A PROM Element Based on Salicide Agglomeration of Poly Fuses in a CMOS Logic Process

Mohsen Alavi , Mark Bohr, Jeff Hicks, Martin Denham, Allen Cassens, Dave Douglas, Min-Chun Tsai
Intel Corporation, Portland Technology Development
Hillsboro, OR

Abstract

A novel programmable element has been developed and evaluated for state of the art CMOS processes. This element is based on agglomeration of the Ti-silicide layer on top of poly fuses. Various aspects of these programmable devices including characterization and optimization of physical and electrical aspects of the element, programming yield, and reliability have been studied. Development of a novel programming and sensing circuit is also included.

Introduction

The capability of implementing a small PROM array on logic products at no additional process cost is highly desirable for a number of applications such as redundancy implementation in SRAMs, die identification, electrically programmable feature selection, etc.

As CMOS technology scaled, gate oxides became thin enough that implementation of flash memory cells on standard logic CMOS processes (SPEED) became possible [1]. However, further scaling of CMOS technology resulted in inadequate charge retention in the SPEED device due to tunneling of carriers through the gate oxide.

The element presented here avoids the problem with scaled gate oxide thickness. The results are a fuse element which is reliable under thermo-mechanical and bias-temperature stress while enjoying near 100% programming success used in a specially designed circuit. Programming the fuse does not result in any collateral damage in overlying or underlying layers and may be performed nominally at 2.5V and 10 mA in 100 ms.

General Description of the Element

A. Physical Properties

The poly agglomeration fuse (PAF) is made from a polysilicon line shunted on top by a layer of Ti-silicide which is used as the gate in CMOS processes. It is programmed via current stress which results in temperatures high enough to cause agglomeration of the Ti-silicide [2]. The damage due to programming of the element has been found to be very subtle and confined to the Ti-silicide and its interface with the underlying poly layer and the overlying dielectric. The integrity of the entire overlying stack from the passivation to the overlying ILD is found to be intact and no collateral damage has been observed (see Fig. 1,2). This is in contrast with traditional poly or metal fuses which require openings in the overlying layers to facilitate removal of fuse material, and therefore, a post program passivation step. Typically, a fuse link is drawn at minimum allowable width with a few microns of length (see Fig. 2). The effect of fuse doping and geometry on its performance has been investigated extensively and will follow.

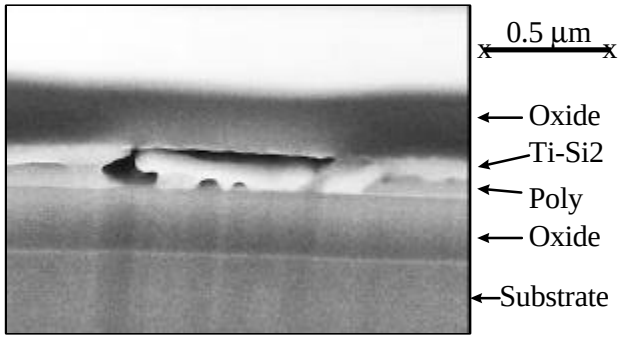

Fig. 1, Cross section of the damaged section of a programmed fuse. Lack of collateral damage to the overlying and underlying layers is evident..

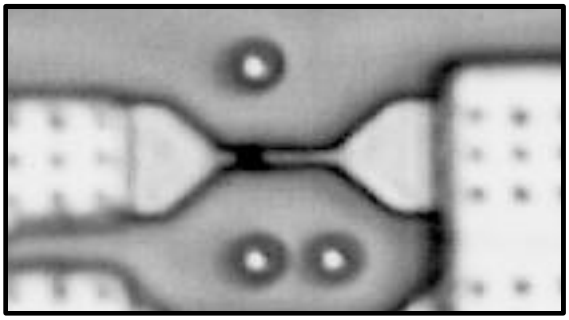

Fig. 2, Top view of a programmed fuse. The subtle damage due to programming is evident on the left side of the element.

