and verification tasks across different simulation platforms and operating systems.

References

[1] ITRS Roadmap Update, 2003, http://www.public.itrs.net

[2] Open Verilog International, "Verilog-AMS, Language Reference Manual", Version 1.9, 1999, http://www.accellera.org/

[3] D. Tomaszewski, "Consistent DC and AC models of non-fully depleted SOI MOSFETS in strong inversion", in *Proc. 9th Int. Conf. Mix.-Sig. Des. Integr. Cir. Syst. MIXDES*, Wrocław, Poland, 2002, pp. 111–114.

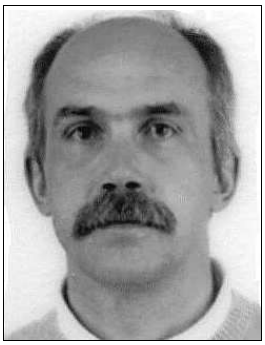
[4] L. Lemaitre, C. McAndrew, and S. Hamm, "ADMS – automatic device model synthesizer", in *Proc. IEEE CICC 2002*, Florida, USA, 2002, pp. 27–30.

[5] J. R. Hauser, "Small signal properties of field effect devices", *IEEE Trans. Electron Dev.*, vol. 12, pp. 605–618, 1965.

[6] D. Tomaszewski, "A small-signal model of SOI MOSFETs capacitances". Ph.D. thesis, Institute of Electron Technology, Warsaw, 1998.

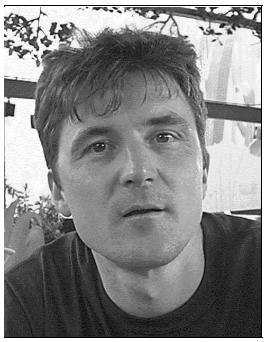
[7] L. Lemaitre, W. Grabiński, and C. McAndrew, "Compact device modeling using Verilog-A and ADMS", in *Proc. 9th Int. Conf. Mix.-Sig. Des. Integr. Cir. Syst. MIXDES*, Wrocław, Poland, 2002, pp. 59–62.

[8] C. Lallement, F. Pêcheux, and W. Grabiński, "High level description of thermodynamical effects in the EKV 2.6 most model", in *Proc. 9th Int. Conf. Mix.-Sig. Des. Integr. Cir. Syst. MIXDES*, Wrocław, Poland, 2002, pp. 45–50.



Władysław Grabiński was a Research Assistant, from 1991 to 1998, at the Integrated Systems Lab, ETH Zürich, Switzerland, supporting the CMOS and BiCMOS technology developments by electrical characterization of the processes and devices. From 1999 to 2000, he was with LEG, EPF Lausanne, and was engaged in the compact MOSFET model developments supporting numerical device simulation and parameter extraction. Since 2000, he is Technical Staff Engineer at Freescale (launched by Motorola), Geneva Modeling Center, Switzerland. He is responsible for modeling, characterization and parameter extraction of MOS transistors for the design of RF CMOS circuits. His current research interests are in high-frequency characterization, compact modeling and device numerical simulations of MOSFETs for wireless applications. He is IEEE member and also involved in activities of the MOS-AK compact modeling group and has been serving as a coordinating manager since 1999.

e-mail: w.grabinski@freescale.com
Freescale, Geneva Modeling Center
207 route de Ferney
CH-1218 Le Grand Saconnex, Switzerland



Laurent Lemaitre is with Freescale (launched by Motorola) since 1995. His main work involves the development and support of compact modeling. He is a member of the Verilog-AMS subcommittee responsible for compact modeling extensions. He received his master's degree and in electrical engineering from ENSERG-Grenoble in 1987 and his Ph.D. degree from the Swiss Polytechnic Institute of Technology-Lausanne in 1994.

e-mail: Laurent.Lemaitre@freescale.com
Freescale, Geneva Modeling Center
207 route de Ferney
CH-1218 Le Grand Saconnex, Switzerland

Daniel Tomaszewski – for biography, see this issue, p. 93.

Andrzej Jakubowski – for biography, see this issue, p. 75.

Regular paper

A new method of frequency offset correction using coherent averaging

Piotr Gajewski, Jerzy Łopatka, and Zbigniew Piotrowski

Abstract— This paper describes a new method of frequency offset correction to improve clock stability in communication systems where the temporary drift or jitter of phase angle is not accepted. This method bases on coherent spectral averaging with a special phase scanning algorithm. Achieved results show that proposed method is effective for strongly degraded signals. Method is useful for precise phase angle reconstruction in these systems where clock stability of the transceiver and receiver is insufficient.

Keywords— coherent averaging, jitter, drift, frequency offset correction.

1. Introduction

The phase noise or fluctuation is defined as a phase shift of the output signal in relation to declared value of the input signal [1–5]. A value of the shift gives us information, how frequency of the received signal differs from the declared value. Proposed method bases on coherent averaging of signal spectrum. To fulfil the coherence conditions during averaging process, a new phase angle scanning method was elaborated and examined.

2. Coherent averaging

Coherent averaging method is used to increase the signal-to-noise ratio (SNR). It is based on multiple calculation of the signal complex spectrum determined in the time domain, and then computation of average spectrum.

Let's assume that the signal is degraded by additive noise:

$$c(n) = y(n) + noise(n)\,, \tag{1}$$

where: $c(n)$ – set of samples degraded by noise; $y(n)$ – periodically repeated signal, with fixed amplitude distribution; $noise(n)$ – n-samples set of noise realization.

The averaged signal can be expressed as:

$$c_{mean}(k) = \frac{1}{M}\sum_{n=1}^{M} c\big(k + L(n-1)\big)\,. \tag{2}$$

Here, it is assumed, that M separated, repeatable sets of samples are averaged, and the signal $y(n)$ possesses repetition period L.

Output deviation δ_{mean} for set $c_{mean}(k)$ depends on signal deviation δ_{org} before averaging and is expressed as:

$$\delta_{mean} = \frac{\delta_{org}}{\sqrt{M}}\,. \tag{3}$$

Basing on expression (3), it is possible to decrease a measurement uncertainty by decreasing the signal deviation during averaging.

Let us consider a signal with a given amplitude y and deviation δ_{org}. The SNR is here defined as:

$$SNR_{org} = \frac{y}{\delta_{org}}\,. \tag{4}$$

For the averaged signal, we have:

$$SNR_{mean} = \frac{y}{\delta_{mean}}\,. \tag{5}$$

Thus, calculated SNR gain coefficient equals:

$$SNR_{coh} = \frac{\delta_{org}}{\delta_{org}/\sqrt{M}} = \sqrt{M}\,. \tag{6}$$

The SNR gain coefficient can be expressed as logarithm value for fixed number of averaging M:

$$\begin{aligned} SNR_{coh}[\mathrm{dB}] &= 20\lg\left(SNR_{coh}\right) \\ &= 20\lg\left(\sqrt{M}\right) = 10\lg(M). \end{aligned} \tag{7}$$

In frequency domain, coherent averaging is realised as separate averaging of real and imaginary parts of particular components of power spectrum. We decided to use fast Fourier transform (FFT) as an effective method of power spectrum density (PSD) computation.

The composition of complex values of FFT bins formed in this way gives as result the coherent averaged spectrum. For given FFT bin, it can be expressed as:

$$\begin{aligned} F_{coh}(N\text{–}1) = {} & \frac{F_1(M\text{–}1)_{\mathrm{Re}} + F_2(M\text{–}1)_{\mathrm{Re}} + \ldots + F_k(M\text{–}1)_{\mathrm{Re}}}{k} \\ & + \mathrm{j}\frac{F_1(M\text{–}1)_{\mathrm{Im}} + F_2(M\text{–}1)_{\mathrm{Im}} + \ldots + F_k(M\text{–}1)_{\mathrm{Im}}}{k}\,. \end{aligned} \tag{8}$$

When frequency offset takes place, it introduces a phase shift between successive spectrum repetitions and the co-

herence condition is not fulfilled, so we cannot significantly increase the *SNR*:

$$SNR_{coh} = 20 \lg \left(\frac{A_{bin}}{noise_{bin}} \right), \tag{9}$$

where: A_{bin} – signal amplitude in given FFT bin; $noise_{bin}$ – noise envelope in given FFT bin.

3. Phase angle scanning method

To fulfil coherence condition during averaging process, a new phase angle scanning method is proposed. This method assumes that accuracy of phase restoration should be lower than:

$$\eta_{scan} = 1 \cdot 10^{-5} \,\mathrm{rad}/dl_{seg}\,, \tag{10}$$

where: dl_{seg} – length of averaged segment (frame).
Assuming that: length of averaged frames equals 512 samples, fixed sampling frequency is 44100 Hz and fixed signal duration is 1 second, the method resolution η_{scan} is:

$$\eta_{scan} = 0.00086 \,\mathrm{rad/s} = 0.05 \,\mathrm{deg/s}\,. \tag{11}$$

Algorithm for phase scanning method is shown in Fig. 1.

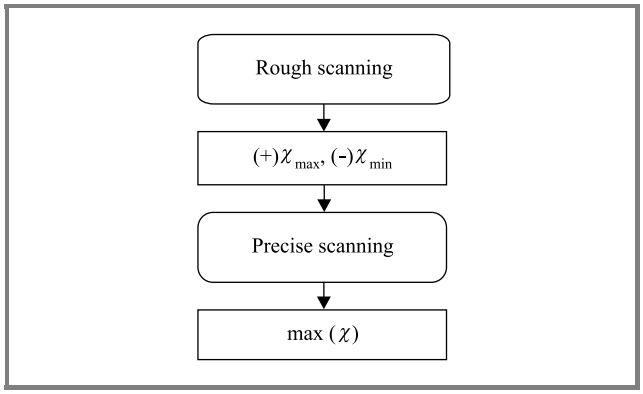


Fig. 1. Algorithm for phase scanning method.

