computers MDPI

Reliability of NAND Flash Memories: Planar Cells and Emerging Issues in 3D Devices

Alessandro S. Spinelli *, Christian Monzio Compagnoni and Andrea L. Lacaita

Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133 Milano, Italy; christian.monzio@polimi.it (C.M.C.); andrea.lacaita@polimi.it (A.L.L.)
* Correspondence: alessandro.spinelli@polimi.it; Tel.: +39-02-2399-4001; Fax: +39-02-2399-3574

Abstract: We review the state-of-the-art in the understanding of planar NAND Flash memory reliability and discuss how the recent move to three-dimensional (3D) devices has affected this field. Particular emphasis is placed on mechanisms developing along the lifetime of the memory array, as opposed to time-zero or technological issues, and the viewpoint is focused on the understanding of the root causes. The impressive amount of published work demonstrates that Flash reliability is a complex yet well-understood field, where nonetheless tighter and tighter constraints are set by device scaling. Three-dimensional NAND have offset the traditional scaling scenario, leading to an improvement in performance and reliability while raising new issues to be dealt with, determined by the newer and more complex cell and array architectures as well as operation modes. A thorough understanding of the complex phenomena involved in the operation and reliability of NAND cells remains vital for the development of future technology nodes.

---

1. Introduction

The concept of electrically programming and erasing a semiconductor memory cell dates back to the 1970s, when a floating-gate device with a thinned oxide layer in a portion of the channel area was successfully demonstrated [1] as a viable alternative to ultra-violet erasable cells. However, it was not until the development of the one-transistor Flash memory device [2], in which the entire chip could be erased at once, that solid-state non-volatile code and data storage began to experience a tremendous growth. Data storage, in particular, was fostered by the invention of the NAND Flash [3], which is possibly one of the most successful semiconductor devices ever developed. Thanks to its extremely compact area occupation and to clever system design and optimization, it benefited from the market explosion of digital cameras first, followed by smartphones and tablets and culminating in the development of NAND solid-state drives (SSDs) [4], that have recently overtaken hard-disk drives (HDDs) in terms of areal storage density [5]. Although the discussion on whether SSDs will eventually replace HDDs aside from small consumer devices is far from settling, it is indubitable that more and more applications are falling within reach of SSDs, that are now targeting enterprise-class storage [6,7].

The NAND success story is the result of continuous scaling of the planar cell dimensions, that has led all major NAND Flash manufacturers to reach the 15 nm planar node [8–11]. However, several constraints have severely challenged the latest nodes development [12–14], eventually putting an end to the scaling of the planar Flash technology. Some of these constraints are process-based and reflect the increasing difficulty in patterning smaller and smaller features with the ArF immersion lithography. Then, traditional device design issues arise, such as the need to curb short-channel effects (exacerbated by the thick gate oxide, which has very limited scaling possibilities) and avoid degradation of the subthreshold slope (reducing the operation margin) while limiting junction leakage (which impairs

channel boosting during program). Finally, there are issues that are specific to memories, such as cell-to-cell interference and program noise, as discussed in detail in [15]. On top of all this, reliability issues, to be discussed in the following, played a key role in challenging further scaling.

This increasingly dramatic scenario prompted all major NAND Flash manufacturers to seek an alternative way to continue to deliver increasing storage capacity, that has resulted in the demise of the planar technology and the pursuit of three-dimensional (3D) memory arrays. In this field, a pioneering proposal was put forward in 2001 [16], evolving into early proposals based on layer stacking [17–19] and reaching the first cost-effective solution of a 3D-integrated array in 2007 [20]. From there on, 3D NAND based on vertical channels have taken the lead in the quest for higher density, resulting in the first commercial product in 2013 [21,22] and reaching a capacity of 786 Gb with a density of 4.29 Gb/mm^2 [23] while stacking up to 64 layers [23–25]. Chip capacity was boosted not only by 3D integration, but also by multi-bit storage: while two bit-per-cell (i.e., multi-level cells, MLCs) were used in the first prototype [21], today's technology relies on storing three bit-per-cell (triple-level cells or TLCs) and quadruple-level cells (QLCs) are under active development. Recent reviews of the different 3D NAND technologies and architectures can be found in [26–28].

The revolutionary move to 3D memories was not without problems, as discussed in [29]. In fact, while traditional device reliability issues must still be accounted for, new phenomena came to prominence, which are not merely an extension of what has already been observed on planar devices, but rather a consequence of the different device and array fabrication. The aim of this paper is then to review the main reliability issues that play a key role in planar NAND Flash devices and discuss the main experimental data and current understanding. Then, 3D NAND reliability is addressed, highlighting the role of such "traditional" mechanisms in the new cell size and geometry and introducing issues that are specific to the new memory cell. The analysis of the literature—though selected—reveals on the one hand the complexity of Flash reliability and, on the other hand, the huge efforts undertaken by the scientific community to understand the root causes of the experimental observations and develop effective predictive models that allowed a successful product design. Three-dimensional NAND bring further challenges to the table, with tighter requirements in an increasingly complex device, meaning that Flash reliability will offer an exciting field for scientific investigation for many years to come.

2. The NAND Memory Array

This Section contains a brief discussion on the planar NAND memory array architecture and program operations, with the aim of providing the basic knowledge to understand the rest of the paper. We will not discuss in detail the different array operations but rather restrict ourselves to the general principles, referring to specific literature for more details [27,30,31].

2.1. Array Architecture and Layout

The basic element at the heart of the NAND memory is the floating-gate (FG) transistor, in which a polysilicon island is sandwiched between a gate (or tunnel) oxide and an interpoly dielectric, in turn connected to a control gate. Whenever a charge Q is stored in the floating gate, simple electrostatics [32] shows that the threshold voltage V_T of the cell becomes

$$V_T = V_{T0} - \frac{Q}{C_{pp}},$$

(1)

where V_{T0} is the so-called *neutral* threshold voltage, corresponding to zero charge on the floating gate, and C_{pp} is the interpoly dielectric capacitance. As charge cannot easily leak from the isolated floating gate, a non-volatile memory element is achieved. In the NAND array, transistors are connected in series along *strings*, while their control gates, connected across different strings, constitute the wordlines (WLs), each making a *page*, as schematically depicted on the left of Figure 1 (more than one page can be associated to the same WL, but this is irrelevant for our purposes). Strings are connected to the

bitlines (BLs) and the common sourceline via select transistors driven by a drain select line (DSL) or a source select line (SSL). This array constitutes a *block*, whose size is a key element in NAND array organization.
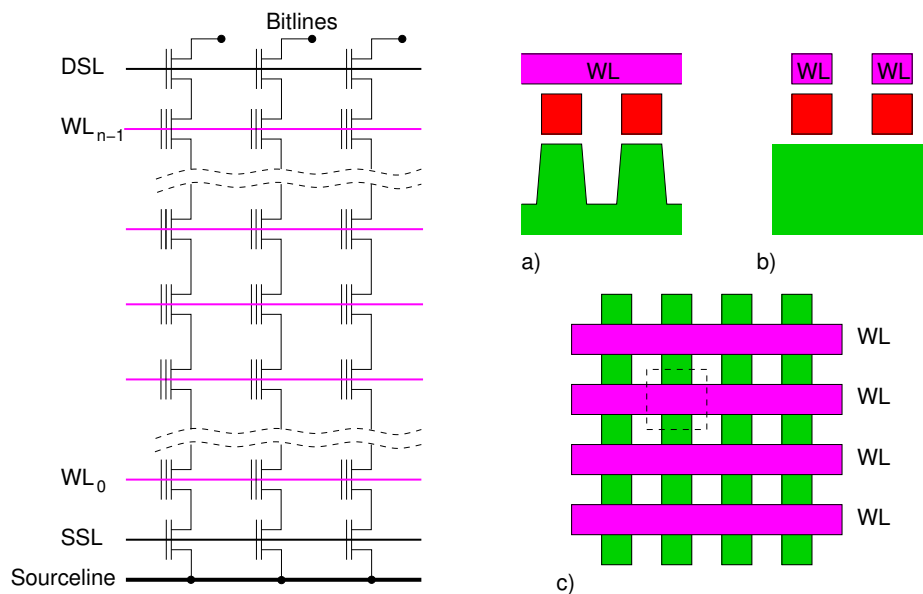

Figure 1. (**Left**) Schematic view of a NAND Flash array. Vertical strings of series-connected devices are attached to the bitlines and the sourceline via select transistors driven by a drain select line (DSL) and a source select line (SSL); (**Right**) schematic cross-sections of the array along the WL direction (**a**) and the string direction (**b**) (green = silicon, red = floating gate, magenta = WL, white = silicon oxide). Inset (**c**) shows the layout of the array with the elementary cell (dashed square).

