

# An Embedded 90nm SONOS Flash EEPROM Utilizing Hot Electron Injection Programming and 2-Sided Hot Hole Injection Erase

E. J. Prinz, G. L. Chindalore, K. Harber, C. M. Hong, C. B. Li, C. T. Swift

Motorola Embedded Memory Center, 6501 William Cannon Dr., MD: OE341, Austin, Texas 78735, U.S.A.

(512) 895-8443, Fax: (512) 895-8605, *Erwin.Prinz@Motorola.com*

## **Abstract**

A SONOS thin film storage non-volatile memory bitcell is programmed with hot electron injection, and erased with 2-sided hot hole injection from both the source and drain sides. Application of source-to-well reverse bias during programming reduces the programming current, and also enables peripheral voltage scaling to +/- 6V while maintaining a 1-10 $\mu$ sec byte programming time. The programmed charge in the thin film includes a secondary electron distribution above the channel area, reducing asymmetry between forward and reverse read currents. Erase is performed with band-to-band tunneling induced hot hole injection from both source and drain sides. The bitcell effective channel length is less than twice the width of the hot hole profile injected into the storage film, such that both the primary and secondary electron peaks formed during programming can be compensated with holes during erase. The bitcell has been integrated into a common source NOR Flash EEPROM architecture and embedded into a 90nm high performance CMOS process with fully silicided sources and drains and a copper backend [1, 2]. Bitcell size can approach that of a diffusion ROM.

## **Introduction**

Microcontrollers with embedded fast access time non-volatile memory (NVM) are utilized in applications such as automotive engine control for code and data storage. The established technology for these applications is floating gate bitcells in a common source NOR architecture, operated with various combinations of Fowler-Nordheim tunneling and hot electron injection [3]. Voltage scaling has largely leveled off at ~9V due to the onset of stress-induced leakage current (after write/erase cycling) in tunnel oxides below 90 $\text{\AA}$  [4, 5]. Fast byte programming time favors hot electron injection as a program mechanism. Parasitic drain current can be reduced somewhat by applying a source/well reverse bias [6]. Due to simpler processing, NVM bitcells based on charge storage in non-conducting films such as SONOS have recently been considered, utilizing HCI programming and hot hole erase [7]. Multi-bit storage has been shown in the NROM, requiring charge to be fully localized at the junction edges [8]. In this work, an embedded NVM combining the process simplicity of thin film storage with reduced peripheral voltages, fast programming time, and reduced programming current is demonstrated.

## **Bitcell Operation**

A bitcell (see Fig. 1) is operated with the voltages shown in Fig. 2. Programming is performed by hot electron injection with a source/well reverse bias of 2V. Erase is performed by 2-sided hot hole injection resulting from band-to-band tunneling with positive source/drain and negative gate voltages. Hot hole injection can be simultaneous from source and drain, or sequential. The bitcell is read and programmed in the same direction. The bitcell length is shorter than twice the width of the distribution of injected holes, such that the entire channel area can be covered by hot hole injection, as shown in Fig. 3. The programming time with 3V drain bias and -2V well-to-source reverse bias is in the 1-10 $\mu$ sec range, similar to what can be achieved with no reverse bias and increased drain and/or gate biases (see Fig. 4). The highest peripheral voltage required during program and erase factors prominently into Flash EEPROM module size. Therefore, a 6V gate voltage level is preferred over 9V. Among the three cases shown, the 3V drain and -2V well-to-source bias has the lowest programming current, as shown from the drain current/ drain voltage data in Fig. 5. Applying a back bias reduces the programming current by 40%.

Programming with a reverse source/well bias results in a charge distribution in the nitride which is less localized as compared to the case with no reverse bias. Therefore, due to the limited extent of the distribution of injected hot holes, hot hole erase becomes less efficient with longer bitcells (see Fig. 6). For the bitcell considered, a bitcell length of less than 0.22 $\mu$ m is required for efficient 2-sided hot hole erase. The maximum erase current per side for 2-sided hot hole erase is in the 1-10nA range, as shown in the erase current curve of Fig. 7.