Proposed method bases on computation of maximum value of the virtual spectral line F_v that is formed as a sum of absolute values of N spectral lines (pilots) with fixed amplitudes in given FFT bins:

$$F_v = \sum_{i=1}^{N} \left(\mathrm{abs}(F_{i\chi}) \right), \tag{12}$$

where: F_v – virtual spectral line; $F_{i\chi}$– complex value of ith pilot spectral line.
We notice that:

$$F_{i\chi} = \sum_{k=1}^{M} \mathrm{Re}(F_{ik}) + \sum_{k=1}^{M} \mathrm{Im}(F_{ik})\,. \tag{13}$$

The $F_{i\chi}$ is computed by averaging (in M-iterations) values of the spectral lines possessing the same index i. Here the real and imaginary parts of the complex values are separately averaged for each phase angle correction coefficient χ and for all M iterations.
The iterative procedure for rough and precise blocks in scanning algorithm is used (Fig. 2). Value of the virtual spectral line is computed for the selected value of correction

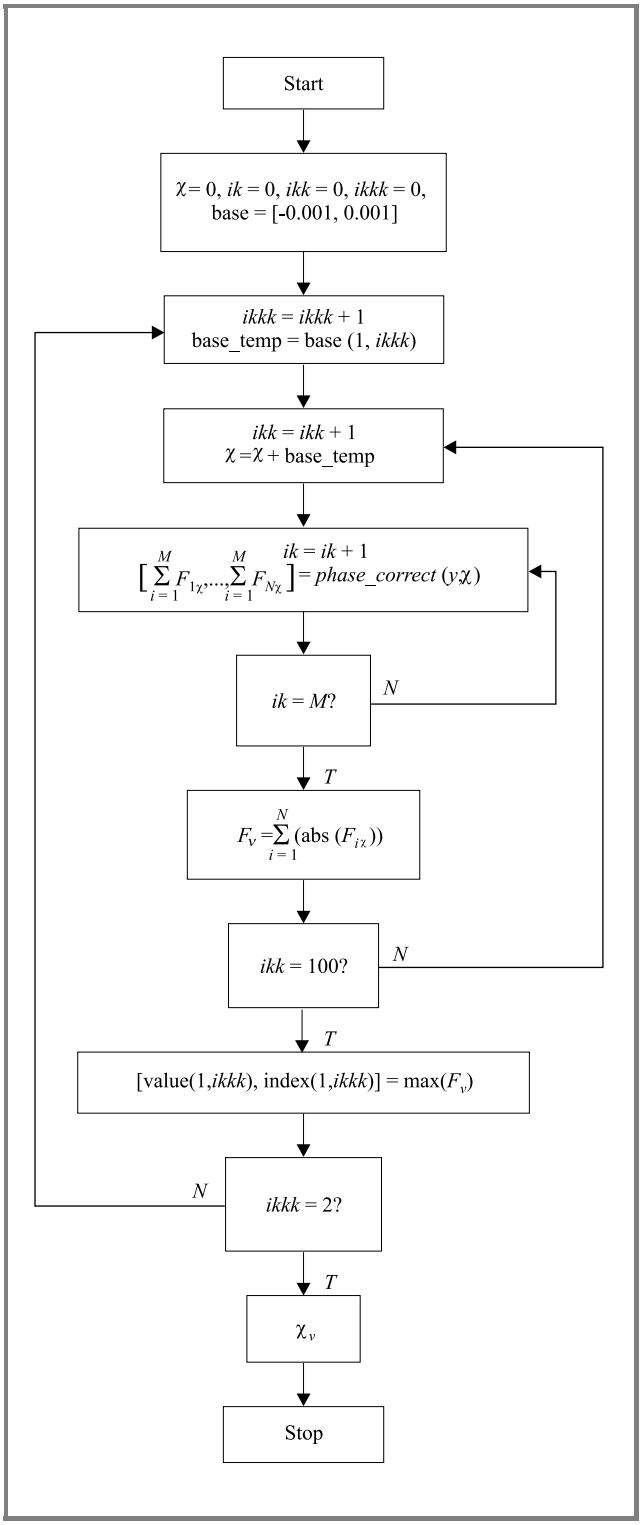


Fig. 2. Iterative algorithm used in rough and precise scanning stages.

coefficient χ [rad]. This value is constant for given iteration but it is changing between successive iterations according to the scanning range $\chi \in \langle \chi_1, \ldots, \chi_D \rangle$.

The virtual spectral line value F_v is buffered for each correction coefficient and the maximum value is found after all iterations. This maximum value represents the searched value of the phase correction. Index D is dependent on method resolution. For our laboratory experiment, it was given as: $D = D_{\mathrm{I}} + D_{\mathrm{II}} = 400$, and a scanning range was fixed as:

$$\chi \in \langle -1 \cdot 10^{-5}\,\mathrm{rad},\ 1 \cdot 10^{-5}\,\mathrm{rad} \rangle . \tag{14}$$

The two stage algorithm for searching of rough and precise values χ is proposed to decrease the computation time for declared scanning range with given resolution.

The following procedure *phase_correct* is used for phase angle correction for selected spectral line value:

$$\begin{aligned}
&[F_1, \ldots, F_N] = phase_correct(y, \chi)\\
&\mathrm{all}_fft = fft\big(y(:, ik)\big)\\
&F_{1\mathrm{temp}} = \mathrm{abs}(\mathrm{all}_fft(1, ik)).^{*} \exp(\mathrm{j}^{*}(\mathrm{angle} - \chi))\\
&F_{N\mathrm{temp}} = \mathrm{abs}(\mathrm{all}_fft(n, ik)).^{*} \exp(\mathrm{j}^{*}(\mathrm{angle} - \chi))\\
&F_1 = F_1 + F_{1\mathrm{temp}}\\
&F_N = F_N + F_{N\mathrm{temp}}
\end{aligned} \tag{15}$$

As a result, the pair of values $[(-)\chi_{\max}, (+)\chi_{\max}]$, for positive and negative phase correction coefficients is computed in a rough stage. The same procedure is carried out in precise stage where range of scanning is specified between $(-\chi_{\max})$ and $(+\chi_{\max})$.

After two stages the maximum value χ_v is obtained.

4. Experimental results

The method described above has been implemented to correct the frequency drift (jitter), which was result of clocks frequencies offset in D/A and A/D converters used in transmitter and receiver of telecommunication system.

The system transmits the same periodically repeated sequences of signal with uncorrelated noise. At the receiver, noise component is reduced by using the coherent averaging process.

Pilot signal consists of six harmonics in dedicated frequency bins, was interleaved with information spectral bins, modulated with transmitted data.

The results of the signal processing are shown in Figs. 3–9. The module of the automatic drift correction computed coefficient equals: $\chi = (-0.00040)$ rad/frame and it was used during coherent spectrum averaging of the signal. The *SNR* gain for 1148 frames was: $SNR_{coh} = 30$ dB.

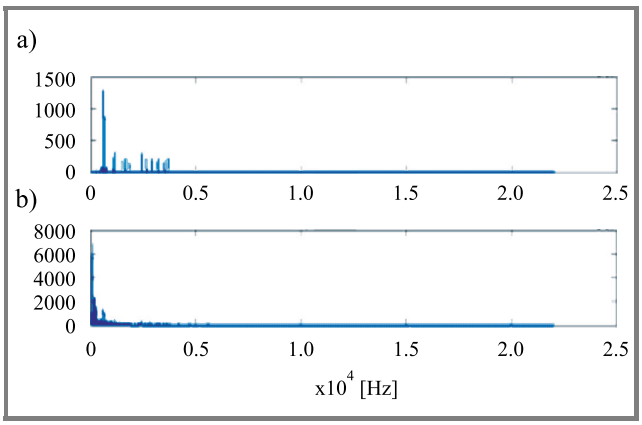


Fig. 3. Frequency domain structures of the signal: (a) original signal; (b) the same signal degraded by another signal.

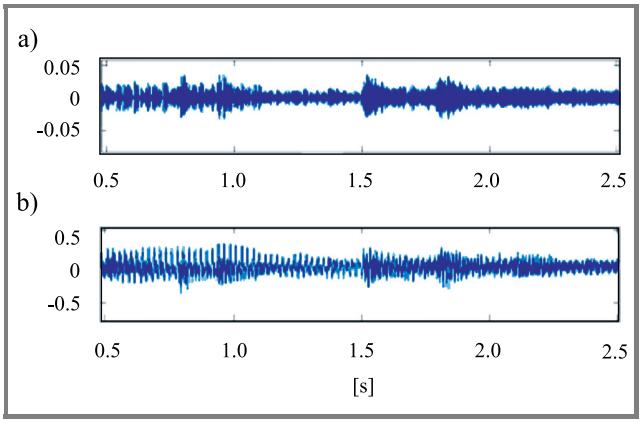


Fig. 4. Time domain structure: (a) of the original signal; (b) the same signal degraded by another signal.

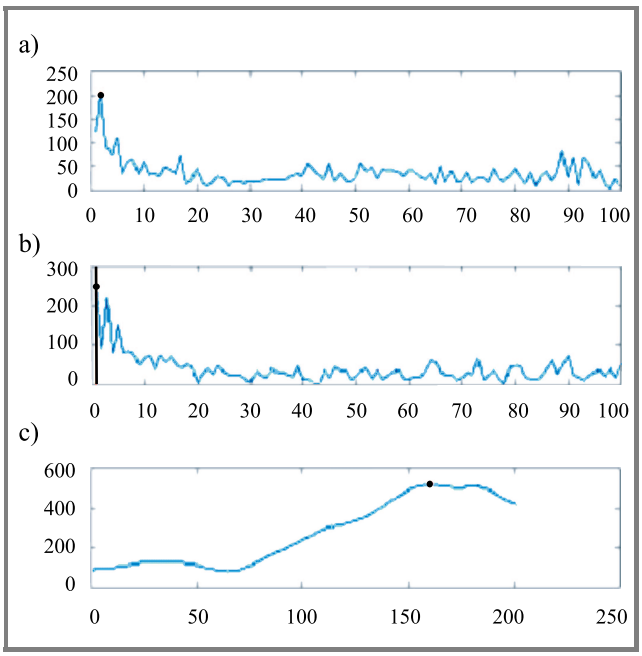


Fig. 5. Two stages of scanning method: (a),(b) rough scanning for positive and negative offsets; (c) precise scanning. Scanning results: $\chi = -0.00040$ rad/frame.

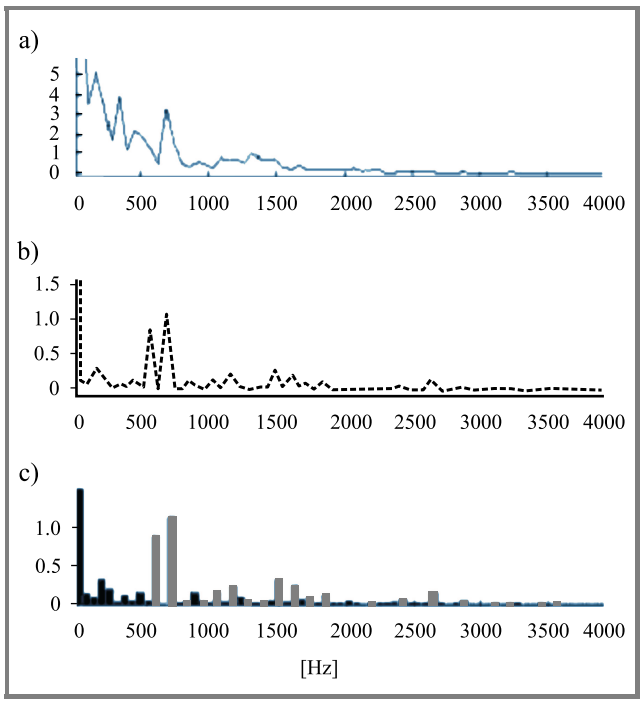


Fig. 6. Degraded spectrum signal: (a) before and (b),(c) after coherent averaging with phase correction coefficient $\chi = -0.00040$ rad/frame.