Cross-sections of a current planar NAND array are shown on the right of Figure 1: Figure 1a depicts a typical cut along the WL (green = silicon, red = floating gate, magenta = WL, white = silicon oxide), highlighting the shallow trench isolation used to separate the active area of adjacent devices [33] and the planar structure of the cell, without any wrapping of the control gate along the floating gate sidewalls [34]. Figure 1b depicts instead the cross-section along the string direction. Figure 1c shows the layout of the memory array, which becomes extremely compact thanks to these solutions. In particular, given the feature size F of a given technology, i.e., the size of the smallest feature that can be created on silicon, the elementary memory cell (shown as a dashed square in Figure 1c) occupies the ideal area of $4F^2$. However, the *effective* cell area, i.e., total chip area divided by the number of cells, is somewhat larger than that because of the overhead circuitry, string contacts with BLs, dummy cells and select transistors. To decrease such overhead, the number of cells in a NAND string has reached the impressive number of 150 in the latest planar nodes [10].

2.2. Array Operation

As mentioned above, digital information is stored as charge onto the floating gate of a NAND cell. While the first devices relied on two discrete charge levels to store a single bit (the so-called single-level cells, SLCs), it was soon realized that the linear relation between floating-gate charge and cell V_T could be exploited to store more than one bit per cell [35,36]. Nowadays, TLCs are in production and used in SSDs [6], while quadruple-level cells (QLCs) are under active development [37]. Figure 2 shows the V_T distribution of a TLC NAND array: eight different levels are necessary to encode three bits. In the NAND architecture, all but one V_T level are positive (shown in different shades of blue) and are referred to as the *programmed* levels. The negative (red) V_T level is instead the *erased* one.
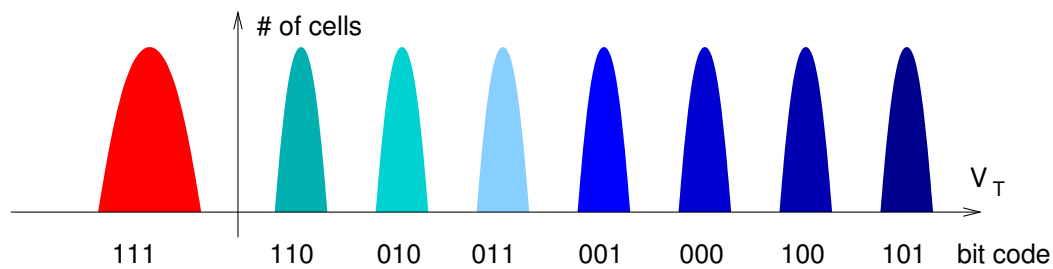
Figure 2. Pictorial view of the threshold voltage (V_T) distribution in a TLC cell. A three-bit code is associated to each of the V_T levels.

2.2.1. Read

Reading the state of a cell means discriminating one V_T level against another and irrespective of the V_T levels of all other cells in the string. To accomplish this, the string has first to be connected to the sensing circuit, i.e., the select lines are high and the BL is biased; therefore, a current can flow to the (grounded) sourceline. Unselected cells are then placed in a pass-transistor mode, biasing their WLs at a *pass voltage* larger than the maximum V_T. The actual read operation is then performed, sensing the cell current and comparing it against fixed reference values to assess the corresponding binary code [38]. The operation is performed more than one time for MLC and TLC cells. To increase parallelism, strings are read in parallel across a page (typically 8–16 kB), resulting in a throughput higher than 150 MB s^{-1}.

2.2.2. Program

NAND Flash cells adopt channel Fowler–Nordheim (FN) tunneling to store electrons onto their floating gates [39]. To this aim, the selected cell is first connected to the external circuitry by biasing the unselected WLs to a pass voltage and turning on the DSL, while the BL is grounded and the SSL is left off. A high voltage is then applied to the desired WL, resulting in tunnel electron injection from the inverted substrate to the floating gate. In MLCs or TLCs, fine placement of the V_T levels is needed because of the reduced distance between adjacent distributions (see Figure 2). This is achieved via an *incremental step-pulse programming* (ISPP) technique [40], in which fast pulses of increasing amplitude are used to inject a constant, controlled amount of charge onto the floating gate at each step, interleaved with *program verify* phases in which V_T is monitored. Several refinements are applied to this basic concept in 8- or 16-level cells [38].

Since the program operation is controlled by the WL bias, it acts on a page of the array. However, not all cells must be programmed at the same level (or programmed at all), which means that a program-inhibit strategy must be implemented for those cells. Inhibiting programming means reducing the gate-channel bias of such cells, thereby lowering their tunnel oxide field responsible for the electron injection. In NAND arrays, this is achieved via a *self-boosting* procedure [36], where the inhibited strings are first floated and their potential is temporarily raised via capacitive coupling to the WLs, that are raised at the passing voltage [38]. Program throughput in excess of 50 MB s^{-1} has been reported [10].

2.2.3. Erase

Erase operation is performed on blocks, i.e., a set of pages that can be 4–8 MB in size, exploiting the same FN mechanism used in program. As NAND chips do not allow the use of high negative voltages, WLs are grounded and a high positive voltage is applied to the *p* and *n* wells of the array. A verify phase follows, aimed at ensuring that all cells have negative V_T, which triggers another erase pulse if the condition is not met: the technique is similar to what is used during program, and is called *incremental step-pulse erasing*. Floating WLs are used to inhibit erase on unselected blocks.

3. NAND Flash Memory Reliability

From a general standpoint, the reliability of a product is related to its capability to continue to work as expected as a function of time. Reliability is enhanced by screening, testing and qualification and, eventually, by error management procedures. In all these respects, Flash memories are peculiarly different from other semiconductor devices, as they intrinsically operate under high-field conditions that other devices purposely avoid. Moreover, reliability is usually extrapolated from accelerated tests, that should replicate as close as possible the real operating conditions. Unfortunately, the development of such tests is complicated by the many different patterns of program/erase and read cycles under which Flash memories usually operate.

As a matter of fact, there are a number of failure modes and mechanisms that are specific to Flash, not to mention error-management procedures that are put forward by the manufacturers to increase reliability and yield. These depend on the type of failure, which can be recoverable (i.e., it can be fixed by erase and reprogram) or not. Recoverable errors are usually dealt with via error-correction codes (ECC). Initially, when SLCs were prevalent, simple Hamming algorithms were employed, which could correct single-bit errors and detect two-bit ones. Later, more efficient BCH (Bose, Ray-Chaudhuri and Hocquenghem) algorithms were implemented in the Flash controller [41]. However, as Flash capacity continued to rise and MLCs gave way to TLCs, more efficient ECCs were needed, resulting in a transition to soft decoding schemes. Low-density parity-check (LDPC) ECCs [42–45] are gaining traction in modern SSDs, while concatenated codes for NAND Flash have been also proposed [46–49].

Non-recoverable errors, or errors exceeding the ECC capabilities, result in *bad blocks* [50], i.e., blocks that are excluded from the memory operation and are replaced by available valid blocks. A small number of initial bad blocks is allowed, which can increase with use. Eventually, when the number of bad blocks exceeds the device specifications, the entire chip fails. It is worth mentioning that both ECC and bad blocks management require extra cells/blocks, thereby contributing to an increase of the cell effective area as discussed in Section 2.1, i.e., a reduction in the array efficiency.

Reliability specification by the manufacturers typically includes the number of bad blocks, the endurance and retention properties of the array and the number of error-correction code bits. In the following, however, we will go beyond the manufacturer specifications and look at the basic physical mechanisms at the heart of the reliability issues. Since this is an extremely large field, we restrict our discussion to issues that develop with memory usage, neglecting for example defect-related issues, time-zero problems such as V_T placement, cell-to-cell interference and many others, for which we refer the reader to other overviews of Flash memory reliability that are available in the literature, leaning either toward a device/technology [27,51–59] or user [60–62] viewpoint.

3.1. Endurance

This parameter specifies the maximum number of program/erase (P/E) cycles N_C that the memory chip can tolerate. The onset of this limitation can be observed by repeating simple, single-pulse P/E operations on a cell and monitoring the resulting V_T. An example of such a dependence in a NAND device is shown in Figure 3 [63]: the V_T levels experience an increase for high values of N_C, eventually leading to a read failure. Higher increase is usually observed on the erased level.

Like most reliability concerns for Flash memories, this phenomenon is one consequence of the electron flow through the oxide during FN tunneling, which has been observed to generate oxide traps and fixed charge in MOS (Metal-Oxide-Semiconductor) devices since the 1980s [64–69]. At first glance, such charges should result in a constant shift of both V_T levels, while the actual behavior features a milder dependence on N_C by the programmed V_T with respect to the erased one, pointing toward an interplay among different mechanisms. This picture is confirmed by early results on different non-volatile memory technologies, where different P/E schemes (i.e., different electron injection conditions into the oxide) resulted in different V_T dependences on N_C [70].
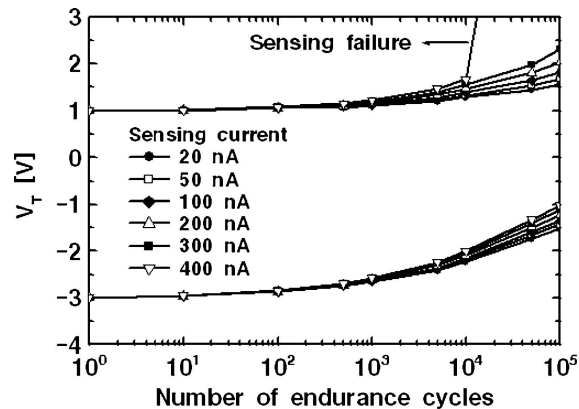
Figure 3. Endurance characteristics of a NAND cell [63], © 2004 IEEE.