The bitcell (W/L=0.27 $\mu$ m/0.22 $\mu$ m) has a read current at 1V gate overdrive of about 20 $\mu$ A, see Fig. 8. The localization of the charge distribution in the nitride can be assessed by comparing forward to reverse read operations (compared to the direction of HCI programming). It is shown in Fig. 8 that application of a reverse well-to-source bias greatly reduces the asymmetry, so the bitcell is operated with forward read. Even with fixed program/erase voltages and times, no charge is accumulated above the channel of the device, resulting in approximately flat program/erase cycling curves (see Fig. 9). Data retention and read disturb at 150 $^{\circ}$ C after write/erase cycling demonstrates an operating window, as shown in Fig. 10. Preliminary data from a 0.25 $\mu$ m test chip (1024 bits of 3 parts) show an acceptable 0.7V charge loss (with no write/erase cycling) of the programmed state after 24 hours at 250 $^{\circ}$ C (see Fig. 11).

In the Flash EEPROM module, the bitcells are arranged in a common source NOR architecture with silicided active area source connections, similar to a diffusion ROM layout, as shown in Fig. 1. The bitcell size is reduced compared to a floating gate bitcell due to the absence of the floating gate pattern over field oxide separating two floating gates. Also, the application of reverse well-to-source bias enables further channel length scaling.

## **Process Integration**

The bitcell shown above has been embedded into a 90nm high performance logic host process [1, 2]. The bitcell trench isolation module is shared with logic and peripheral transistors. The bitcell is located in a P-well which is fully isolated from the P-type substrate by a high energy implanted deep N-well. The charge-storing film consists of a 70 $\text{\AA}$ -thick thermally grown bottom oxide, 150 $\text{\AA}$  CVD nitride, and a 50 $\text{\AA}$  top oxide thermally grown in steam ambient. All NVM-related high temperature processing steps are done before the low voltage logic well implants to maintain the logic transistor characteristics. The bitcell source/drains are implanted with pre-spacer arsenic extension and optional boron halo implants. The SONOS nitride in the bitcell source/drain areas is removed simultaneously with the gate poly antireflective coating, such that all sources, wordlines, and drains are fully silicided with cobalt silicide, resulting in a sheet resistance of about 10  $\Omega$ /sq. A copper backend with 4 layers of metal has been used in this work. The bitcell processing adds 2-3 masks to the process.

## **Conclusion**

An embedded non-volatile memory based on a common source NOR array of SONOS bitcells, operated with HCI programming with reverse source-to-well bias, and erased by 2-sided hot hole injection, has been proposed for the first time, and has been integrated into a 90nm high performance logic technology. Peripheral voltage reduction to +/- 6V has been achieved. The proposed memory is suitable for fast read access time even with large block sizes due to low source/drain resistance.

## **Acknowledgments**

We acknowledge Bryan Acred, Jerry Peschke, and Sandra Altmeyer for technician support; Motorola MOS12, Dan Noble Center, and MOS13, for silicon fabrication; and Ko-Min Chang for managerial support.

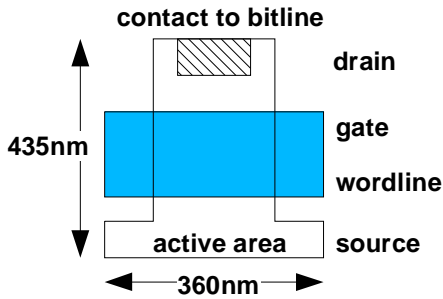


Fig. 1: Bitcell layout.

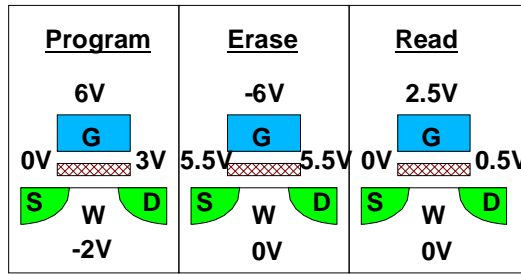


Fig. 2: Schematic of write, erase, and read operations.