An example of the coherent averaging procedure for system with drift value: $\chi = 0.00003$ rad/frame is shown in Fig. 7.

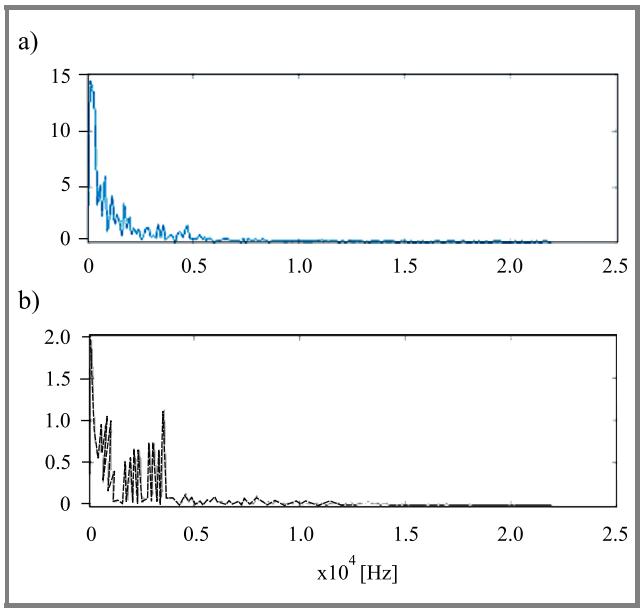


Fig. 7. Spectrum of the degraded signal before (a) and after (b) coherent averaging.

First four pilots in (Re, Im) planes before and after correction with value $\chi = -0.00040$ rad/frame are shown in Figs. 8 and 9. The signal duration was 10 seconds. Correction coefficient was found using scanning phase method.

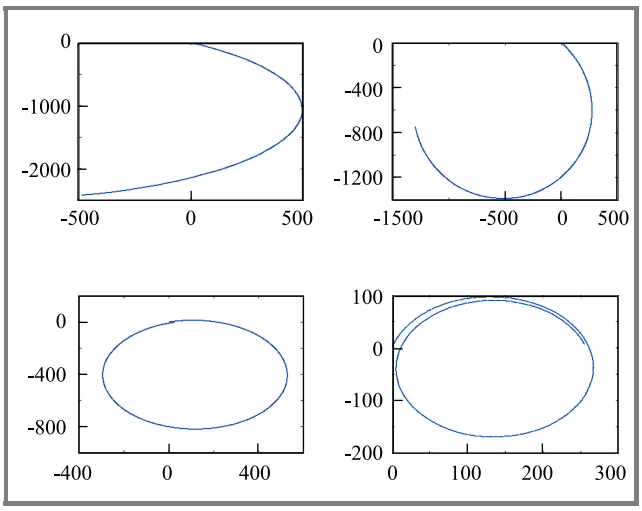


Fig. 8. First four pilots in (Re, Im) planes before phase correction.

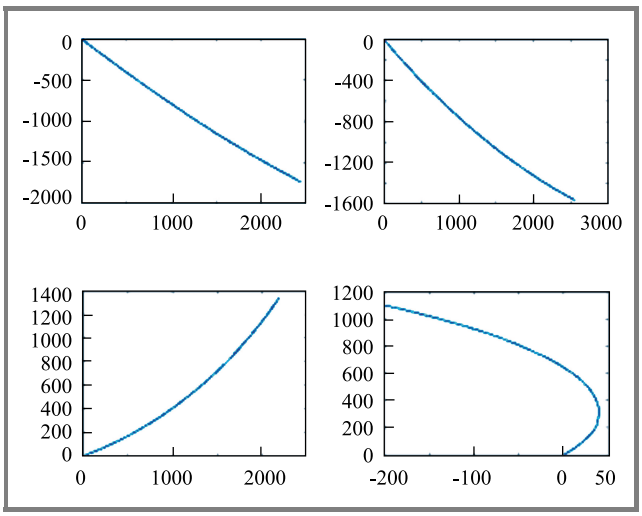


Fig. 9. First four pilots in (Re, Im) planes after phase correction with value $\chi = -0.00040$ rad/frame.

5. Conclusions

The proposed method of coherent averaging is very effective. As experimental results shown that acceptable results were obtained even if the original signal was degraded by another signal stronger by 30–40 dB.
It is possible to obtain further extension of this method to improve the calculation accuracy of drift χ. This method can be also optimised to minimise its time consumption, when is used as a component of standard digital signal processing (DSP) library.

References

[1] R. G. Lyons, *Understanding Digital Signal Processing*. New York: Prentice Hall, 1996.

[2] R. G. Lyons, *Wprowadzenie do cyfrowego przetwarzania sygnałów*. Warszawa: WKŁ, 1999 (in Polish).

[3] K. G. Beauchamp, *Przetwarzanie sygnałów metodami analogowymi i cyfrowymi*. Warszawa: WNT, 1978 (in Polish).

[4] W. Butryn, *Dźwięk cyfrowy*. Warszawa: WKŁ, 2002 (in Polish).
[5] J. Izydorczyk, *Teoria sygnałów*. Gliwice: Helion, 1999 (in Polish).

Piotr Z. Gajewski received the M.Sc. and D.Sc. degrees from Military University of Technology (MUT) Warsaw, Poland, in 1970, and 2001, respectively, both in telecommunication engineering. Since 1970 he has been working at Electronic Faculty of Military University of Technology (EF MUT) as a scientist and lecturer in communications systems (radios, cellular, microcellular), signal processing, adaptive techniques in communication and communications and information systems interoperability. He was an Associate Professor at Telecommunication System Institute of EF MUT from 1980 to 1990. From 1990 to 1993 he was Deputy Dean of EF MUT. Currently he is the Director of Telecommunication Institute of EF MUT. He is an author (co-author) of over 80 journal publications and conference papers as well as four monographs. He is a member of the IEEE Vehicular Technology and Communications Societies. He is also a founder member of the Polish Chapter of Armed Forces Communications and Electronics Association.

e-mail: pgajewski@wel.wat.edu.pl
Military University of Technology
Kaliskiego st 2
01-489 Warsaw, Poland

Jerzy Łopatka received the M.Sc. and Ph.D. degrees in communications from the Military University of Technology (MUT), Warsaw, Poland. At present he is a Head of Radiocommuniation Section in the Telecommunication Institute (MUT). His main research interests include digital signal processing in wireless systems.

e-mail: jlopatka@wel.wat.edu.pl
Military University of Technology
Kaliskiego st 2
01-489 Warsaw, Poland

Zbigniew Piotrowski was born in Gdynia, Poland, in 1973. He received the M.Sc. degree from Military University of Technology (MUT), Warsaw, Poland, in 1997 in telecommunication engineering. After one-year training period he works as an assistant in Telecommunication Institute of EF MUT. His main areas of interest are speech and audio processing and telecommunication systems engineering.

e-mail: zpiotrowski@wel.wat.edu.pl
Military University of Technology
Kaliskiego st 2
01-489 Warsaw, Poland

Regular paper

Comparison of traffic performance of QPSK and 16-QAM modulation techniques for OFDM system

Imdadul Islam and Siddique Hossain

Abstract— Orthogonal frequency division multiplexing (OFDM) provides better spectral efficiency than frequency division multiplexing (FDM), while maintaining orthogonal relation between carriers; hence traffic is better carried by OFDM than FDM within the same spectrum. This paper reveals a comparison of spectral efficiency, performance of communication system in context of bit error rate (BER) for the same information rate and peak to average power ratio (PAPR) of quadrature amplitude shift keying (QPSK) and 16-quadrature amplitude modulation (16-QAM) technique.

Keywords— OFDM, QPSK, 16-QAM, IFFT, frequency spectrum, PAPR, BER.

1. Introduction

Today major challenge in telecommunication is to convey as much information as possible through limited spectral width. Orthogonal frequency division multiplexing (OFDM) introduces the concept of allocating more traffic channels within limited bandwidth of physical channel. Here the available bandwidth is split into several narrow band channels for simultaneous transmission. In frequency division multiplexing (FDM) a guard band is provided between individual channels, which separates the spectrum of different channels, and enables a practical band pass filter to detect individual channel. But the situation is completely different in OFDM where spectrums of adjacent channels are overlapped which resembles adjacent channel interference, but interference is avoided by maintaining orthogonal relation between sub-carriers. First of all high speed serial data is converted to low speed parallel data, as shown in Fig. 1 based on [1, 2]. Therefore transmitted signal is a vector addition of orthogonal modulated carriers, makes large peak to average power ratio, therefore dynamic range of devices should be large enough, as summarized in [3–5].

Output of each parallel line is modulated; here two different types of modulation quadrature amplitude shift keying (QPSK) and 16-quadrature amplitude modulation (16-QAM) are selected for this paper, whose constellations are shown in Fig. 2. QPSK waves have constant peaked sinusoidal wave but phase angle is different for four different combinations of 2 bits. In 16-QAM both amplitude and phase of the wave varies according to 16 different combination of 4 bits. In this paper 7 parallel lines are used, hence 7 different carrier frequencies are used for simulation. Parallel waves are again converted to an instantaneous serial waves prior to transmission. This phenomenon resembles inverse first Fourier transform (IFFT) mentioned in [3, 6–8]. At receiving end signals are detected by coherent or envelope detection but this paper considers only coherent detection. In Section 2 complete analysis of transmitted signal in both time and frequency domain is done explicitly along with carrier waves (both before and after modulation) using constellation vectors of QPSK and 16-QAM. All the equations needed to detect signal at receiving end along with evaluation of peak to average power ratio (PAPR) [9, 10] are also summarized in this section. Section 3 deals with simulation of OFDM in additive white Gaussian noise (AWGN) environment to evaluate the performance of both modulation techniques in context of bit error rate (BER) and PAPR. Finally a comparison of both modulation techniques is given in a nutshell in Section 4 based on complete analysis previous sections.