A thorough analysis of the phenomenon was first conducted in [71], using charge-pumping and V_T measurements to assess the roles of interface and oxide charges, and was later corroborated by subsequent studies [72]. The authors' conclusion was that trapped oxide charge should not influence the programmed V_T value because tunnel injection from the channel during the program pulse is controlled by the oxide field at the cathode interface, meaning that the final condition reached by the cell will be of equal field, irrespective of the starting V_T and trapped charge (if not too close to the cathode interface). In turn, the same oxide field at the channel interface means the same final V_T, explaining the experimental behavior of a nearly constant programmed V_T level. Simply put, as trapped negative oxide charge builds up, it reduces injection, giving a negative V_T shift that is balanced by the positive V_T shift given by the charge itself. This "converging" behavior of the FN tunneling operation has been in fact exploited to tighten the V_T distribution after erase [73].

The same reasoning holds for the erase operation, but now the final condition is the one with the same field at the oxide-floating gate interface (the cathode during erase), which does not translate into the same V_T, now dictated by the oxide trapped charge: a larger increase in V_T vs. N_C is then expected. This model was then refined [63] to include the effect of charged stress-generated interface states, that provide an additional positive V_T shift for both levels (while not influencing injection) for high cycling conditions, explaining the slight increase in the programmed V_T. Finally, the effect of non-uniform tunneling current flow and oxide charge distribution over the area of scaled NAND devices have been investigated in [74,75], pointing out their effect on the endurance curve and on the assessment of the endurance degradation with scaling. A recent technique for separating the endurance degradation components, including subthreshold slope and transconductance, which also influence the V_T values, has been reported in [76].

From the product viewpoint, the V_T window degradation is managed by the program and verify ISPP algorithm, meaning that any variation in the P/E characteristics of the cell will result in a change in the number of ISPP pulses. Eventually, when—say—the number of erase pulses will become higher than the maximum allowed by the erase time specifications, an erase failure will be signaled. This is an unrecoverable error, resulting in the block being marked as *bad* (overprovisioning ensures that failed blocks do not reduce the device capacity). When the number of bad blocks exceeds the device specifications, the entire chip fails. Several techniques of *wear leveling* [77] are adopted by current NAND suppliers to distribute the number of P/E cycles among the blocks and make degradation uniform. Typical endurance values for SLC cells are in the 10^5 P/E cycles range, which drops by an order of magnitude or even more when moving to MLC [62]. An even lower value (around 1000) can be expected for TLCs. For SSDs, this specification translates into the total amount of writable data in—say—five years for a given workload.

3.2. Random Telegraph Noise (RTN)

3.2.1. RTN Amplitude

Random telegraph noise (RTN) in MOS transistors is an abrupt two-level fluctuation in drain current I_D following trapping and release of a single electron by an oxide defect. This well-known phenomenon has been observed and studied since the 1980s [78–86], resulting in a prediction for the RTN amplitude [79] equal to

$$\frac{\Delta I_D}{I_D} = \frac{q}{C_{ox}WL(V_G - V_T)} + \frac{\Delta\mu_n}{\mu_n}\frac{\Delta A}{WL}, \tag{2}$$

where q is the electron charge, C_{ox} the oxide capacitance per unit area, W and L the channel width and length, respectively, and μ_n the electron mobility in the inversion channel, which experiences a fluctuation $\Delta\mu_n$ over an area ΔA as a consequence of the additional scattering with the trapped electron. Typical values for the first term are smaller than 10%, while the scattering part is typically negligible in the subthreshold regime, where NAND cells are usually read; therefore, RTN has long been regarded as a curious scientific phenomenon not constituting a noticeable reliability threat. This held true even if unusually large values for the RTN amplitude had been consistently reported on some devices, together with complex time behaviors [79,82,87–89], that seemed to defy Equation (2). When Flash memories made a large quantity of data readily available, making it possible to pinpoint and study anomalous behaviors occurring with extremely low probability, larger and larger fluctuation amplitudes emerged [90,91], and RTN became a serious reliability constraint in static random-access [92–94] and Flash memories [90,91,95–99], prompting an intensive research effort to understand its root cause and impact on memory array operation.

An example of the relevance of RTN is shown in Figure 4, where results obtained on a selected decananometer Flash cell are reported [91]: current fluctuations up to 60% were detected, which cannot be explained with the above theory. Though early explanations pointed to multiple carrier emissions [88,89], change in defect properties [87] or interacting quantum wells [82], a different approach based on the percolative nature of the current conduction around the charged trap site had been mentioned in [100] and thoroughly revisited in [101], exploiting a concept applied to amorphous semiconductors [102]. The authors pointed out that no uniform conduction in a MOSFET (Metal-Oxide-Semiconductor Field-Effect Transistor) channel could account for the anomalous RTN amplitudes. Rather, a non-uniform conduction with filamentary current flow was needed as a starting point; therefore, an RTN trap could effectively modulate a current-crowding path, resulting in a large current fluctuation. Fixed oxide and interface charges were proposed as sources of percolation, following earlier studies [103,104]. This is now regarded as the correct interpretation of the RTN amplitudes, with the caution to add the actual main source of percolation: random dopant fluctuation (RDF). In fact, statistical variation in number and position of dopants was known to affect MOSFET parameters such as V_T and subthreshold slope [105–108], and was shown to have a significant impact on RTN in [109] and in [110,111], where large amplitude values compatible with experimental data were reported.

For the evaluation of the RTN effects in Flash memories, the statistical distribution of its amplitude is needed, accounting for realistic device geometry and doping. This task was undertaken in [112,113], where 3D Monte Carlo simulations of electron transport were carried out, accounting for randomized doping and trap positions. Simulations were conducted within the drift-diffusion framework, neglecting the field dependence of mobility to avoid unphysical effects close to the ionized dopants and RTN trap. The impact of Coulomb scattering on the RTN and variability characteristic of such MOSFETs has been studied in [114–117], showing that those effects are barely relevant for NAND Flash, that are read in the subthreshold regime. Charge quantization was instead accounted for via the density-gradient correction to the drift-diffusion formalism [118,119]. Its relevance from the variability viewpoint was discussed in [120].
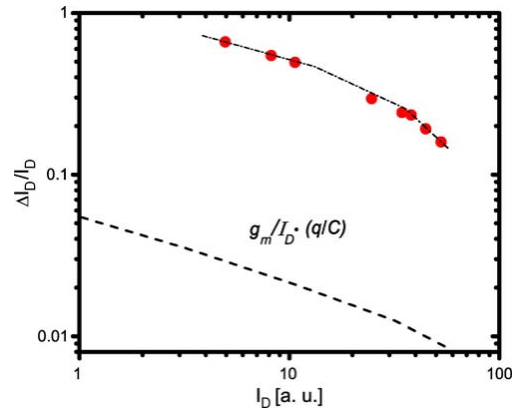
Figure 4. RTN fluctuation amplitude measured on a decananometer Flash cell (points = experimental data, dashed line = predictions from (2) when mobility modulations are neglected) [91], © 2007 IEEE.