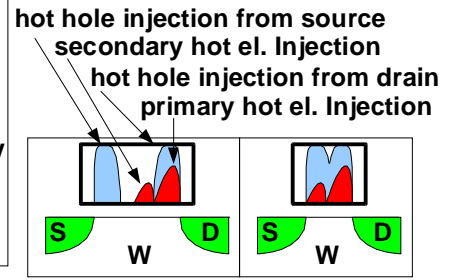


Fig. 3: Charge distribution after hot hole injection for a "long" and a "short" bitcell.

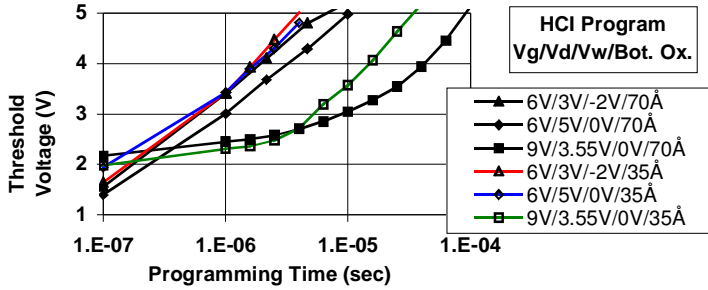


Fig. 4: Programmed threshold voltage ( $V_t$ ) vs. time with and without back bias.

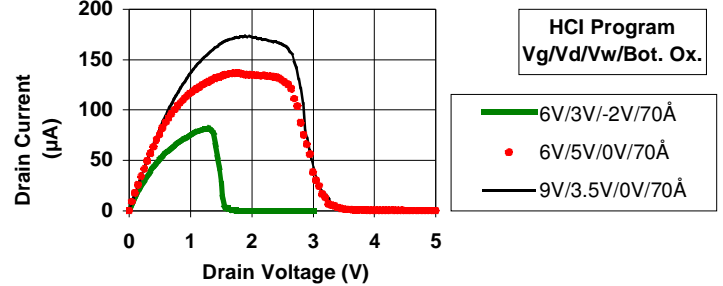


Fig. 5: Programming current vs. drain voltage (one shot curve) with and without back bias.

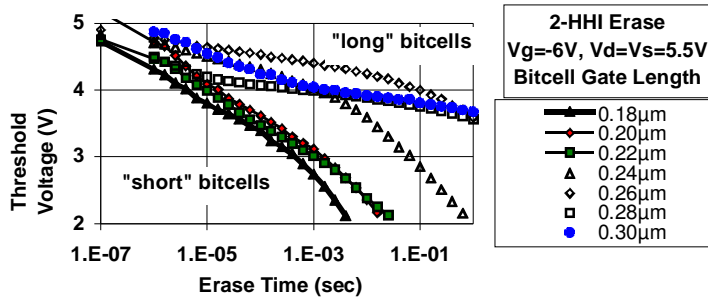


Fig. 6: Erased threshold voltage ( $V_t$ ) vs. time with and without negative gate bias.

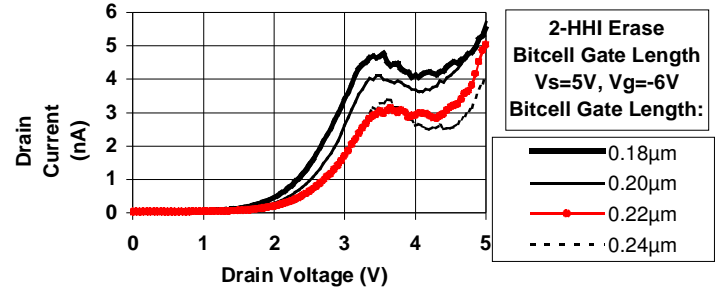


Fig. 7: Erase current vs. source/drain voltage (one shot curve) with and without negative gate bias.