2. Methodology

Typical FFT-based OFDM communication system is shown in Fig. 1. Modulator part of the figure will use only QPSK and 16-QAM technique whose constellation is shown in Fig. 2. In OFDM each sub-carrier is modulated independently with complex modulation symbol vector and added for simultaneous transmission; it is expressed like [6, 8]:

$$v(t) = \tilde{v}(t)\mathrm{e}^{\mathrm{j}2\pi f_c t}\,. \tag{1}$$

Complex envelope $\tilde{v}(t)$ of above equation is summarized succinctly in [1–3, 6] given by

$$\tilde{v}(t) = A_c \sum_{n=0}^{N-1} \omega_n \phi_n(t)\,; \quad 0 > t > T\,, \tag{2a}$$

where A_c is the peak carrier amplitude and ω_n is the N-element parallel vector.

For orthogonal relation the sub-carrier frequencies are related as

$$\phi_n = \mathrm{e}^{\mathrm{j}2\pi f_n t} \quad \text{and} \quad f_n = \frac{1}{T}\left(n - \frac{N-1}{2}\right). \tag{2b}$$

Figure 3a shows the OFDM signal, i.e., summation of sub-carriers prior to modulation and Fig. 3b depicts the same signal after QPSK modulation. Before modulation, the waves have the same starting and ending point since each

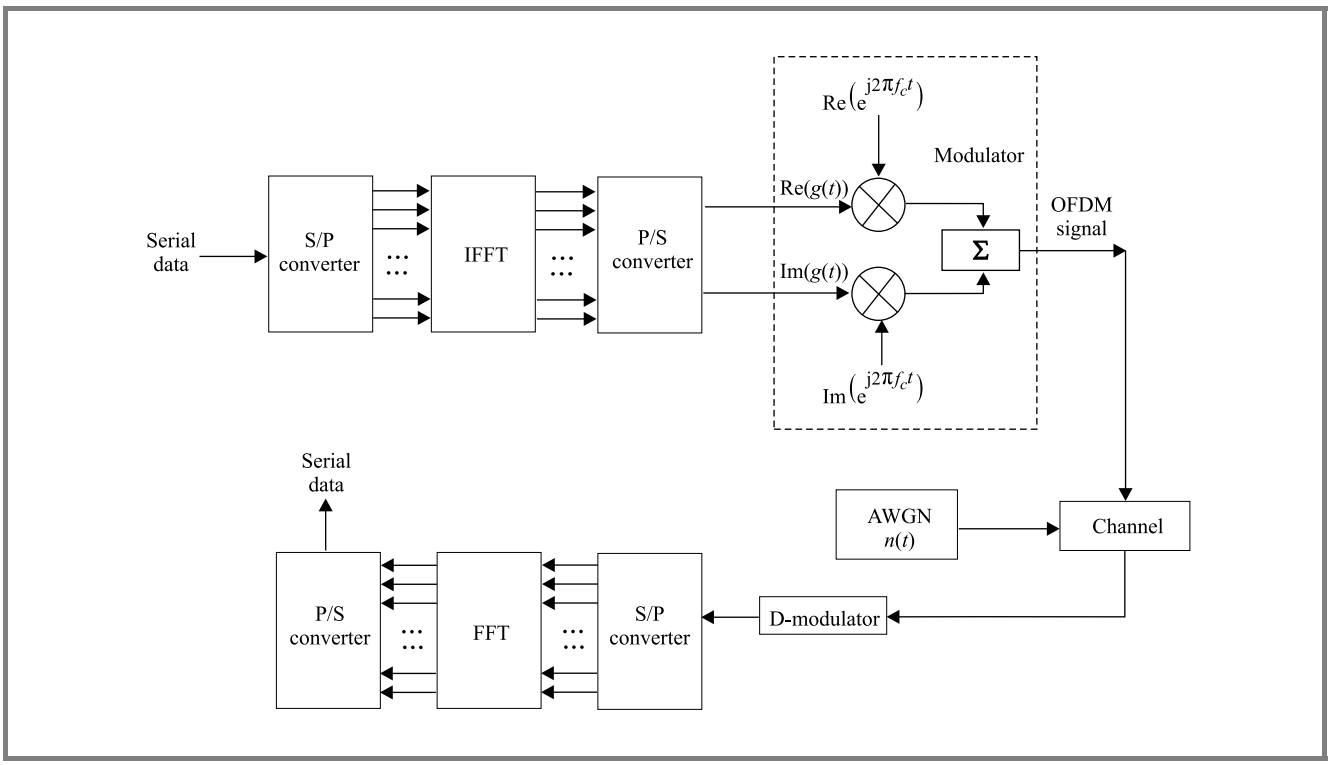


Fig. 1. OFDM communication system.

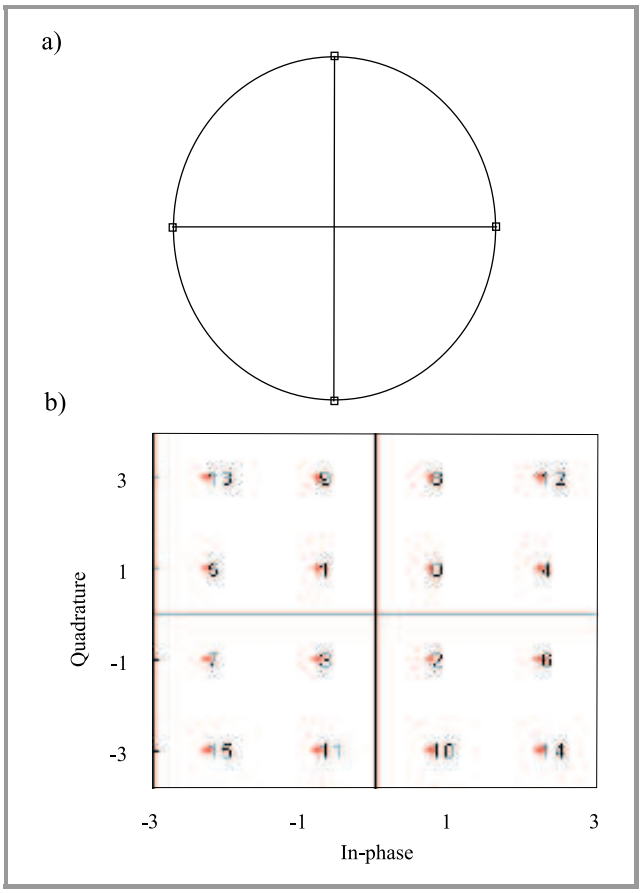


Fig. 2. Constellation vector (a) of QPSK and (b) of 16-QAM.

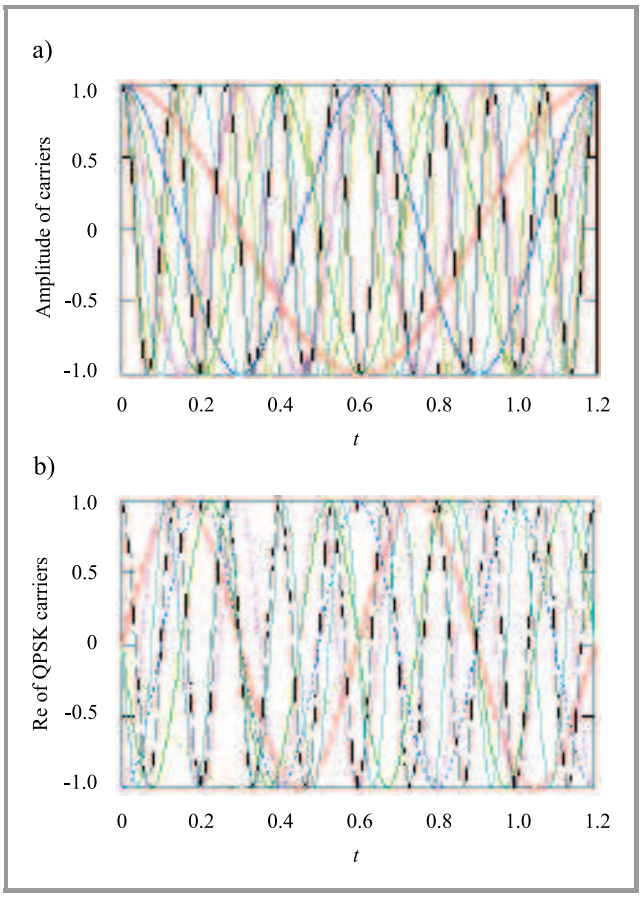


Fig. 3. Sub-carriers (a) befor and (b) after modulation.

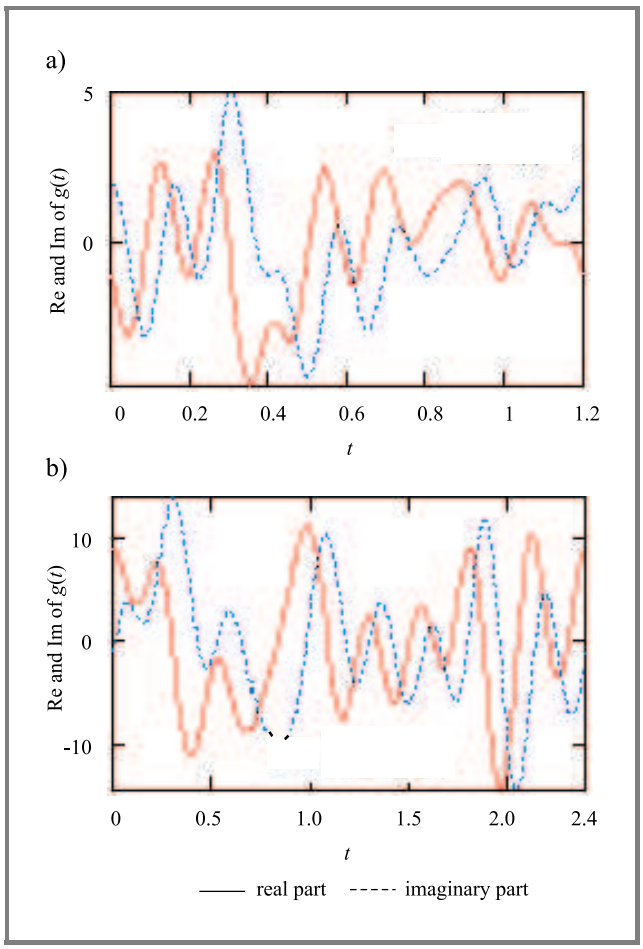


Fig. 4. Real and imaginary part of complex envelope of OFDM (a) of QPSK in time domain and (b) of 16-QAM in time domain.