Figure 5 shows results for the current density profile in the channel of a decananometer Flash memory biased around a threshold. The two figures on the left-hand side correspond to a case of high V_T shift: transport in the empty trap case (upper figure) has a percolative nature because of RDF and field intensification around the active area corners. The RTN trap (black square in the lower figure) is located right above one main percolation path, close to one edge: its charging will shut off such path, leading to a large variation in current and V_T. On the other hand, if the trap is located over a region in which the current density is already negligible, its charging and discharging will barely affect the overall current flow, as shown by the figures on the right-hand side.
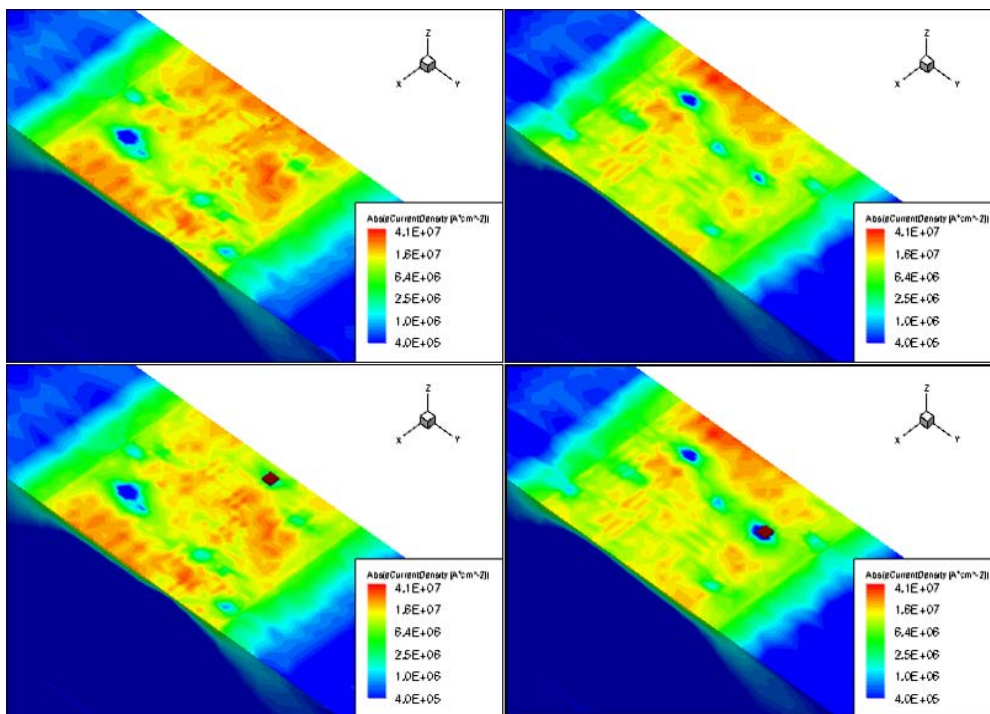

Figure 5. Current density profiles in the channel of decananometer Flash memories with and without a charged RTN trap. Left-hand side = high V_T shift case, right-hand side = low V_T shift case. Upper figures represent the empty trap case, lower figures the charged trap ones. The trap position is marked by the black squares in the lower figures.

Monte Carlo results for the statistical distribution of the V_T shift due to a single RTN trap, ΔV_T^1, are shown in Figure 6: an exponential distribution was found, which can be easily characterized by its slope λ, usually expressed in mV/decade. A thorough parametric analysis of λ was carried out in [113,121], resulting in the following dependence:

$$\lambda = \frac{K}{\alpha_G}\frac{t_{ox}^{\alpha}\sqrt{N_A}}{W\sqrt{L}},\tag{3}$$

where t_{ox} is the tunnel oxide thickness, N_A the average substrate doping, α_G the control-gate coupling coefficient and K is a constant dependent on device design, doping profile and so on. This relation has proven itself to reproduce quite successfully the scaling dependence of the RTN slope over several different technologies.
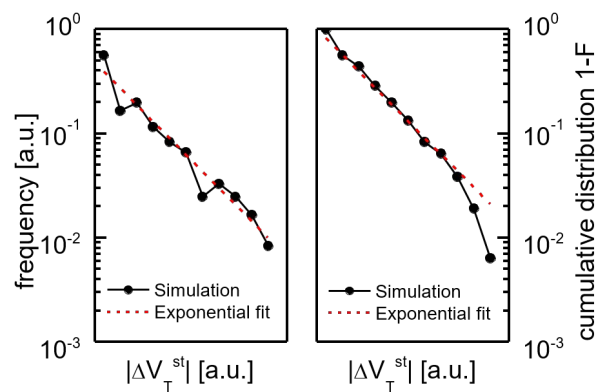

Figure 6. Probability density function (**left**) and cumulative distribution (**right**) for the threshold voltage shift determined by the charging of a single interfacial trap, ΔV_T^1, in a decananometer NAND Flash cell [112], © 2008 IEEE.

3.2.2. Main Experimental Data

The first experimental data on the statistical distribution of RTN amplitude in Flash memory arrays were published in 2006 [95]. For each cell in the array, the authors applied a staircase control-gate bias, recording the V_T value at which the cell switched from on to off state. Because of RTN, different V_Ts are recorded every time the experiment is repeated, and their maximum and minimum values will give the amplitude of the RTN fluctuation (two staircases are actually employed, with increasing and decreasing control-gate bias, but the working principle does not change). Results for the fluctuation amplitude on a 90 nm cell are reported in Figure 7: note that the majority of cells exhibit negligible V_T fluctuations, but a small number of them can reach V_T shifts of a few hundred mV. Moreover, the number of cells exhibiting this behavior increases with P/E cycling.
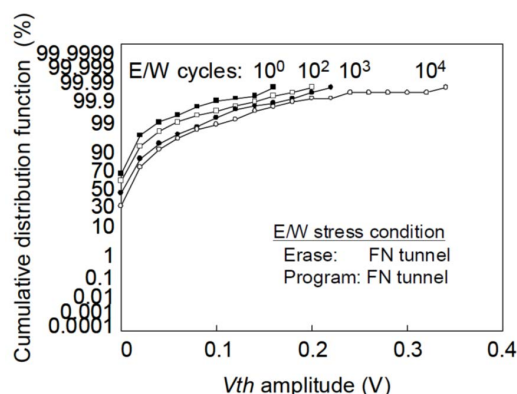
Figure 7. Experimental cumulative distribution function for the V_T shift induced by RTN in a Flash array [95], © 2007 IEEE.

A simpler characterization technique for RTN is reported in Figure 8 [96], where the V_T values at two different times are recorded for each cell. The RTN amplitude is now defined as the V_T difference between the two values. Note that this technique does not measure the *maximum* RTN fluctuation amplitude, as accomplished by [90,95,97,99], because a noisy cell can give either positive, negative or zero fluctuation, depending upon the sampling time (see the different cases Figure 8). However, this technique will return the actual RTN amplitude distribution that matters under real operating conditions, when a cell is read after being programmed.
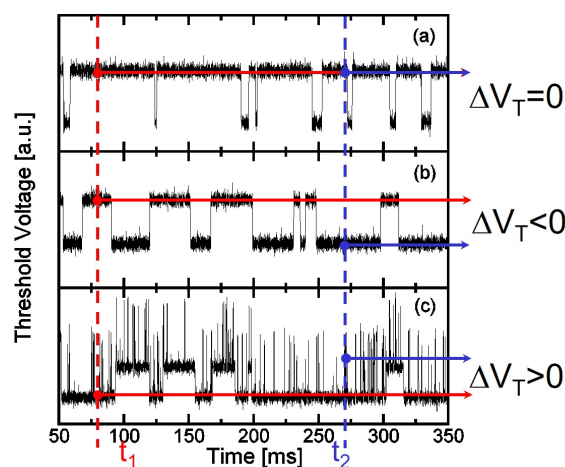

Figure 8. Time dependence of V_T for three Flash cells affected by RTN. V_T is sampled at times t_1 and t_2 and the RTN amplitude is defined as the V_T difference. Note that (**a**) null, (**b**) negative and (**c**) positive ΔV_T can result.

Typical results obtained with this technique are shown in Figure 9 for NOR (left) and NAND (right) Flash arrays [96,122,123]. Figures show the cumulative density function (cdf) and its complement, highlighting clear exponential tails down to very low probability values [124] for both high and low ΔV_T values. The region around $\Delta V_T = 0$ is affected by Gaussian noise, and the exponential tails depart from such region at different probabilities. Note also that when ΔV_T is evaluated between the *n*th and the first read operation, the tails slightly increase with *n*. Differences between NOR and NAND RTN levels and amplitude are determined by cell design parameters (see Equation (3)) as well as by the cycling conditions.
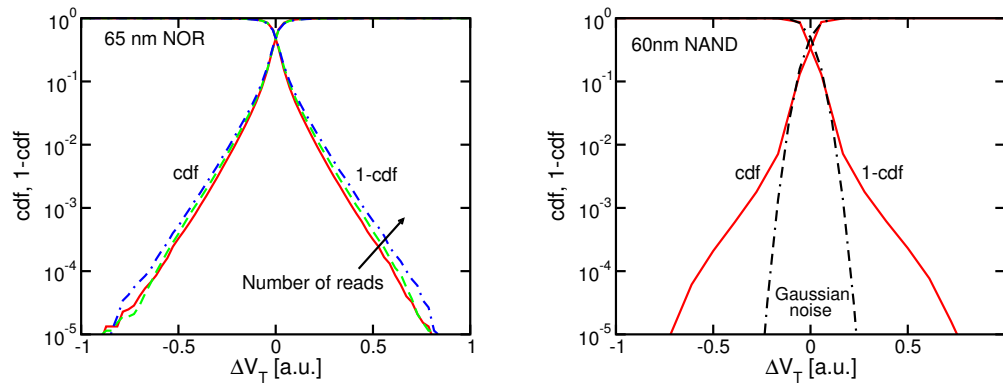
Figure 9. Experimental results for the ΔV_T statistical distribution on a 65 nm NOR (**left**) and a 60 nm NAND (**right**) Flash array, obtained following the procedure highlighted in Figure 8. Both the cumulative distribution function (cdf) and its complement are displayed, to highlight the exponential tails. Results for different numbers of reads are also shown for the NOR case.