B. Electrical Properties

Prior to programming, electrical properties of the fuse are determined by the salicide layer on top which has a sheet resistance of about 4 ohms per square in our study, resulting in a typical resistance of about 50-100 ohms depending on the dimensions of the fuse. Injection of current beyond a certain level results in a sudden increase in resistance indicating formation of discontinuities in the silicide layer. The value of this resistance varies greatly from device to device. In our structures, post program resistance varied from several hundred Ohms to several hundred kOhms. Post program I-V characteristics are found to be nonlinear and therefore, the value of resistance varies with applied bias.

Element Characterization

A. Test Structures

The element described above has been implemented in a 0.25um CMOS process with a poly thickness of about 0.2um [3]. Ti-silicide films resulting in sheet resistance ranging

from 3 to 4 ohms per square have been studied. Initial and post program electrical characteristics of a variety of element designs have been investigated. This includes the effects of poly doping (n, p, undoped), fuse length and width, fuse shape, programming and sensing voltage and current, and programming time.

B. Programming dynamics

In order to program an element, a certain amount of current is needed. The voltage needed for injecting this current must obviously be smaller than the available power supply voltage. Under constant voltage stress, as the element gets hot enough, agglomeration starts to occur, thereby, increasing the element resistance. As a result, the current through the element drops to a low value consistent with the elements final resistance and the element cools down. This mechanism is one with negative feedback. Therefore, a given fuse may be stressed only once and it's post program resistance will not increase with additional voltage stress.

Figure 3 shows the I-V characteristics of a typical fuse element. As the voltage is increased, current increases in a nonlinear fashion due to resistance change caused by self heating. When the dissipated power reaches a critical value, fusing occurs and element goes to a much higher resistance.

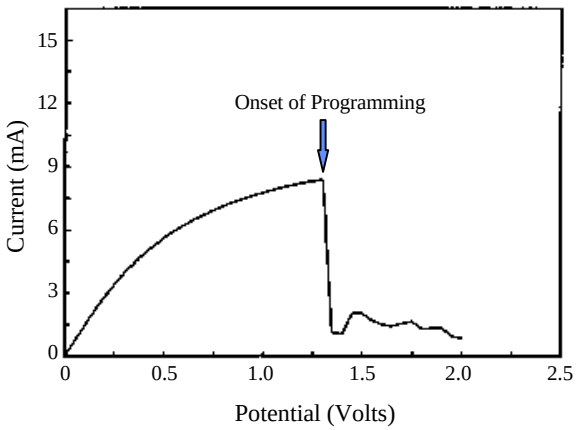

Figure 3, I-V characteristics of a typical element upon programming.

C. Response parameters

Initial and post program resistance of the element are the two key parameters affecting any circuit meant to sense the state of the element. A maximum value of initial resistance and a minimum value of post program resistance are needed to guarantee proper circuit function (about 100Ω and 1kΩ respectively in our circuit).

Initial fuse resistance depends on element geometry and silicide thickness and quality. Silicide quality in turn depends on process conditions, poly line width, and doping [2,4]. Silicide imperfections are more likely for long narrow elements and best silicide lines were found to be the ones made from p-doped poly. Imperfections in the silicide layer (cracks, high resistance Ti-Si phase) result in a resistive element Figure 4 shows cumulative distribution of the resistance of a typical fuse structure made with two processes with different thermal cycles and Ti thickness. A high resistance tail corresponding to silicide imperfections is evident in the distribution of the resistance of the unoptimized process.

Post program resistance varies greatly from device to device and depends on the shape and size of the discontinuity in the link. Due to this variation, any aspect of this resistance must be studied statistically. Many factors affect the level of fusing and therefore, post program resistance. They include:

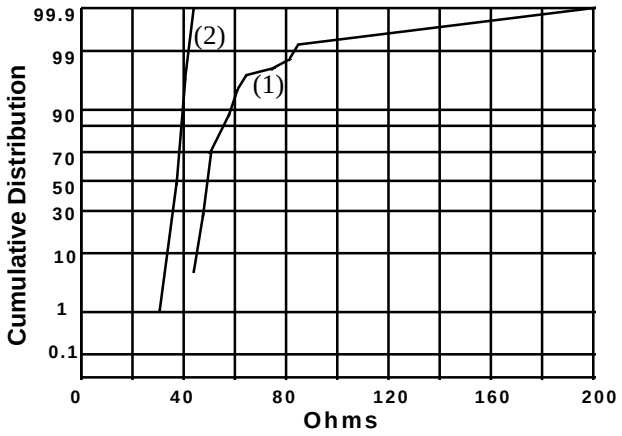

Figure 4, Resistance of a typical fuse. (1) Unoptimized silicide process, (2) Optimized silicide process.