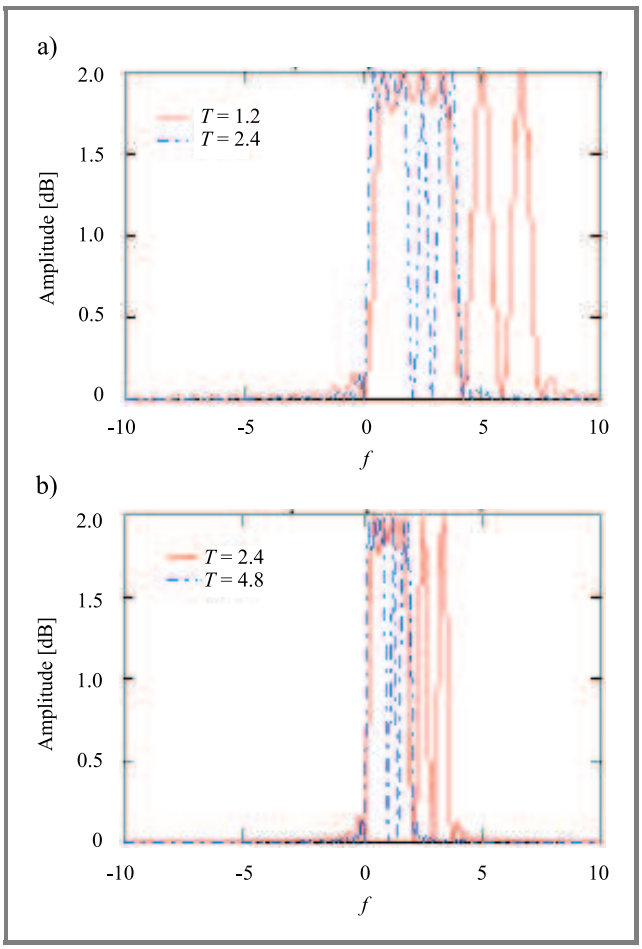


Fig. 5. Frequency spectrum of complex envelope (a) of QPSK and (b) of 16-QAM.

carrier has an integer number of cycles over a symbol period to maintain orthogonal relation but after modulation start and end points are shifted due to multiplication of constellation vectors. For Fig. 3b constellation vectors for QPSK and 16-QAM are taken as

$$W_{\text{QPSK}} = \begin{bmatrix} e^{j\pi/2} \\ e^{j3\pi/2} \\ e^{j\pi} \\ e^{j\pi/2} \\ e^{j\pi} \\ e^{j\pi/2} \\ e^{j0} \end{bmatrix} \qquad W_{\text{16-QAM}} = \begin{bmatrix} 1+j \\ 3+j \\ 3-3j \\ -1+j \\ -3-3j \\ 3-i \\ 3+3i \end{bmatrix}$$

Real and imaginary part of complex envelope of 7 simultaneously transmitted signal is shown in Fig. 4 for both 16-QAM and QPSK. Signals have very wide dynamic ranges for both cases.

Frequency spectrum of complex envelope [16] is given by

$$\Psi(f) = C \sum_{n=0}^{N-1} |\sin c(f - f_n)T|^2 . \tag{3}$$

Spectrum of QPSK and 16-QAM signals is depicted in Fig. 5 for a symbol period of $T = 1.2$ and 2.4 units for QPSK modulation and $T = 2.4$ and 4.8 units for that of 16-QAM. Symbol period of 16-QAM is taken twice compare to that of QPSK, since each modulation symbol of 16-QAM holds 4 bits but that of QPSK holds only two bits.

If there is N different users, i.e., N sub-carriers OFDM system, nth signal block [7, 8, 11] is represented as

$$S_n(t) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} S_{n,k}\, g_k(t - nT) . \tag{4a}$$

Entire continuous time signal:

$$S(t) = \frac{1}{\sqrt{N}} \sum_{n=0}^{\infty} \sum_{k=0}^{N-1} S_{n,k}\, g_k(t - nT) . \tag{4b}$$

Where the constellation vector $S_{n,k}$ of kth sub-carrier is recovered using cross correlation of following equation:

$$S_{n,k} = \frac{\sqrt{N}}{T_S} \left\langle S_n(t),\, \overline{g_k(t - nT)} \right\rangle , \tag{5}$$

where

$$\langle g_k,\, g_l \rangle = \int g_k(t)\, \overline{g_l(t)} dt .$$

At receiving end, the constellation vector becomes [4]:

$$R_{n,k} = \frac{\sqrt{N}}{T_S}\left\langle r_n(t),\, \overline{g_k(t-nT)}\right\rangle, \tag{6}$$

where $r_n(t) = S_n(t) + n(t)$; $n(t)$ is AWGN of environment. A maximum likelihood sequence estimator would have to choose one out of all possibly transmitted symbol sequence μ. The sequence estimator determines an estimated $<S_{n,k}>$ according to the following criterion:

$$\langle \hat{S}_{n,k} \rangle = \min \sum_k \left| R_{n,k} - H_{n,k} S_{n,k}(\mu) \right|^2, \tag{7}$$

where μ is the types of possible modulation symbols and $H_{n,k}$ is the transfer function of channel [12].
Finally peak to average power ratio is evaluated as

$$\text{PAPR} = \frac{\max\langle |s(t)|^2 \rangle}{\text{mean}\langle |s(t)|^2 \rangle}. \tag{8}$$

3. Simulation and results

A simulation work is done based on Eqs. (1)–(8) by the authors using MATLAB-6.5 in their own way to evaluate

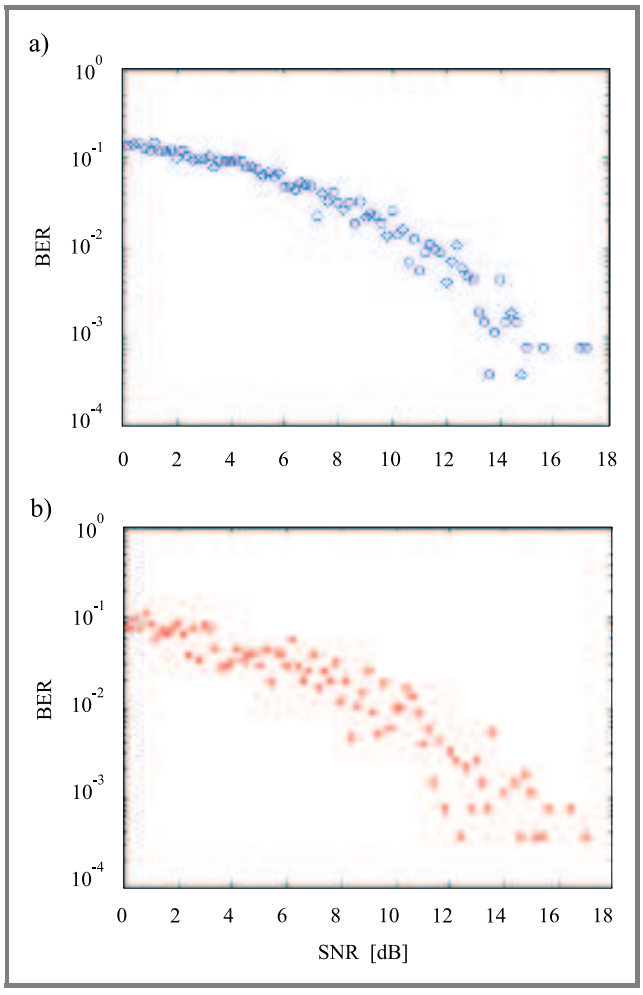


Fig. 6. Comparison of performance (a) of 16-QAM and (b) of QPSK under AWGM.

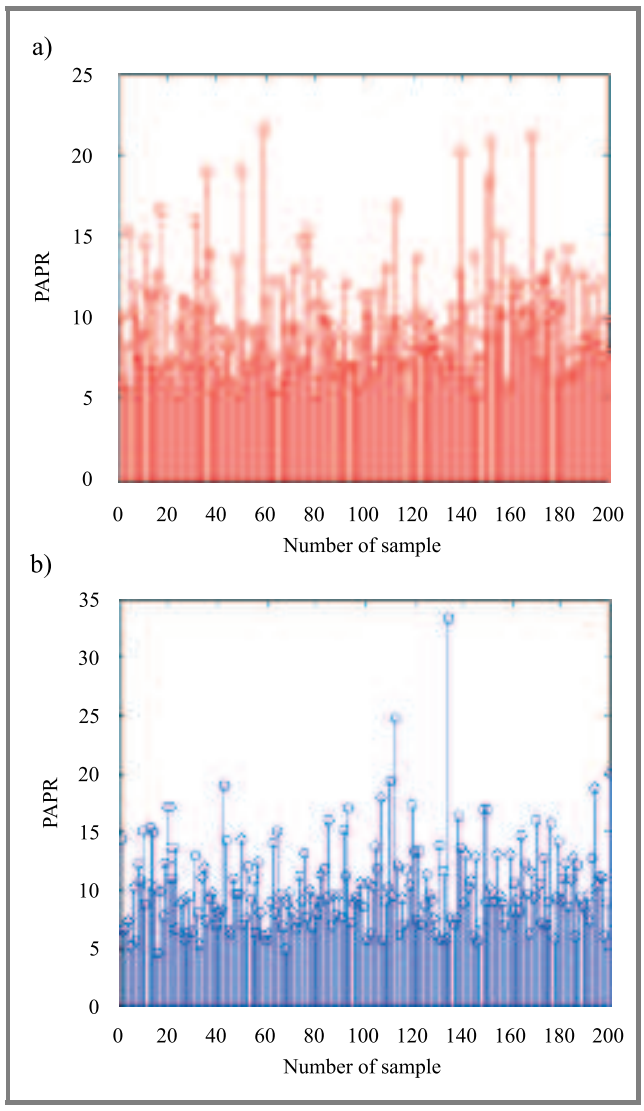


Fig. 7. Comparison of PAPR (a) of 16-QAM and (b) of QPSK.

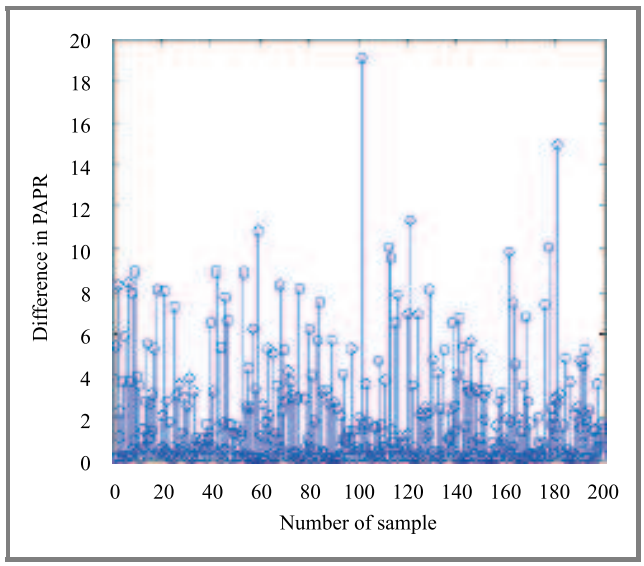


Fig. 8. Difference between PAPR of QPSK and 16-QAM for same throughput.

performance of QPSK and 16-QAM for OFDM in AWGM environment in context of spectral width, BER and PAPR shown in Figs. 5–7. Spectrum of narrower time slot becomes wider in frequency domain, visualized by solid lines of Fig. 5. Each symbol of QPSK convey 2 bits but that of 16-QAM is 4 bits/symbol therefore in time domain equivalent symbol period of 16-QAM is twice as long.