Figure 7 shows that RTN depends on P/E cycling. An increase in the tail height was also reported in [98,123,125], where different technologies were compared. Figure 10 (left) shows experimental data for RTN in a NAND array for different cycling conditions: note that the exponential tails grow while their slope is hardly affected. This is reported in the right-hand side of Figure 10 where the slope λ of the exponential tails and the probability at a fixed $|\Delta V_T|$ value are shown as a function of N_C, for the same technology. Note the negligible dependence of λ on N_C, confirming that cycling results in a parallel shift of the RTN tails. The increase in tail height follows a power-law dependence with exponent 0.29 for this technology [123]. Different values were found in NOR arrays, owing once more to differences in cell design and cycling conditions. Such dependences prompted several works and studies on the optimization of the cell process/architecture in order to reduce the extent of the RTN [124,126–134]. In the NAND array, moreover, the RTN amplitude was shown to depend on the cell position and state along the string [135,136] (because of the different transconductances, depending on the source and drain series resistances) and on the state of cells on adjacent BLs [137] (because of the modification in the electron density profile induced by electrostatic interference).
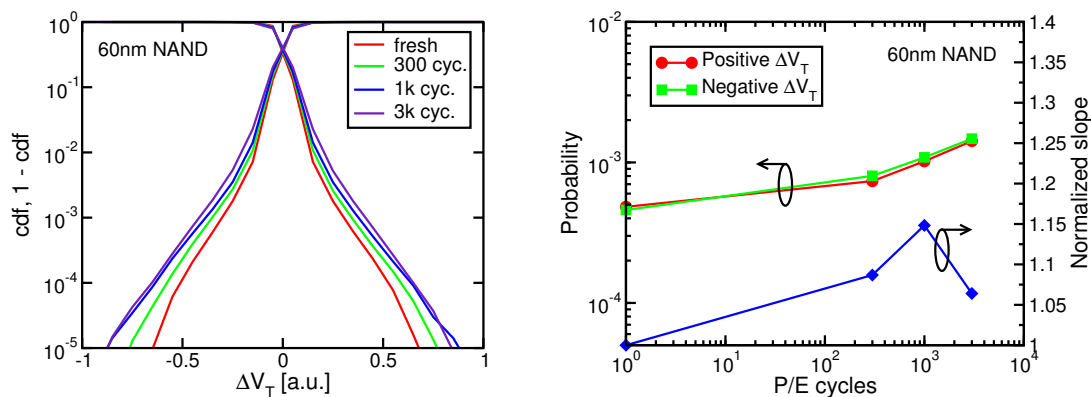

Figure 10. Experimental data for the RTN distribution in NAND arrays as a function of N_C (**left**) and RTN tail height at a fixed $|\Delta V_T|$ value and slope as a function of N_C for a 60 nm NAND (**right**) Flash array [123], © 2008 IEEE.

The cycling dependence was also investigated in [138], in connection with the temperature dependence of RTN. Two cycling schemes were adopted, involving different temperatures and cycling times in accordance with the activation energy of damage recovery which will be discussed in Section 3.4. Then, RTN was read at room temperature. Figure 11 (left) shows that no significant

difference can be detected in the RTN cdfs, meaning that the high-temperature phase applied during the 500 h scheme is not annealing any defect responsible for RTN. The right-hand side of Figure 11 shows instead the dependence upon the read temperature. Again, no dependence appears when the temperature is changed from room temperature to 55°C. This result may seem surprising, as it is well known that trap capture and emission times are temperature-dependent [136,139,140], but comes from the distribution of such times among the cells, as discussed in the next section.
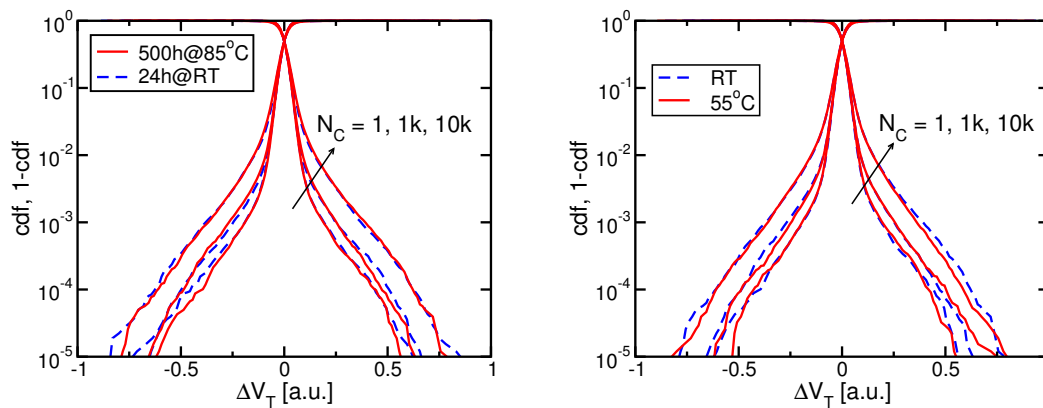

Figure 11. RTN cdfs for two cycling schemes and room-temperature read (**left**) and for one cycling scheme and different read temperatures (**right**). No significant differences appear in all cases.

3.2.3. Models

The experimental data just reported call for a consistent model able to explain the different features of the phenomenon. While data shown in Figure 7 can be reproduced by a simple exponential fit [98], the explicit time dependence captured by Figures 8 and 9 requires a time-dependent analysis, that was first put forward in [96,122]. The starting point is the exponential distribution of V_T shift extracted from simulation results described in Section 3.2.1. Such a distribution describes a positive V_T shift between time t_1 and t_2 if an empty trap at t_1 is filled at t_2 and a negative shift if the opposite is true. If the trap state does not change, a zero V_T shift holds. The final time-dependent distribution is then the one shown in Figure 12 (left), where the positive and negative exponentials and the zero-shift delta function must be weighted by their occurrence probabilities, representing their areas. This is accomplished describing the RTN process as a two-state Markov process [141,142], as depicted in Figure 12 (right). The two states correspond to the empty (low-V_T, state 0) and filled (high-V_T, state 1) trap. Considering constant capture and emission times τ_c and τ_e, the transition probabilities between state i and j, P_{ij}, over a time interval T become:

$$P_{11}(\tau_c, \tau_e) = \frac{1}{\tau_c + \tau_e} \left(\tau_c \, e^{-T/\tau_{eq}} + \tau_e \right) \tag{4a}$$

$$P_{10}(\tau_c, \tau_e) = 1 - P_{11} = \frac{\tau_c}{\tau_c + \tau_e} \left(1 - e^{-T/\tau_{eq}} \right) \tag{4b}$$

$$P_{00}(\tau_c, \tau_e) = \frac{1}{\tau_c + \tau_e} \left(\tau_e \, e^{-T/\tau_{eq}} + \tau_c \right) \tag{4c}$$

$$P_{01}(\tau_c, \tau_e) = 1 - P_{00} = \frac{\tau_e}{\tau_c + \tau_e} \left(1 - e^{-T/\tau_{eq}} \right), \tag{4d}$$

where $\tau_{eq}^{-1} = \tau_c^{-1} + \tau_e^{-1}$. Note that these expressions describe the probabilities to find the trap in a certain state after a time interval T, given the initial condition, irrespective of the number of transitions taking place in-between.
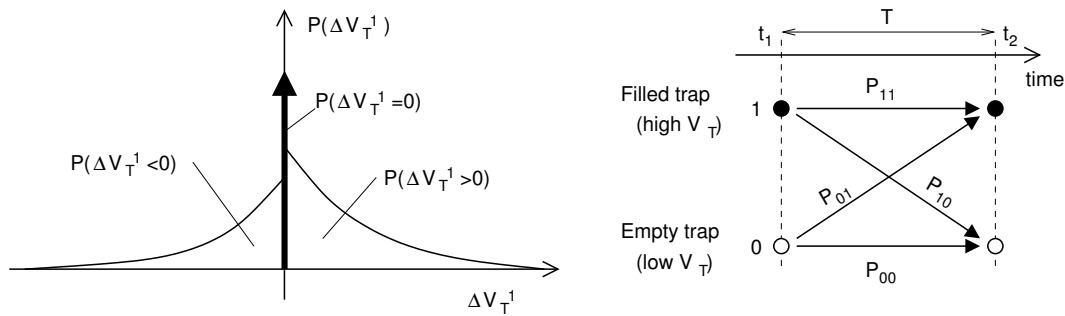
Figure 12. PDF (Probability Density Function) of the single-trap threshold voltage shift $P(\Delta V_T^1)$ (**left**) and Markov process view of the single-trap RTN waveform between times t_1 and $t_2 = t_1 + T$ (**right**).