Programming voltage, current, and time: Even though fusing can occur quickly and at fairly low currents and voltages (in the order of 1V, 8mA, 1mS), post program resistance is significantly enhanced if more energy is dumped into the element Therefore, increased voltage and current levels are needed for a longer time to guarantee a sufficiently large resistance. In this work, minimum programming conditions which resulted in statistically adequate post program resistance were a current of 20mA injected for 100ms with a voltage compliance of 2.5V.

Initial fuse integrity: Measured data shows that fuses that are more robust initially (by process or geometry) result in more successfully programmed elements. This is due to the fact that for a given voltage compliance and a given value of fusing current, a smaller resistance results in a larger amount of energy transferred to the device. The fortuitous result is that process conditions which result in good silicide formation and robust unprogrammed fuses also produce elements which program successfully.

Fuse shape: In addition to the relation between fuse size and it's initial resistance, the shape of the fuse has a marked effect on the distribution of its post program resistance. This has been found to be due to the fact that in addition to the high temperature necessary for agglomeration, the level of

temperature gradient (and therefore stress) in the element plays a key role in the fusing event. Fusing has been found to occur near the sides of the element close to the point which has the highest temperature gradient (see Figure 2,5). Additionally, line width plays a significant role in fusing success with narrower lines having the advantage of better fusing. Figure 6 shows four different fuse shapes of the same length. Figures 7,8 show the distribution of initial and post program resistance for these elements. The difference between post program resistance of elements a,b corresponds to the effect of element width while differences between structures c,d show the effect of temperature gradient.

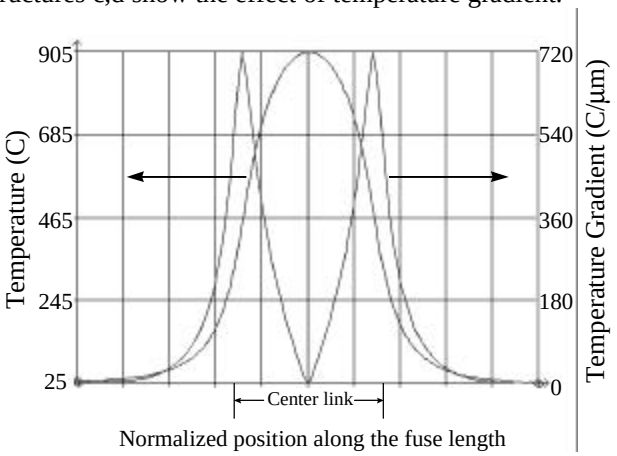

Figure 5, Simulation results of profiles of temperature, and temperature gradient along the length of a typical fuse at nominal programming bias.

Modeling and Simulation

In order to look for an optimum fuse design, numerical simulation of temperature in the element under current stress has been performed. The simulation is based on a two dimensional model with an added loss term to the overlying and underlying layers. Thermal conductivity of the silicide layer and the heat loss coefficient were fitting parameters. Assuming that fusing occurs when the temperature of the fuse reaches 800C (silicide agglomeration temperature [2]), the simulation is able to predict fusing current using a single set of fitting parameters for various fuse geometry with good accuracy (see Fig. 5,9) and provides insight into the distribution of temperature and its gradient in the element.

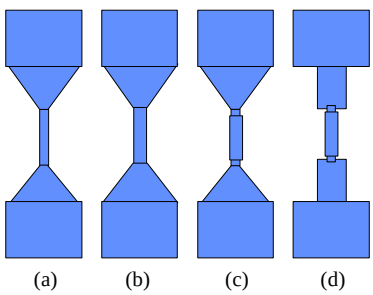

Figure 6, various fuse shapes. All elements are p-type, about 2um long. (a) width=0.22um, (b) width=0.27um, (c, d) width = 0.22um/0.27um.