In this paper 10 000 random bits are generated to detect channel performance in AWGN environment. Rising cosine filter is used to emulate transmission medium and SNR is varied from 0 to 18 dB depicted in Fig. 6. One of the major problems in OFDM is the peak to average power ratio of un-coded signals. Here no coding technique is used to improve PAPR like [5, 13] since our aim is to compare performance of modulation technique in severe environment. Here PAPR is evaluated for 200 samples for both modulation techniques depicted in Fig. 7. Variation of PAPR lies between 5 to 15 units in Fig. 7 also verified in Fig. 8 where difference between PAPR of two modulation technique is measured shows the same difference. PAPR of QPSK and 16-QAM appear identical and it is really difficult to make command about improvement of PAPR but performance of both could be improved using coding technique summarized in [5, 13, 14].

4. Conclusion

It is obvious from Fig. 5 that spectral width of 16-QAM is narrower than that of QPSK for same information rate. Each symbol of QPSK conveys 2 bits but that of 16-QAM is 4 bits/symbol therefore in time domain equivalent symbol period of 16-QAM is twice as long. This phenomenon is verified from the simulation program. In context of BER, QPSK yields better performance than that of 16-QAM, shown in Fig. 6. Finally it could be concluded that BER performance of QPSK is better than that of 16-QAM at the expense of spectral width. Therefore 16-QAM can carry more traffic than QPSK at the expense of BER which is obvious in context of digital modulation technique hence analysis of the paper yield logical results in context of OFDM. PAPR solely depends on coding technique not on modulation technique, which is also verified from the simulation.

References

[1] B. J. Periris, R. M. A. P. Rajatheva, and V. Tralli, "Performance of self ICI cancellation based OFDM systems with carrier frequency offset estimation", in *3rd Int. Symp. Wirel. Pers. Multimed. Commun*., Bangkok, Thailand, 2000, pp. 680–683.

[2] P. H. Moose, "A technique for orthogonal frequency division multiplexing offset correction", *IEEE Trans. Commun.*, vol. 42, no. 10, pp. 2908–2914, 1994.

[3] A. D. S. Jayalath and C. Tallambura, "Reducing the peak to average power ratio of an OFDM signal by interleaving" in *3rd Int. Symp. Wirel. Pers. Multimed. Commun*., Bangkok, Thailand, 2000, pp. 698–703.

[4] T. May, H. Rohling, and V. Engels, "Performance analysis of Viterbi decoding for 64-DAPSK and 64-QAM modulated OFDM signals", *IEEE Trans. Commun.*, vol. 46, pp. 182–190, 1998.

[5] D. H. Guo and Ch.-Y. Hsu, "Minimization of the PARP of OFDM by linear systematic block coding", in *Asia-Pacific Conf. Commun. APCC*, Malaysia, 2003, pp. 129–133.

[6] L. W. Couch, *Digital and Analog Communication Systems*. 6th ed. Singapore: Pearson Education, 2002.

[7] D. A. Wegandt, Z. Wu, and C. R. Nassar, "High throughput, high performance OFDM via pseudo-orthogonal carrier interformely spreading codes", *IEEE Trans. Commun*., vol. 51, no. 7, pp. 1123–1134, 2003.

[8] A. F. Molish, *Wideband Wireless Digital Communications*. Singapore: Pearson Education, 2003, Chapt. 18 and 20.

[9] H. Zhang and Y. Li, "Optimum frequency domain partial response encoding in OFDM system", *IEEE Trans. Commun*., vol. 51, p. 1068, 2003.

[10] X. Cai, S. Zhou, and G. B. Giannakis, "Group-orthogonal multicarrier CDMA", *IEEE Trans. Commun*., 2003 (to appear).

[11] P. Banelli and S. Cacopardi, "Theoretical anlysis and performance of ODDM signal in nonlinear AWGN channels", *IEEE Trans. Commun*., vol. 48, no. 3, pp. 430–441, 2000.

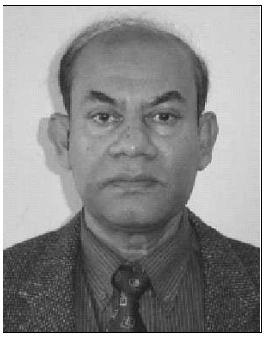
[12] M. A. Saeed, B. M. Ali, and M. H. Habaebi, "Performance evaluation of OFDM scenes over multipath fading channel", in *Asia-Pacific Conf. Commun. APCC*, Malaysia, 2003, pp. 415–419.

[13] D. H. Guo and C.-Y. Hsu, " Systematic reducing the PARP of OFDM by cyclic coding", in *Asia-Pacific Conf. Commun. APCC*, Malaysia, 2003, pp. 133–137.

[14] S. K. Yusof, N. Fisal, and T. S. Yin, "Reducing PAPR of OFDM signals using partial transmit sequences", in *Asia-Pacific Conf. Commun. APCC*, Malaysia, 2003, pp. 411–414.

Imdadul Islam has received his B.Sc. in electrical and electronic engineering from Bangladesh University of Engineering and Technology, Dhaka, in 1993 and the M.Sc. from the same institute in 1998. Now he is perusing Ph.D. at the department of EEE, BUET, Dhaka, in the field of teletraffic engineering. He worked as an Assistant Engineer in Sheba Telecom (Pvt.) Ltd. (a joint venture company between Bangladesh and Malaysia, for mobile cellular and WLL), from Sept'94 to July'96. He has very good field experiences in installation of radio base station and switching center for WLL. He is now working as an Associate Professor, at the Department of Electronics and Computer Science, Jahangirnagar University, Savar, Dhaka, Bangladesh. His research field is network traffic, OFDMA, WCDMA and array antenna systems. He has published more than 20 papers in national and international journal and conference proceedings.

e-mail: imdad22000@yahoo.com
Department of Electronic and Computer Science
Jahangirnagar University, Savar, Dhaka
Bangladesh



Siddique Hossain is the most senior Professor of Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology, Dhaka. He has more than 32 years of teaching, research and administrative experience both at home and abroad. Worked as the Head, EEE Dept., Head, CSE Dept., Dean, EEE Faculty and Director, BUET Computer Centre. He worked as a visiting faculty member of more than ten universities. He is a Senior Member of IEEE and was engaged as Chief of IEEE in Bangladesh in 1997–1998. Dr. Siddique Hossain is interested in the field of computers and communication engineering, 3G mobile communication, WCDMA for UMTS, etc. He has more than 30 publications in national, international journals and conference proceedings.

e-mail: sdq@eee.buet.ac.bd
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka-1000, Bangladesh

Regular paper

A newly developed random walk model for PCS network

Imdadul Islam and Siddique Hossain

Abstract— Different types of random walk models are prevalent in mobile cellular network for analysis of roaming and handover, being considered as important parameters of traffic measurement and location updating of such network. This paper proposes a new random walk model of hexagonal cell cluster, exclusively developed by the authors and a comparison is made with two existing models. The proposed model shows better performance in context of number of probability states compared to existing models.

Keywords— cell cluster, random walk, subarea n, state transition, probability matrix, expected number of steps.

1. Introduction

In a mobile cellular network, the service area is represented as an array of hexagonal cells in a continuous fashion. In any cell two types of offered traffic take place: one is new call arrival and the other is handover arrival. The latter one solely depends on mobility of users hence mobility is an important parameter for measurement of quality of service (QoS) of a network. In teletraffic engineering mobility is measured as probability $P_{i,j}$, i.e., probability of an mobile station (MS) to make transition from cell i to cell j. User's mobility is random and usually does not follow any particular probability distribution function, but could be analyzed based on random walk model summarized in [1, 2] where model in [2] shows better performance than that of [1] in context of number of states. This paper proposes a new random walk model considering that any user makes transition from its current cell to any neighboring cells with equal probability, i.e., 1/6 like existing model but our aim is to reduce number of states of state transition diagram hence gives less process time in detection of state of an MS. In Section 2 previous two models are depicted in a nutshell and referred to as "model 1" and "model 2", but the proposed model is summarized in detail and designated as "proposed model".

2. Methodology

Any MS can make transition from its present cell to any one of surrounding cells with equal probability of 1/6 for hexagonal cell structure shown in Fig. 1. Each cell in a mobile cellular network has its own identification number hence in random walk model each cell has to be designated by a number based on certain criteria. Two dimensional cell identification technique is used in both previous and proposed model based on [1–3]. Probability of transition of mobile stations from one cell to another is depicted by both state transition and probability matrix. Finally, the expected number of steps to make transition from each cell to most peripheral cell is detected to get idea of mobility at a glimpse.

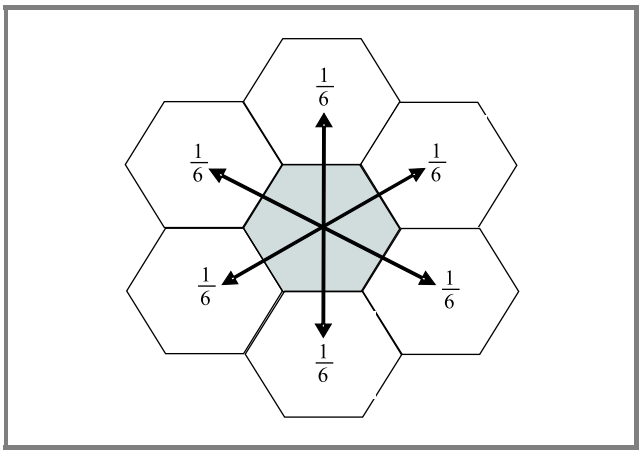


Fig. 1. State transition of hexagonal cell cluster/pattern.

2.1. Model 1

In this model [1] the entire cell pattern is divided into different subareas with respect to the cell at the center of the cluster labeled (0, 0) and this cell at the centre called subarea 0. All the cells surrounding (0, 0) are marked as (1, 0) and designated as subarea 1. All the cells surrounded by subarea 1 is called subarea 2 and cells are marked as (2, 0) and (2, 1) in an alternate fashion. Similarly cells of subarea 3 are marked as (3, 0), (3, 1) and (3, 2) and so on, is given in Fig. 2. In recursive form the cells surrounded by subarea x are called subarea $x+1$. This type of two dimensional model was first proposed in [3] and modified by the same author in [1]. In this model, the cell cluster/pattern is symmetrical in six wedges marked in alternate shade of white and dark. Cells in a single wedge suffice for analysis since they are symmetrical.

