

A Comparative Study of Non-melt Laser Spike Annealing and Flash Lamp Annealing in Terms of Transistor Performance and Pattern Effects on SOI-CMOSFETs for the 32 nm Node and Below

R. Illgen, S. Flachowsky, T. Herrmann, T. Feudel*, D. Thron*, B. Bayha*, W. Klix, M. Horstmann* and R. Stenzel

Department of Electrical Engineering, University of Applied Sciences Dresden
Friedrich-List-Platz 1, D-01069 Dresden, Germany
E-mail: ralf.illgen@et.htw-dresden.de

*AMD Saxony LLC & Co. KG, Wilschdorfer Landstrasse 101, D-01109 Dresden, Germany

Due to the continuous CMOS transistor scaling requirements, millisecond annealing has been introduced in 45 nm CMOS technology to enhance dopant activation with minimal dopant diffusion. This paper considers two different ultra fast annealing technologies as alternative to the conventional rapid thermal annealing strategy for the 32 nm node. We compared a long wavelength non-melt laser spike annealing and a flash lamp annealing in terms of CMOSFET device performance. We also investigated possible temperature variations induced by shallow trench isolation density variations of these two annealing techniques by means of electrical parameters. The comparison was made without the introduction of an absorbent layer to take into account the different absorption mechanism between laser spike annealing and flash lamp annealing. The results show that both approaches despite their different annealing techniques are full comparable in terms of device performance without any concerns in pattern effects at least for SOI-CMOSFETs and therefore equal useable for the 32 nm node and beyond.

Introduction

During the last few years, millisecond annealing (MSA) technologies like non-melt laser spike annealing (LSA) [1] or flash lamp annealing (FLA) [2] have been widely investigated in order to achieve high performance CMOSFET devices. In the 45 nm node these new annealing strategies are mainly applied in combination with a conventional rapid thermal

annealing (RTA) [3] [4]. In order to reach the 32 nm node requirements and beyond, it is necessary to replace the conventional RTA by MSA techniques as dopant activation technology to allow further transistor scaling. But there is a concern about performance variation between these both annealing schemes for different structures [5]. To solve this problem an absorbent layer above the wafer could be added to the FLA process that cancels out the light absorption variation [6]. But adding and removing this additional cap layer increases process complexity and costs and could affect the yield adversely.

In this paper, we compare the impact of LSA and FLA on performance of a diffusionless CMOSFET without the introduction of an absorbent layer. Furthermore, we investigate the dependency of these two annealing schemes on the shallow trench isolation (STI) layout density on a specific test structure.

Experimental

N- and p-MOSFET were fabricated on SOI wafers using SiON as gate dielectric and polysilicon as gate material. Before gate patterning, the poly was doped for both n- and p-MOSFET. To avoid poly depletion issues, a gate annealing was added before SDE definition. After halo and extension implants, the spacer was formed followed by deep source-drain implants. To activate the dopants in the SDE and deep source-drain area, only LSA or FLA was carried out before silicidation. All other process parameters were chosen to be identical.

The light source of LSA is a CO₂ laser at a wavelength of 10.6 μm whereas the emission spectrum distribution of FLA is white light and the intensity is high in the broad visible spectral region, particularly at wavelength in the range from 300 nm to 600 nm. The radiation absorption for these two annealing techniques is therefore extremely different. The dominant mechanism for FLA is interband absorption because this mechanism occurs for wavelengths shorter than 1.1 μm (Fig. 1). For the mid-far infrared light source that LSA uses, the principal absorption mechanism is by free carrier absorption. In this case the density of free carriers in a semiconductor material can be easily changed by heating the entire wafer.

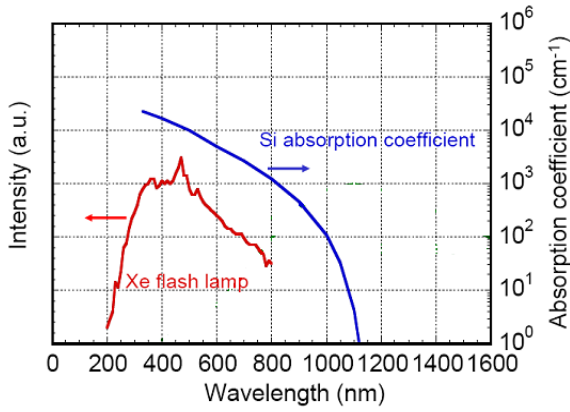


Fig. 1: Emission spectrum of Xenon flash lamp with optical absorption spectrum of silicon [2].