From (4), the probability for a single trap to provide a V_T shift, ΔV_T^1, can be easily calculated:

$$P(\Delta V_T^1 > 0|\tau_c, \tau_e) = P_0^{t_1} P_{01}(\tau_c, \tau_e) \tag{5a}$$

$$P(\Delta V_T^1 < 0|\tau_c, \tau_e) = P_1^{t_1} P_{10}(\tau_c, \tau_e) \tag{5b}$$

$$P(\Delta V_T^1 = 0|\tau_c, \tau_e) = P_0^{t_1} P_{00}(\tau_c, \tau_e) + P_1^{t_1} P_{11}(\tau_c, \tau_e), \tag{5c}$$

where $P_i^{t_1}$ is the probability for the trap to be in state $i = 0, 1$ at time t_1.

Once the analysis for a given set of τ_c, τ_e is assessed, we must account for their statistical distribution among the traps. The original approach in [122] relied on the classical theory of direct tunneling into oxide defects [80,81], explicitly linking τ_c and τ_e to the space and energy position of the trap inside the oxide. Later results have, however, shown that capture/emission times in MOSFETs are not compatible with such a description [143,144], due to device variability effects on the time constants [145–147] and structural relaxation of defects [140,148], generating a large spread in time constants even for traps located close to the silicon/oxide interface and disrupting the classical correlation between τ_c and τ_e. The study of the microscopic properties of RTN traps is still an active research topic [149–156], but for our purposes we will adopt a pragmatic approach, assuming given distributions of capture/emission time constants that may be consistent with observations, without linking them to any particular trap location. A constant PDF on a log-time scale was successfully assumed in [138], such as the one that will be used to describe detrapping after cycling in Section 3.4.2. It is worth noting that the constant trap density adopted in [122] implicitly provides a similar time constant distribution, due to the exponential nature of the tunneling probability, the only difference being the lack of correlation between τ_c and τ_e and between the time constant and the spatial location of traps. Within this framework, the expression for the single-trap V_T shift pdf, $P(\Delta V_T^1)$, becomes

$$P(\Delta V_T^1) = \int \int P(\Delta V_T^1|\tau_c, \tau_e) p(\tau_c, \tau_e) d\tau_c d\tau_e, \tag{6}$$

where $p(\tau_c, \tau_e)$ is the probability for a trap to possess such capture/emission time constants.

Once $P(\Delta V_T^1)$ is computed, multiple trap effects can be easily accounted for assuming statistical independence of their V_T shifts. This means that the overall V_T shift is the sum of the individual contributions, which translates into a convolution operation for pdfs. In particular, for N_t independent traps, each with the same $P(\Delta V_T^1)$ pdf, we have

$$P(\Delta V_T^{N_t}) = \bigotimes_1^{N_t} P(\Delta V_T^1) \tag{7}$$

and, since the number of traps per cell can be assumed to be statistically distributed according to a Poisson statistics $p(N_t)$, we finally get the RTN pdf:

$$P(\Delta V_T) = \sum_{N_t=1}^{\infty} P(\Delta V_T^{N_t}) p(N_t). \tag{8}$$

Typical comparisons between model results and experimental data are reported in Figure 13 for NAND [138] and NOR [122] arrays. The left-hand side shows that a nice agreement can be achieved for different technologies and cycling conditions. In particular, because of the exponential nature of $P(\Delta V_T^1)$, its slope λ matches what is obtained from array data and resulting from the convolution operation in (8). The right-hand size is instead related to data reported in Figure 9 (left), where RTN distributions for increasing read operations (i.e., time) are reported, and shows the probabilities at fixed $|\Delta V_T|$ values as a function of the number of read operations. The increase in the RTN figure is well reproduced by the model and can now be explained: the increase in the elapsed time between the read operations (i.e., T in Equation (4)) allows for more and more traps to be activated, as their τ_{eq} becomes comparable with T. This increases the randomness of V_T, enlarging the RTN distribution. This effect was nicely shown in [96,122] as an extension with time of the oxide region whose traps are sampled by RTN.
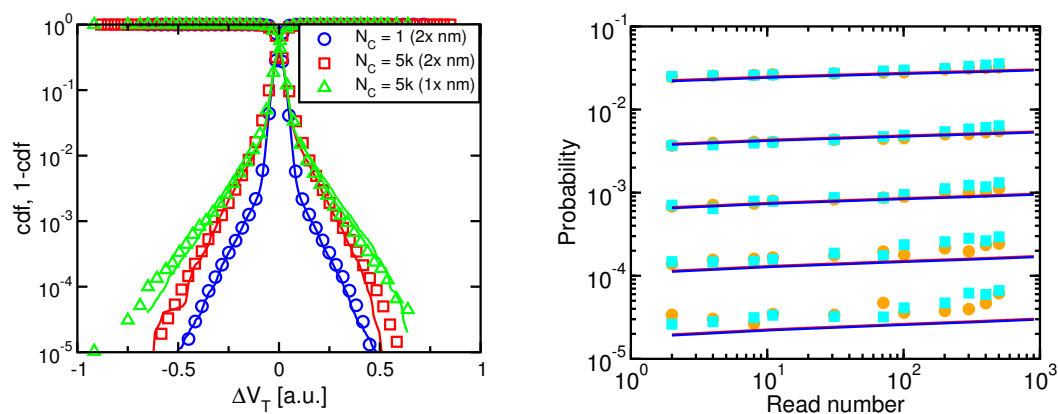

Figure 13. (**Left**) Comparison between experimental data and model for the RTN distribution in $1\times$ and $2\times$ NAND arrays and different N_C; (**Right**) same comparison for the tail probability at a fixed $|\Delta V_T|$ as a function of the number of reads for a 65 nm NOR Flash array.

Finally, the temperature-independent RTN behavior observed in [138] was explained in terms of a lack in correlation between the time constants and their activation energy values; therefore, temperature will not affect the overall time constant distribution. Of course, different defects will contribute to the RTN fluctuations, as some are accelerated and will end up being filtered by the read time while slow ones will come into play as a result of thermal activation.

3.2.4. Effect on Programmed V_T Distribution

Figure 14 shows pictorially what the effect of RTN on the V_T levels is and how it affects the array reliability, for the case of an MLC array. Exponential tails due to RTN extend from the ideal Gaussian distribution, their magnitude increasing with N_C and time. This results in a reduced separation between adjacent V_T levels (the so-called *window*), compromising the accuracy of the read operation and leading eventually to a read failure: the dashed lines in Figure 14 indicate the reference V_T levels used during read, to which the cell V_T is compared. When a distribution crosses a reference level, a read error is bound to occur.
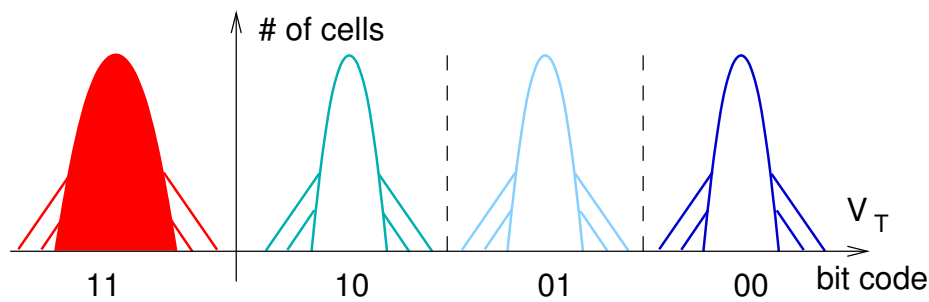
Figure 14. V_T distribution of an MLC array in the presence of RTN. Dashed lines represent the reference levels used during read.

To investigate this issue in more detail, Figure 15 shows the effects of RTN on the programmed V_T distribution [157] following ISPP and successive read. The left-hand side shows the V_T pdf right at the end of the ISPP algorithm, for two values of the ISPP step: V_{s1} and $V_{s2} > V_{s1}$. In the ideal case, the algorithm is expected to result in a V_T distribution exceeding the program-verify level PV by an amount at most equal to the ISPP step (dashed lines in Figure 15). In reality, the distributions are widened by the program noise (not discussed here, see [158,159]) and by RTN and display the characteristic exponential tails. Note that no cell can have V_T values below the program-verify level, as the algorithm would apply an additional step to them. However, a negative RTN tail also arises as soon as the cells are read (right-hand side of Figure 15). This figure also shows that the RTN tail height is smaller for larger values of the ISPP step; however, increasing its value will lead to an overall larger V_T distribution. On the other hand, reducing the ISPP step is effective—apart from the increase in program time—only until its value becomes comparable with the RTN distribution width.
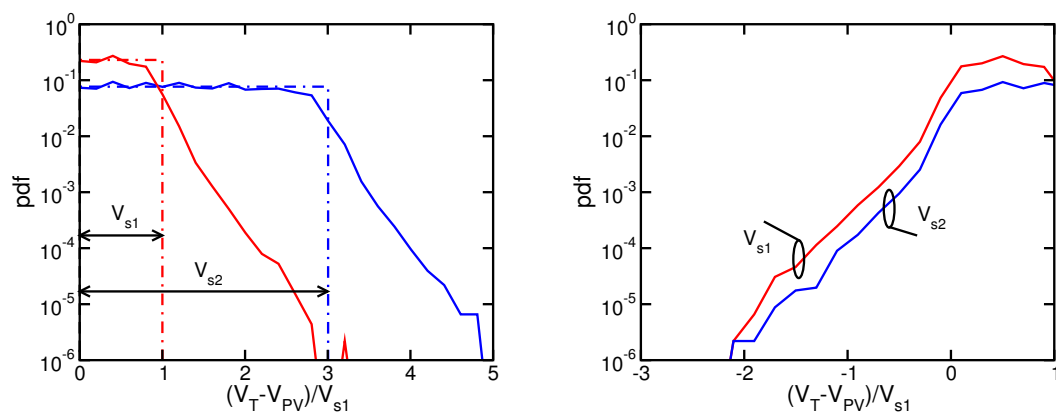

Figure 15. V_T pdf at the end of the ISPP operation (**left**) and after a successive read operation (**right**). Results are shown for two values of the ISPP step, V_{s1} and $V_{s2} > V_{s1}$.