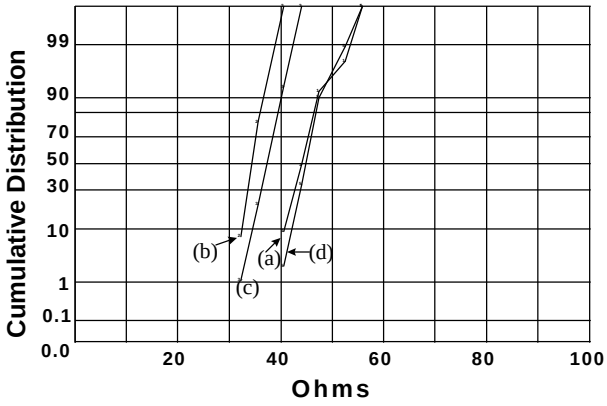

Figure 7, Pre-programmed fuse resistance of structures in figure 6.

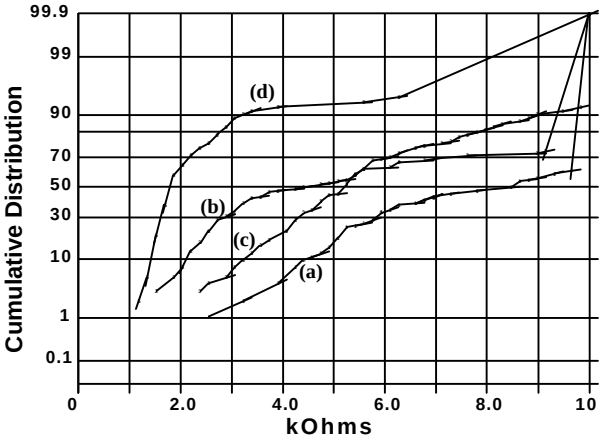

Figure 8, Post program resistance of fuses in Figure 6.

Sensing Circuit

A special circuit has been developed for programming and sensing the element. Figure 10 shows a simplified schematic of this circuit. Programming occurs when a logic LO is asserted on the gate of a large PMOS transistor. Since the fuse programs at relatively low bias, logic and programming circuits share a common supply voltage.

The sensing circuit is a novel and well-balanced solution to a stringent set of requirements, the foremost being that the sensing currents must be kept very low. The core of this circuit comprises a pair of matched N-channel transistors, which perform the sensing, and a pair of matched P-channel devices, which act as current-sensing output loads. The N-channel sensing transistors are connected in a current mirror configuration, such that, if the fuse-reference resistance on the left were equal to the unburned fuse resistance on the right, both circuit branches would have equal current. In practice, the reference resistance is set to about 8 times that of the unburned fuse. This ratio of reference to fuse creates a default (unburned) output voltage that is low enough to be interpreted as a logic LO value. Additionally, for a programmed fuse, the resulting output voltage is sufficiently high to be interpreted

as logic HI. Therefore, the gain of the circuit is sufficient for single-ended voltage outputs.

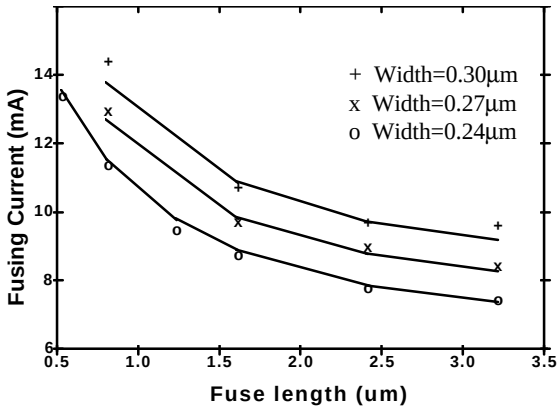

Figure 9, Measured and simulated current at the onset of fusing. Solid lines show simulation results. Symbols show measured data points

In this circuit, the ratio of reference to unburned fuse resistance represents a balanced tradeoff between output high voltage (VOH) and output low voltage (VOL) levels. With a ratio of 8, noise margins for VOH and VOL signals are roughly equal. The resulting noise margin is adequate to guard-band the circuit from expected manufacturing variations in transistor Vt and channel length.