Due to the use of almost white visible light as in the case of FLA, constructive and destructive interference of wavelength by each device pattern can lead to microscopic emissivity non-uniformities which cause microscopic temperature fluctuation. These temperature variations result in variations in dopant activation from pattern to pattern. In the case of LSA, diffraction effects are minimized due to the use of a wavelength that is much longer than the film stack thicknesses. To minimize reflectivity deviations of patterned wafers the LSA equipment utilizes a p-polarized laser beam incident at Brewster's angle [7].

Microscopic emissivity non-uniformity within a patterned wafer should cause microscopic temperature fluctuation in annealing techniques by light irradiation in principle. Annealing time in MSA is extremely short and

lateral thermal diffusion length in MSA is about 100 μm [8]. Then emissivity non-uniformity within sub-100 μm-scaled area causes microscopic temperature distribution. To study possible pattern-related temperature non-uniformities between LSA and FLA within sub-100 μm-scaled areas we designed a test structure where the STI density, given by

$$\text{STI density} = \frac{L_{\text{STI}}}{L_{\text{STI}} + L_{\text{Active}}}$$

was changed to cover the variation range that can be observed on a chip. The size of the silicon islands within the test structure was fixed to an area of 0.320 μm x 3.0 μm. The STI pitch was varied in x- and y-direction from 0.1 μm to 2 μm which leads to an absolute STI density variation of around 0.2 to 0.8 (Fig. 2). The variation in the STI pattern density modified the absorptivity of the incident light and therefore the temperature. This structure includes a transistor in the middle of the configuration to measure possible temperature variations on the basis of electrical parameters.

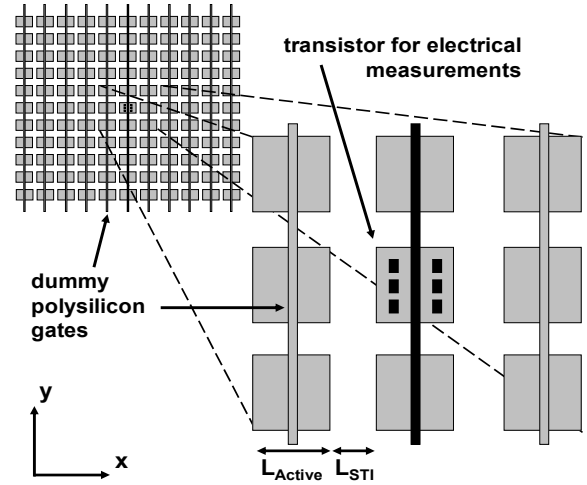


Fig. 2: Test structure for pattern effect evaluation; STI pitch was varied in x- and y-direction from 0.1 μm to 2 μm at a fixed active area size of 0.320 μm x 3.0 μm.

Results and Discussion

MOSFET characteristics

The comparison of $I_{D,\text{sat}}-I_{D,\text{off}}$ characteristics of n- and p-MOSFETs annealed only by LSA or FLA is shown in Fig. 3 and Fig. 4, respectively. In contrast to [5] there are no performance differences between LSA and

FLA for both n- and p-MOSFET devices even without the use of an absorbent layer. There are also no obviously discrepancies in overlap capacitance between gate and source/drain (Miller capacitance) which indicates the same overlap and activation conditions for both anneal schemes as shown in the inset in Fig. 3 and Fig. 4.