Here the number of distinguished cells increases by one with each increment of level, hence number of states for n subarea cell pattern would be

$$1+2+3+4+\ldots+(n-1)+n=n(n+1)/2. \quad (1)$$

This model is summarized in [1] hence state transition diagram and probability matrix is avoided since authors are

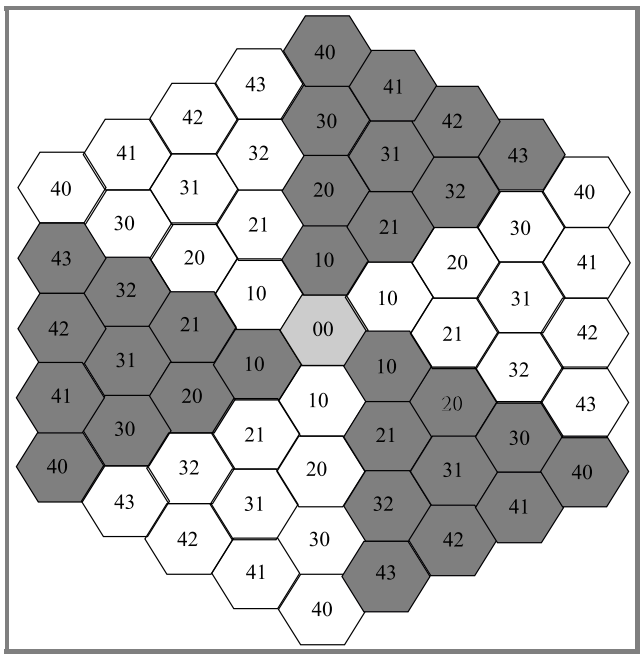


Fig. 2. Cell cluster of 4 subareas defined in model 1.

only interested in the proposed model developed by themselves in details.

2.2. Model 2

Advanced form of previous model is proposed in [2], where marking of cells is a little bit different than that of [1] given in Fig. 3. Like in model 1, the cell at the center of the cluster is called subarea 0 and the cells surrounded by that cell are called subarea 1, and are marked as (1, 0). Cells

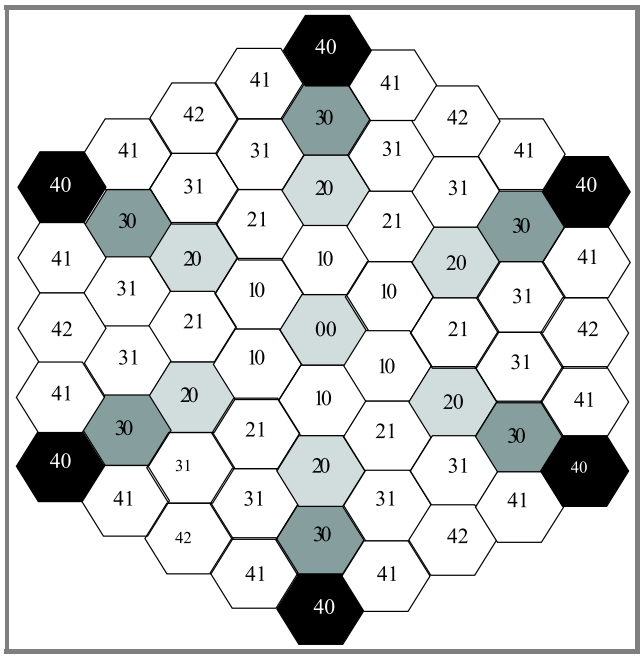


Fig. 3. Cell cluster of 4 subareas defined in model 2.

of subarea 2 are marked like [1] but that of subarea 3 are marked as (3, 0), (3, 1) and (3, 1). Subarea 4 is marked as (4, 0), (4, 1), (4, 2) and (4, 1) and so on. This model also shows the same symmetrical characteristics like the previous one. Here number of states for a cell cluster of n subarea is evaluated as

$$1+(1+1)+(2+2)+(3+3)+\ldots+(k+k) = \frac{(n^2+2n+4)}{4}; \quad \text{where } n=2k, \text{ i.e., } n \text{ is even,} \quad (2)$$

$$1+(1+1)+(2+2)+(3+3)+\ldots+(k+k)+k+1 = \frac{(n^2+2n+5)}{4}; \quad \text{where } n=2k+1, \text{ i.e., } n \text{ is odd.} \quad (3)$$

State transition diagram and probability matrix of this model are also excluded for the same reason as mentioned in previous section.

2.3. Proposed model

This is the model proposed by authors where the number of states is reduced compared to [1, 2] at the expense of complexity of determination of probability of transition from one state to another. Here the cell at the center is marked as (0, 0) and called subarea 0, cells at subarea 1 are marked alternately (1, 0) and (1, 1), cells at subarea 2

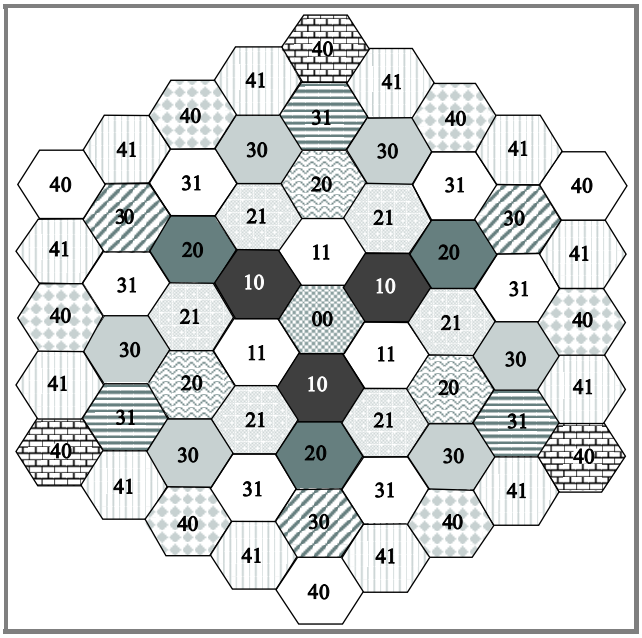


Fig. 4. Cell cluster of 4 subareas defined in the proposed model.

are alternately marked as (2, 0), (2, 1), that of subarea 3 as (3, 0) and (3, 1) and so on. The symmetric cells are marked with same brightness as shown in Fig. 4 but its symmetrical characteristics are different from both of previous two. Here number of states for a cell cluster of n subarea is

$$1+\big(2+2+2+\ldots+(n-1)\text{th } term\big)+1 = 2\big(1+1+1+\ldots+n\text{th } term\big) = 2n. \quad (4)$$

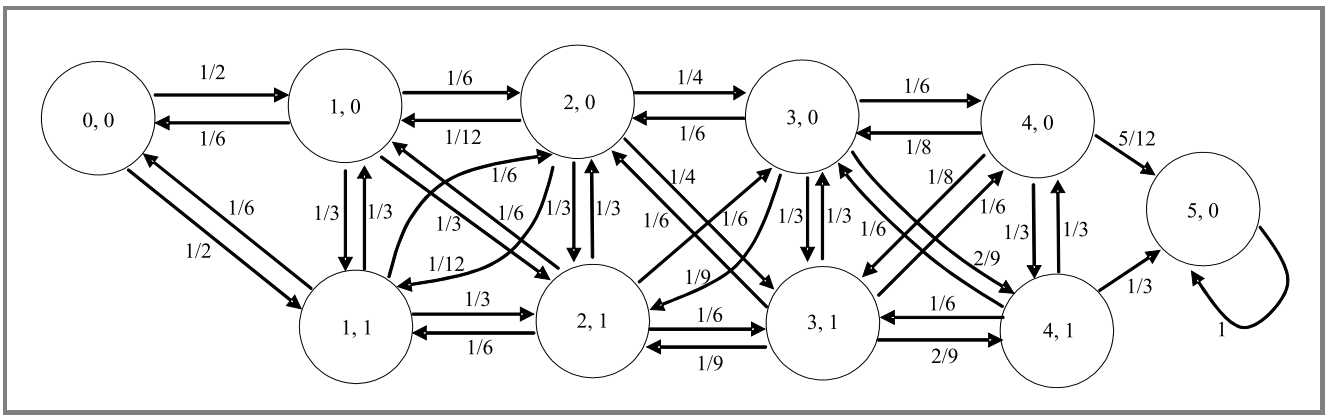


Fig. 5. State transition for 4 subareas of proposed model.

Figure 5 shows the state transition diagram for a cell pattern/cluster of 4 subareas, which resembles a finite stochastic process. Probability of transition from any arbitrary state (p, q) to $(p \pm 1, q \pm 1)$ doesn't remain constant in generalized form like [1, 2]. Probability of transition between very few states is constant, but in most of the cases probability is evaluated based on symmetrical relation among subarea i, subarea $(i+1)$ and number of cells s_i or $s_{(i+1)}$ of subarea i or subarea $(i+1)$. Probability of transition between any two states of cell cluster of n subarea is derived as

1. $P_{00,1i} = \frac{1}{2}$ for $i = 0$ and 1.
2. $P_{i0,(i+1)0} = \begin{cases} \frac{1}{6} & \text{for } 1 \le i \le n-1, \quad i \text{ is odd} \\ \frac{1}{4} & \text{for } 2 \le i \le n-2, \quad i \text{ is even.} \end{cases}$
3. $P_{i0,(i+1)0} = \begin{cases} \frac{1}{6} & \text{for } 1 \le i \le n-1, \\ & \text{and } i \text{ is odd} \\ \left[0 \times 3 + \left(\frac{s_i}{2} - 3\right)\frac{1}{6}\right]\frac{2}{S_i} & \text{for } 2 \le i \le n\text{–}2, \\ & \text{and } i \text{ is even.} \end{cases}$
4. $P_{i0,(i+1)1} = \left[3 \times \frac{1}{3} + \frac{1}{6}\left(\frac{s_i}{2} - 3\right)\right]\frac{2}{s_i}$ for $1 \le i \le n-2$.
5. $P_{i0,i1} = \frac{1}{3} = P_{i1,i0}$ for $1 \le i \le n-1$.
6. $P_{i1,(i-1)0} = \frac{1}{6}$ for $1 \le i \le n-1$.
7. $P_{i1,(i+1)1} = \frac{1}{6}$ and $P_{(i+1)1,i1} = \left[0 \times 3 + \frac{1}{6}\left(\frac{s_{i+1}}{2} - 3\right)\right]\frac{2}{s_{i+1}}$ for $1 \le i \le n-1$.
8. $P_{i1,(i+1)0} = \frac{1}{6}$ and $P_{(i+1)0,i1} = \left[0 \times 3 + \frac{1}{6}\left(\frac{s_{i+1}}{2} - 3\right)\right]\frac{2}{s_{i+1}}$ for $1 \le i \le n-1$.
9. $P_{(n-1)0,n0} = \frac{5}{12}$ and $P_{(n-1)1,n0} = \frac{1}{3}$.
10. $P_{n0} = 1$.