3.3. Retention after Cycling and SILC

Data retention specifies the maximum period of time that allows a correct retrieval of stored data, and is determined by the faint but nonzero leakage of electrons from the floating gate. Early experimenters ascribed such leakage to electron tunneling through the dielectric and realized that temperature [160] or bias [161] could be used to accelerate the phenomenon, while devising appropriate design criteria [162–164]. In particular, it was found that a ten-year retention requirement could be satisfied with a minimum tunnel oxide thickness around 3.5–4.5 nm [56,165], much thinner than the 7–8 nm usually adopted. However, the oxide thickness has barely scaled over the many Flash technology generations, remaining far away from its theoretical limit, because it was soon realized [70,166,167] that repeated P/E cycles led to degradation of the tunnel oxide characteristics, enhanced leakage and worse retention. The enhanced oxide leakage at low fields after electrical stress

is known as *stress-induced leakage current*, or SILC [168,169]. SILC has been largely studied on capacitors and described in terms of a trap-assisted tunneling (TAT) process through oxide traps [170–179] (see Figure 17a for a pictorial view of the process), in agreement with experimental evidence [180,181]. In the realm of non-volatile memories, SILC can lead to charge loss/gain from/to the floating gate after stress and to retention errors. To keep this contribution negligible, the oxide thickness of Flash cells has barely scaled with the technology generations, going from about 10 nm in the first prototype [3] to about 7–8 nm in the latest nodes [182]. However, experimental data [183–188] have demonstrated that, even if the SILC is reduced, there is a small number of cells that can exhibit a high leakage after stress, in analogy with the statistical behavior of RTN. An example of this effect is reported in Figure 16, where thin-oxide (6.5 nm) NOR Flash arrays were cycled heavily (10^4 P/E cycles) before being subjected to positive (left) or negative (right) oxide field to induce a V_T shift [189]: note the distribution tails made up of cells featuring an enhanced leakage with respect to the main part of the distribution, whose shift is due to the intrinsic FN tunnelig current.

This enhanced leakage, called *anomalous SILC*, has been also identified as a trap-related phenomenon, because of its low activation energy [188,189] and fluctuations resulting from capture/emission processes [185,190]. However, a two-trap assisted tunneling process (2TAT) has now been invoked to explain the experimental data of individual cells [191,192], as depicted in Figure 17c. From this view, a picture of the SILC distribution in a Flash array can be built, assuming a statistical distribution of stress-generated oxide defects (Figure 17b): those cells that happen to have two traps close enough to allow significant tunneling will be affected by anomalous SILC [193,194], while the majority will only experience SILC or the FN tunneling current. Figure 18 shows an example of gate leakage current of two cells affected by anomalous SILC compared with the 2TAT model predictions (left), and (right) the statistical distribution of SILC at a given control-gate bias for different oxide thicknesses, again compared with results of the statistical model. From there, reliability extrapolations can be performed to estimate the number of failures. Other extrapolations procedures have been proposed, based on the same model or on empirical approximations [195–203], while addressing the optimization of the P/E waveforms and cell design [204–207].

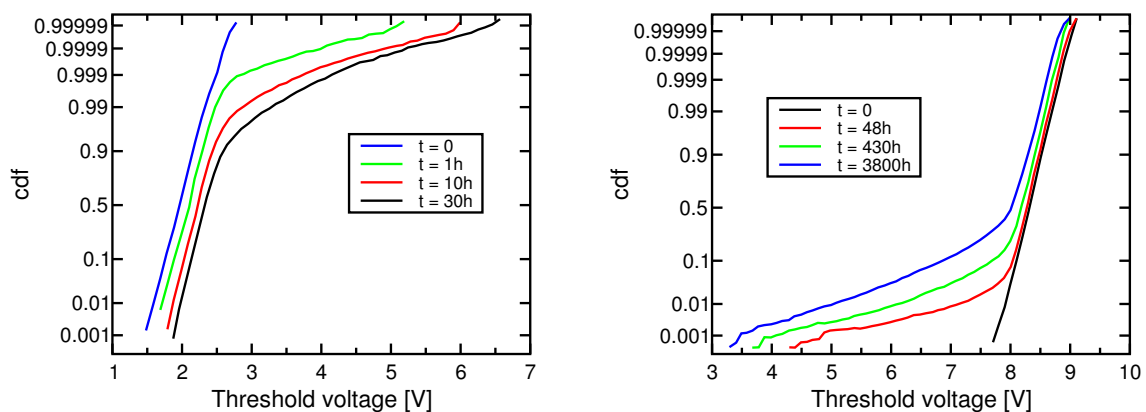
Figure 16. V_T cdf over time for 6.5 nm NOR Flash cells arrays after cycling. Positive control-gate bias (5.5 V) was applied in the left-hand side figure, while the right-hand side shows results of a retention experiment [189], © 2001 IEEE.
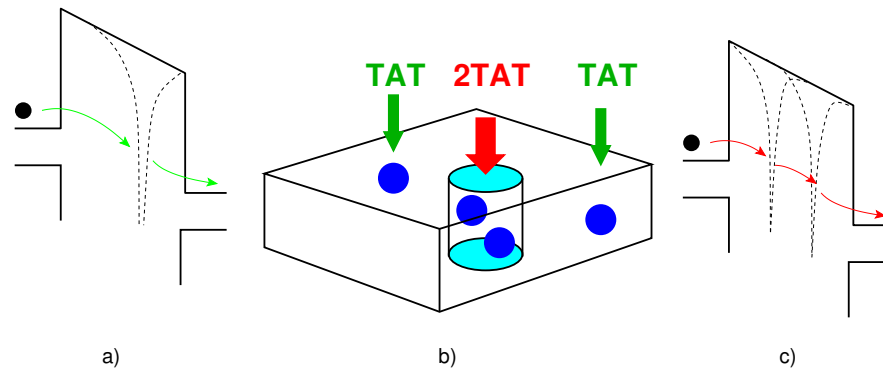
Figure 17. Pictorial view of the SILC distribution in Flash cells. The center picture (**b**) shows a trap distribution in the tunnel oxide region of the cell. One- and two-trap-assisted tunneling mechanisms (respectively TAT and 2TAT) are depicted as (**a**,**c**), respectively.
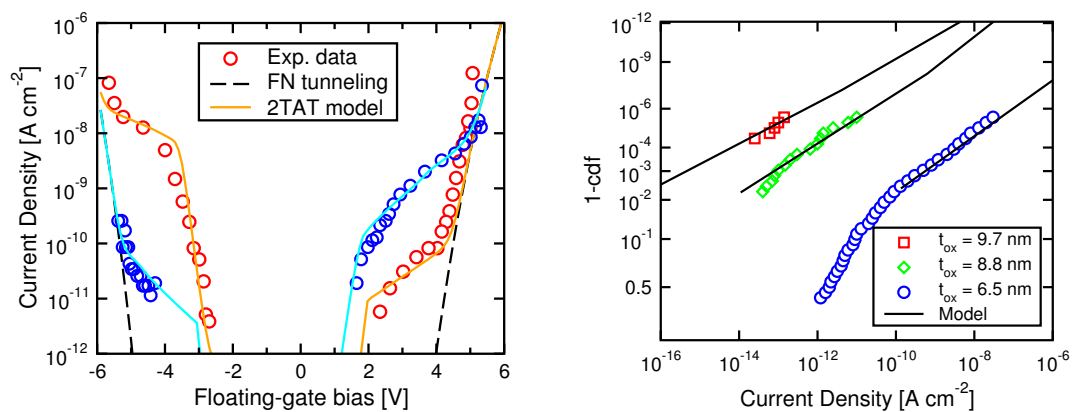

Figure 18. (**Left**) average gate current density vs. floating-gate bias for two cells affected by anomalous SILC, together with model predictions; (**Right**) cdf of the array leakage currents at fixed bias for different gate oxide thicknesses and model results. In all cases, symbols = experimental data, lines = model results.