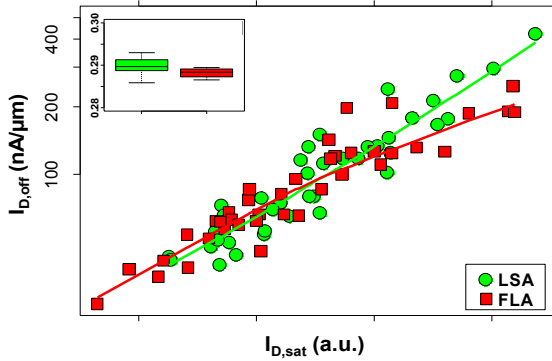


Fig. 3: $I_{D,sat}$ - $I_{D,off}$ characteristics of n-MOSFET devices annealed only with LSA or FLA (Inset: Miller capacitance in fF/ μm &side).

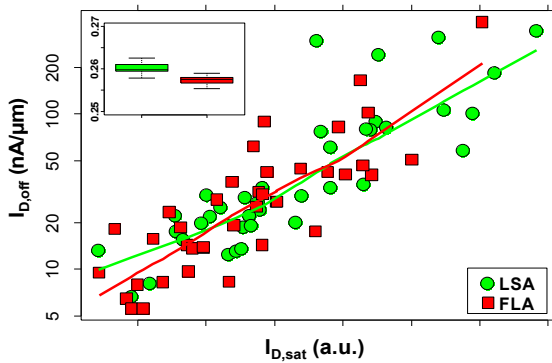


Fig. 4: $I_{D,sat}$ - $I_{D,off}$ characteristics of p-MOSFET devices annealed only with LSA or FLA (Inset: Miller capacitance in fF/ μm &side).

Junction leakage current

A junction leakage current, which degrades the subthreshold leakage current, is also one of the major concerns in case of using millisecond anneal, because such short time annealing seems to be insufficient to recover damages introduced by ion implantation. As shown in Fig. 5 the off-state junction leakage for n- and p-MOSFET are similar for LSA and FLA. This indicates a comparable thermal budget for both annealing schemes which leads to equivalent p/n-junctions in terms of steepness and defect annealing.

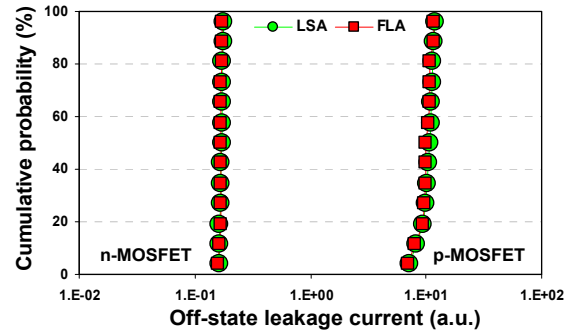


Fig. 5: Off-state junction leakage current of n- and p-MOSFET annealed only with LSA or FLA.

Pattern dependence

Although the discussion above mentioned is applied to structures where pattern effects are already taking into consideration, deeper investigations were done with the help of the test structure in Fig. 2 to reveal possible pattern-related temperature non-uniformities between LSA and FLA. Fig. 6 and Fig. 7 show the relation between the saturation threshold voltage and the STI density respectively for n- and p-MOSFET annealed only with LSA or FLA.

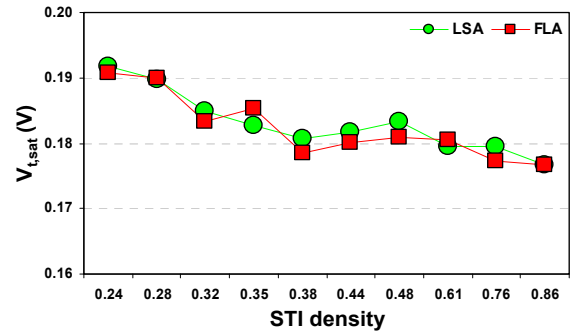


Fig. 6: $V_{t,sat}$ vs. STI density of n-MOSFET devices annealed only with LSA or FLA.

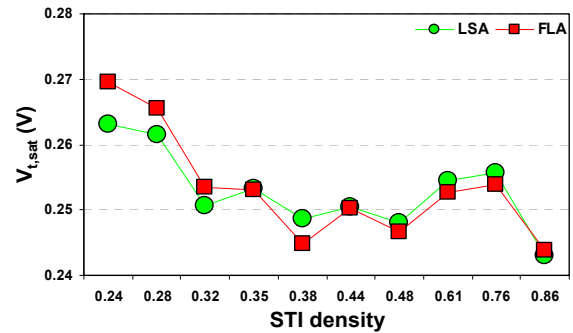


Fig. 7: $V_{t,sat}$ vs. STI density of p-MOSFET devices annealed only with LSA or FLA.