Probability matrix P of state transition in generalized form is like

$$P = \begin{bmatrix} P_{00,00} & P_{00,10} & P_{00,20} & \dots & P_{00,n0} \\ P_{10,00} & P_{10,10} & P_{10,11} & \dots & P_{10,n0} \\ P_{11,00} & P_{11,10} & P_{11,11} & \dots & P_{11,n0} \\ \dots\dots & \dots\dots & \dots\dots & \dots & \dots\dots \\ \dots\dots & \dots\dots & \dots\dots & \dots & \dots\dots \\ P_{n-11,00} & P_{n-11,10} & P_{n-11,11} & \dots & P_{n-11,n0} \\ P_{n0,00} & P_{n0,10} & P_{n0,11} & \dots & P_{n0,n0} \end{bmatrix}. \quad (5)$$

Element of the matrix, $P_{ij,pq}$ is the probability of transition between cell (i, j) to cell (p, q). Probability transition matrix of the proposed model of 4-subarea cell cluster is derived as

$$P = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{6} & 0 & \frac{1}{3} & \frac{1}{6} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{6} & \frac{1}{3} & 0 & \frac{1}{6} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{12} & \frac{1}{12} & 0 & \frac{1}{3} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & 0 & \frac{1}{6} & \frac{1}{6} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{6} & \frac{1}{9} & 0 & \frac{1}{3} & \frac{1}{6} & \frac{2}{9} & 0 \\ 0 & 0 & 0 & \frac{1}{6} & \frac{1}{9} & \frac{1}{3} & 0 & \frac{1}{6} & \frac{2}{9} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{8} & \frac{1}{8} & 0 & \frac{1}{3} & \frac{5}{12} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (6)$$

Now a random walk moves from state (a, b) to (a', b') with s steps summarized in [1, 2] as

$$P_{sab,a'b'} = \begin{cases} P_{ab,a'b'} & \text{for } s = 1 \\ P^{(s)}_{ab,a'b'} - P^{(s-1)}_{ab,a'b'} & \text{for } s > 1. \end{cases} \quad (7)$$

Where $P^{(s)}$ is evaluated by recurrence formulae

$$P^{(s)} = \begin{cases} P & \text{for} \quad s = 1 \\ P \times P^{(s-1)} & \text{for} \quad s > 1\,. \end{cases} \tag{8}$$

Expected number of steps for a MS to leave subarea n is evaluated as

$$N(a,b) = \sum_{s=1}^{\alpha} s \times P_{sab,n0}\,, \tag{9}$$

where $N(a,b)$ is equivalent to $L(x,y)$ of [2]. Taking maximum value of $s = 300$, the values of $N(a,b)$ are derived from different starting states based on [4, 5] shown in Table 1.

Table 1
The values of $N(a,b)$ derived from different starting states

$N(a,b)$	$P_{00,50}$	$P_{10,50}$	$P_{11,50}$	$P_{20,50}$	$P_{21,50}$	$P_{30,50}$	$P_{31,50}$	$P_{40,50}$	$P_{41,50}$
$N(a,b)$	19.5	18.5	18.5	15.64	16.44	12.16	12.16	6.44	7.2

It is obvious that only $N(1,0) = N(1,1)$ and $N(3,0) = N(3,1)$, i.e., very few states show symmetry compare to [1, 2].

3. Conclusion

In model 1 [1] the number of states for n-subarea cell cluster is $n(n+1)/2$, greater than that of model proposed in [2]. The model 2 [2] of n-subarea cell pattern has $(n^2+2n+4)/4$ states when n is even and $(n^2+2n+5)/4$ states when n is odd. It is obvious that $n(n+1)/2 > (n^2+2n+4)/4$ or $(n^2+2n+5)/4$, i.e., model 2 is better than model 1 in context of number of states. In proposed model, the number of states is $2n < (n^2+2n+4)/4$ or $(n^2+2n+5)/4$ for $n > 5$. Therefore the proposed model yields smaller number of states for a network of 6 subareas or more, i.e., performance of our model is better than that of [1, 2] for a large mobile cellular network. Any network planner can use our model quite comfortably since the process time to estimate any probability of state of any MS would be smaller in comparison to any one of the existing models.

References

[1] I. F. Akyildiz, Y.-B. Lin, W.-R. Lai, and R.-J. Chen, "A new random walk model for PCS networks", *IEEE J. Selec. Area Commun.*, vol. 18, pp. 1254–1260, 2000.

[2] G. Xue, "An improved random walk model for PCS networks", *IEEE Trans. Commun.*, vol. 50, no. 8, pp. 1224–1226, 2002.

[3] I. F. Akyildiz, J. S. M. Ho, and Y.-B. Lin, "Movement-based location update and selective paging for PCS network", *IEEE/ACM Trans. Netw.*, vol. 4, no. 4, pp. 638–639, 1996.

[4] T. Veerarajan, *Probability, Statistics and Random Proces*. 2nd ed. McGraw-Hill, 2003.

[5] S. M. Ross, *Introduction to Probability Models*. 7th ed. Harcourt India, 2001.

Imdadul Islam – for biography, see this issue, p. 151.

Siddique Hossain – for biography, see this issue, p. 152.

INFORMATION FOR AUTHORS

The *Journal of Telecommunications and Information Technology* is published quarterly. It comprises original contributions, both regular papers and letters, dealing with a broad range of topics related to telecommunications and information technology. Items included in the journal report primary and/or experimental research results, which advance the base of scientific and technological knowledge about telecommunications and information technology.

The *Journal* is dedicated to publishing research results which advance the level of current research or add to the understanding of problems related to modulation and signal design, wireless communications, optical communications and photonic systems, speech devices, image and signal processing, transmission systems, network architecture, coding and communication theory, as well as information technology. Suitable research-related manuscripts should hold the potential to advance the technological base of telecommunications and information technology. Tutorial and review papers are published by invitation only.

Papers published by invitation and regular papers should contain up to 15 and 8 printed pages respectively (one printed page corresponds approximately to 3 double-space pages of manuscript, where one page contains approximately 2000 characters).

Manuscript: An original and two copies of the manuscript must be submitted, each completed with all illustrations and tables attached at the end of the papers. Tables and figures have to be numbered consecutively with Arabic numerals. The manuscript must include an abstract limited to approximately 100 words. The abstract should contain four points: statement of the problem, assumptions and methodology, results and conclusion, or discussion, of the importance of the results. The manuscript should be double-spaced on only one side of each A4 sheet (210 × 297 mm). Computer notation such as Fortran, Matlab, Mathematica etc., for formulae, indices, etc., is not acceptable and will result in automatic rejection of the manuscript. The style of references, abbreviations, etc., should follow the standard IEEE format.

References should be marked in the text by Arabic numerals in square brackets and listed at the end of the paper in order of their appearance in the text, including exclusively publications cited inside. **The reference entry** (correctly punctuated according to the following rules and examples) **has to contain**.

From journals and other serial publications: initial(s) and second name(s) of the author(s), full title of publication (transliterated into Latin characters in case it is in Russian, possibly preceded by the title in Russian characters), appropriately abbreviated title of periodical, volume number, first and last page number, year. E.g.:

[1] Y. Namihira, "Relationship between nonlinear effective area and modefield diameter for dispersion shifted fibres", *Electron. Lett.*, vol. 30, no. 3, pp. 262–264, 1994.

From non-periodical, collective publications: as above, but after title – the name(s) of editor(s), title of volume and/or edition number, publisher(s) name(s) and place of edition, inclusive pages of article, year. E.g.:

[2] S. Demri, E. Orłowska, "Informational representability: Abstract models versus concrete models" in *Fuzzy Sets, Logics and Reasoning about Knowledge*, D. Dubois and H. Prade, Eds. Dordrecht: Kluwer, 1999, pp. 301–314.

From books: initial(s) and name(s) of the author(s), place of edition, title, publisher(s), year. E.g.:

[3] C. Kittel, *Introduction to Solid State Physics*. New York: Wiley, 1986.

Figure captions should be started on separate sheet of papers and must be double-spaced.

Illustration: Original illustrations should be submitted. All line drawings should be prepared on white drawing paper in black India ink. Drawings in Corel Draw and Postscript formats are preferred. Colour illustrations are accepted only in exceptional circumstances. Lettering should be large enough to be readily legible when drawing is reduced to two- or one-column width – as much as 4:1 reduction from the original. Photographs should be used sparingly. All photographs must be gloss prints. All materials, including drawings and photographs, should be no larger than 175 × 260 mm.

Page number: Number all pages, including tables and illustrations (which should be grouped at the end), in a single series, with no omitted numbers.

Electronic form: A floppy disk together with the hard copy of the manuscript should be submitted. It is important to ensure that the diskette version and the printed version are identical. The diskette should be labelled with the following information: a) the operating system and word-processing software used, b) in case of UNIX media, the method of extraction (i.e. tar) applied, c) file name(s) related to manuscript. The diskette should be properly packed in order to avoid possible damage during transit.

Among various acceptable word processor formats, TEX and LATEX are preferable. The *Journal's* style file is available to authors.

Galley proofs: Proofs should be returned by authors as soon as possible. In other cases, the article will be proof-read against manuscript by the editor and printed without the author's corrections. Remarks to the errata should be provided within two weeks after receiving the offprints.

The copy of the "Journal" shall be provided to each author of papers.

Copyright: Manuscript submitted to this journal may not have been published and will not be simultaneously submitted or published elsewhere. Submitting a manuscript, the authors agree to automatically transfer the copyright for their article to the publisher if and when the article is accepted for publication. The copyright comprises the exclusive rights to reproduce and distribute the article, including reprints and also all translation rights. No part of the present journal may be reproduced in any form nor transmitted or translated into a machine language without permission in written form from the publisher.

Biographies and photographs of authors are printed with each paper. Send a brief professional biography not exceeding 100 words and a gloss photo of each author with the manuscript.

(*continued from outside back cover*)

(continued on inside back cover)

National Institute
of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

Editorial Office

tel. +48(22) 512 81 83
tel./fax: +48(22) 512 84 00
e-mail: redakcja@itl.waw.pl
http://www.itl.waw.pl/jtit