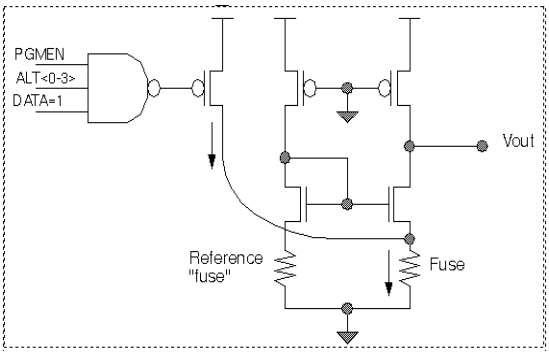

Figure 10, Schematic of a simplified programming and sensing circuit.

Yield and Reliability

A PROM array based on the PAF will suffer yield loss if the programmed fuse does not have a high enough resistance to be properly sensed. Programming yield depends on the fuse design (see Fig. 8), array size, and circuit design. Even after optimizing the element and circuit, the resulting yield may not be as high as expected. In that case, redundant fuse elements are needed such that if programming of a given fuse in a given memory bit is not successful, an additional fuse is available in that bit for an extra attempt. In this work, for a 64 bit array, a programming yield loss of less than 1 in 10,000 was achieved using two fuses per bit (a programmed state in either fuse resulted in a programmed bit).

The reliability of this element was characterized by placing a large number of samples (programmed and unprogrammed) under thermo-mechanical stress (1000 cycles of condition 'C' temperature shock) and in bake (300 hours, 250C). The element was found to be quite stable under these conditions (see Fig. 11). Additional testing was done to characterize the stability of the unprogrammed fuse under bias temperature stress. Results indicated that as long as the sensing current is significantly less than the current at the onset of programming, the device will remain stable.

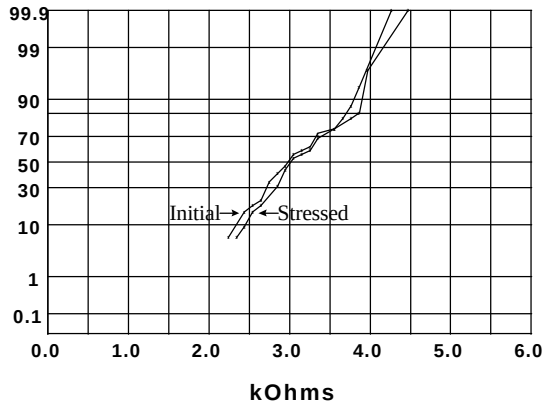

Figure 11, Post program fuse resistance distribution of a typical fuse before and after 300HR 250C bake.

Conclusions

Poly agglomeration fuse is a reliable programmable element which may be implemented in a logic CMOS processes. This element may be programmed under nominal bias and does not introduce any collateral damage. Distribution of the post program resistance depends on silicide quality, fuse shape, doping, and programming conditions. Optimized conditions for fuse shape and programming parameters have been presented using empirical results and numerical simulations and a novel circuit has been presented for the device with a 1 in 10,000 programming yield loss for a 64 bit PROM arrays with 2 fuses per bit. Element reliability has been verified under temperatue shock and bake.

References

1) K. Ohsaki, N. Asamoto, S. Takagaki, "A Single Poly EEPROM Cell Structure for Use in Standard COMS Processes", *IEEE J. Solid State. Circuits*, Vol 29, No. 3, March 1994, PP. 311-316
2) J.B. Lasky, J.S. Nakos, O.J. Cain, P.J. Geiss, "Comparison of Transformation to Low-Resistivity Phase and Agglomeration of TiSi2 and CoSi2, *IEEE Trans. Elect. Devices*, Vol. 38, No. 2, Feb. 1991, pp. 262-269.
3) M. Bohr, et. al., "A High Performance 0.25µm Logic Technology Optimized for 1.8V Operation", *1996 IEDM Tech Digest*, 1996, pp. 847-850.
4) J.A. Kittle, Q-Z Hong, D.A. Prinslow, G.R. Misum, "A Ti Salicide Process for 0.10 um Gate Length CMOS Technology", *1996 VLSI Symp. Tech. Digest*, 1996, pp. 14,15.