Two major facts are obvious:

1. There are only marginal threshold voltage differences between LSA and FLA for both n- and p-MOSFET devices which indicates comparable peak anneal temperatures for both millisecond annealing techniques independent on STI density.
2. The threshold voltage decreases as STI density increases which means that the temperature increases with increased STI density. This can be explained by the higher emissivity of a STI structure (0.7) compared silicon active areas (0.57) [9].

Fig. 8 and Fig. 9 show the relation between the saturation threshold voltage and the STI density respectively for n- and p-MOSFET annealed with RTA+LSA or only LSA.

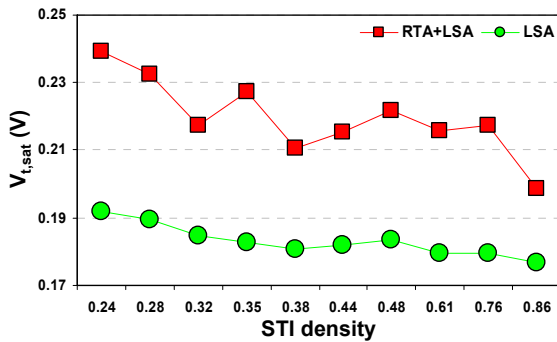


Fig. 8: $V_{t,sat}$ vs. STI density of n-MOSFET devices annealed with RTA+LSA or only with LSA.

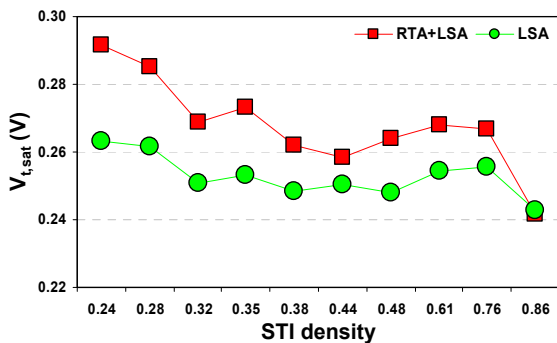


Fig. 9: $V_{t,sat}$ vs. STI density of p-MOSFET devices annealed with RTA+LSA or only with LSA.

A reduction in parameter fluctuation between the highest and lowest STI density of about 50 % is achievable by removing the RTA process. The reason for this enhancement in parameter variation is that few degrees of temperature non-uniformity modify the device performance of a RTA process more than few degrees of temperature variation of a diffusionless process due to more or less dopant diffusion.

Conclusions

Non-melt laser spike annealing and flash lamp annealing were compared in terms of transistor performance and pattern loading effects on SOI-CMOSFETs for the 32 nm node and below. The data shows neither significant differences between the two different millisecond annealing techniques, nor the need for absorption layers. Furthermore, a reduction in parameter fluctuation of about 50 % is achievable with a diffusionless process compared to a process with RTA. Both millisecond annealing techniques are therefore equal useable for implant annealing for next generation CMOSFETs.

Acknowledgement

This project was funded by the German Federal Ministry of Education and Research, registered under funding number 01M3167B. The author named in the publication bears responsibility for all published contents.

References

- [1] S. Talwar, et al., Solid State Techn., 2003.
- [2] T. Ito, et al., ISSM, p. 19, 2002.
- [3] T. Yamamoto, et al., VLSI, p. 122, 2007.
- [4] T. Sanuki, et al., IEDM, p. 281, 2007.
- [5] K. Adachi, et al., VLSI, p. 142, 2005.
- [6] T. Ito, et al., IWJT, p. 63, 2005.
- [7] L.M. Feng, et al., IWJT, p. 25, 2006.
- [8] A. Colin, et al., Material Science and Eng. B, Vol. 154-155, p. 31, 2008.
- [9] T. Sanuki, et al., IEDM, p. 281, 2007.