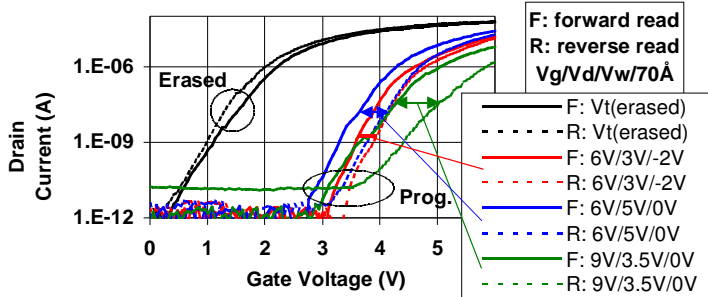


Fig. 8: Bitcell forward and reverse read currents for the erased bitcell and bitcells programmed with various voltages.

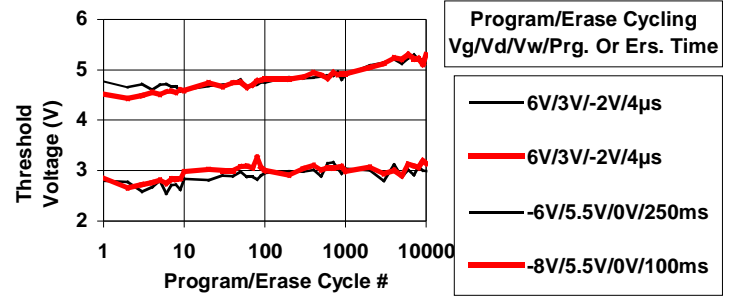


Fig. 9: Write/erase cycling of bitcell, programmed with  $(V_g/V_d/V_w)=(6V/3V/-2V)$ , and 2-HHI-erased with  $(V_g/V_d/V_w)=(-6V/5.5V/0V)$ .

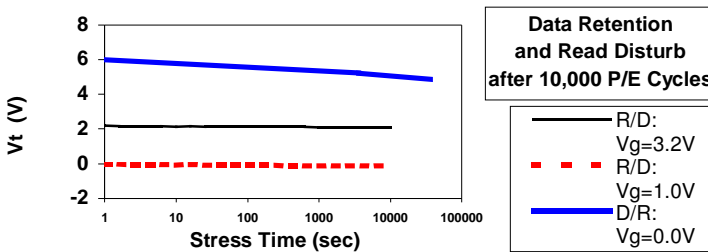


Fig. 10: Data retention and read disturb at  $150^\circ\text{C}$  vs. time after 10,000 program/erase cycles.

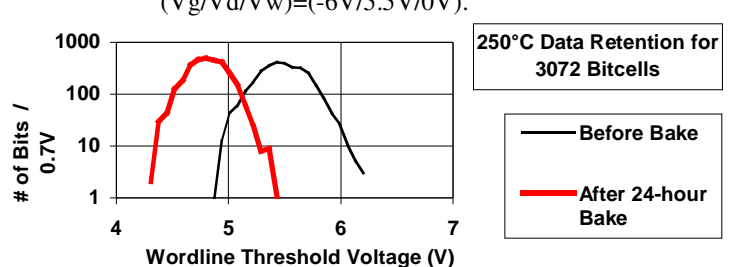


Fig. 11: Programmed threshold voltage distribution from a  $0.25\mu\text{m}$  test chip before and after a 24hr,  $250^\circ\text{C}$  bake.

**References**

[1] S. Parihar et al., 2001 IEDM Tech. Dig., p. 11.4.1, 2001.  
 [3] D. Burnett et al., 2000 NVSM Workshop, p. 59, 2000.  
 [5] H.P. Belgal et al., Proc. 2002 Reliability Physics Symposium, p. 7.  
 [7] T. Y. Chan et al., IEEE Elec. Dev. Lett. 8, p. 93, 1987.

[2] G. Yeap et al., 2002 VLSI Tech. Symp., p. 16, 2002.  
 [4] A. Hoefler et al., Proc. 2002 Reliability Physics Symposium, p.21.  
 [6] J. Bude et al., 1995 IEDM Tech. Dig., p. 989, 1995.  
 [8] B. Eitan et al., IEEE Elec. Dev. Lett. 21, p. 543, 2000.