Like all reliability issues, SILC becomes more of a concern in scaled MLCs or TLCs, because of the reduced separation between V_T levels and smaller floating-gate capacitance, meaning that fewer electrons are stored into the floating gate and that even microscopic leakage currents can cause significant damage. In today's technologies, SILC remains dominant in tail cells [208], that are usually covered by ECC, while its importance as a reliability issue is somewhat decreased with respect to phenomena affecting the entire distribution, such as the next one [209].

3.4. Charge Detrapping

Early investigators of SILC in stressed oxides observed a transient component in the leakage current, that was attributed to charging and discharging of oxide traps [210–214] and was exploited to extract information on the oxide quality [215,216]. The detrapping current followed a $1/t$ dependence down to very short times [217], in agreement with tunneling models [218]. The impact of this phenomenon on Flash reliability was first discussed in [219,220], pointing out that the log-time dependence of the V_T shift could lead to retention errors even in the absence of SILC and could dominate the charge loss with respect to SILC itself [63]. Moreover, as charge is trapped into the oxide during P/E cycling and detrapped in-between, a dependence on the cycling pattern arises, which complicates the development of comprehensive models. Such task began in [221,222] and developed into a full model accounting for the major experimental evidence in [223–225].

3.4.1. Main Experimental Data

Figure 19 shows experimental data from a retention experiment at room temperature on a 16 nm NAND vehicle after 10^4 P/E cycles [226]. Note the difference with respect to Figure 16 (right): no specific tail is detected here, but rather the entire distribution is affected, although different charge loss rates can be detected between the main part and the lower extreme of the distribution (see right-hand side of Figure 19). Note also the logarithmic dependence on time of ΔV_T, which can be fully characterized by its slope α.
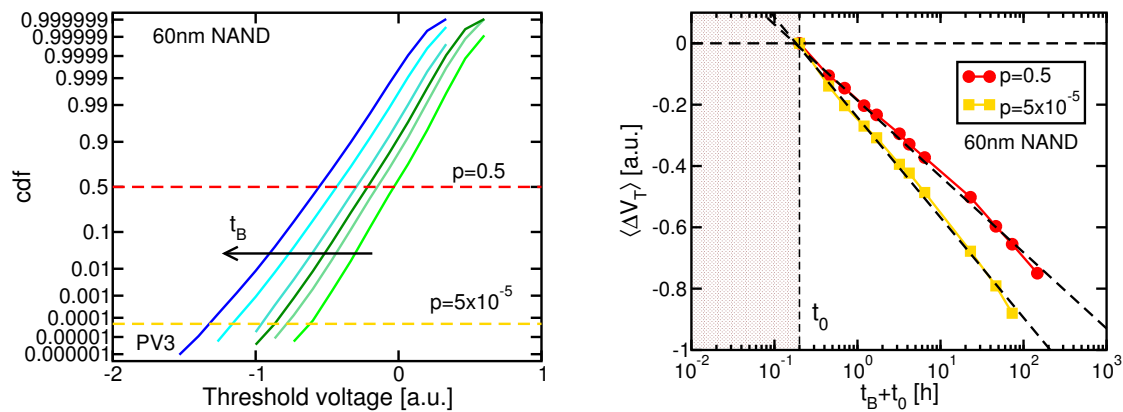

Figure 19. (**Left**) V_T cdf over time for 60 nm NAND Flash cell arrays after cycling; (**Right**) V_T shift vs. time extracted from the cdfs at fixed probability levels [226], © 2011 IEEE.

The dependence of charge detrapping on the cycling conditions is reported in Figure 20: its left-hand side shows that ΔV_T increases with N_C, as more electrons are trapped during P/E operations and emitted afterwards. The right-hand side reports the dependence of α on N_C: a square-root law was found as soon as the cycling-induced damage overcomes the native detrapping, in agreement with observations on RTN-related defects [138]. This dependence is also independent of the programmed level and temperature (note that the curves for different programmed levels have been vertically shifted to highlight the square-root law and avoid excessive overlap). Figure 21 shows instead the dependence on the bake conditions: note that the detrapping dynamics depend on the time t_0 elapsing between the last program and the first read operation (left-hand side), resulting in a shift of the detrapping curve along the log-time axis by a quantity exactly equal to t_0. Also, ΔV_T depends upon the bake temperature T_B (right-hand side of Figure 21), demonstrating that detrapping is thermally activated: the ΔV_T curves at different T_B are shifted according to an Arrhenius law with activation energy $E_A \approx 1.1$ eV [221,222,228,229], a single detrapping curve can be obtained for an equivalent T_B. This value of activation energy had been also observed in earlier retention experiments [51]. Careful account of additional effects which may compromise a reliable extraction of E_A, such as interface-state and mobility recovery [230,231] or disturbs [232], is required.
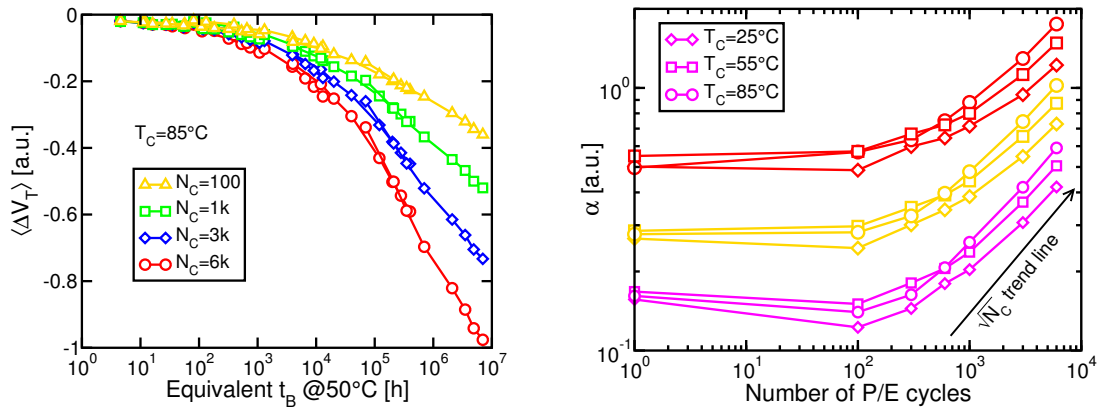
Figure 20. (**Left**) average ΔV_T transients measured on 16 nm NAND Flash arrays for increasing number of P/E cycles [227], © 2016 IEEE; (**Right**) dependence of the slope α on N_C for different programmed levels (colors) and temperatures (symbols). Curves for different programmed levels have been vertically shifted for better clarity.
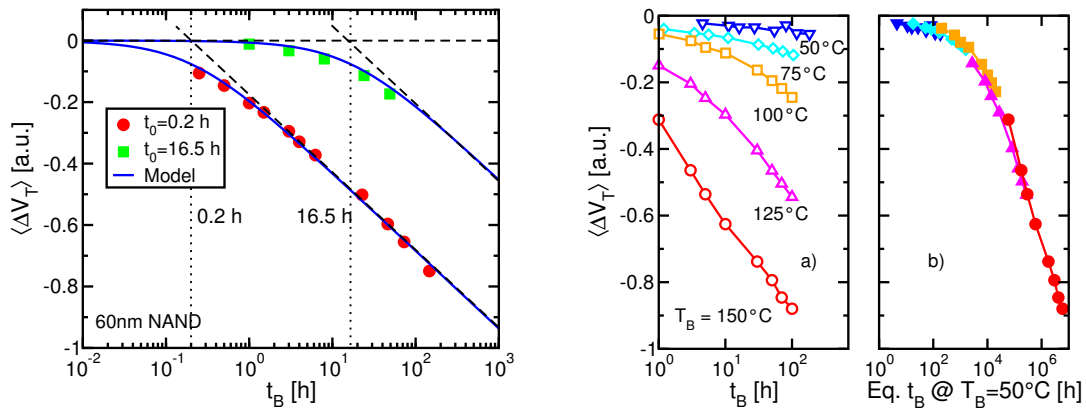

Figure 21. Dependence of ΔV_T on bake conditions, namely time t_0 [226], © 2011 IEEE (**left**) and bake temperature T_B (**right**). Data are referred to 60 and 16 nm NAND Flash vehicles. Data on the right-hand side show a) experimental data for different T_B and b) the same data when t_B is shifted on log scale according to E_A.

The last data to be discussed relate to the cycling conditions dependence. Figure 22 shows ΔV_T for different values of the cycling time t_C (left) and temperature T_C (right). Note that a smaller ΔV_T is detected if cycling is distributed over a longer period of time or performed at higher temperatures. This is consistent with a picture that assumes electron detrapping between one cycle and the next one, reducing the number of trapped electrons at the end of the last cycle that are then lost during the bake phase. Coherently with this interpretation, the activation energy for cycling was found to be the same as for bake. These findings have important implications for the device reliability, that will be discussed after reviewing the quantitative understanding of the phenomenon together with other experimental data.
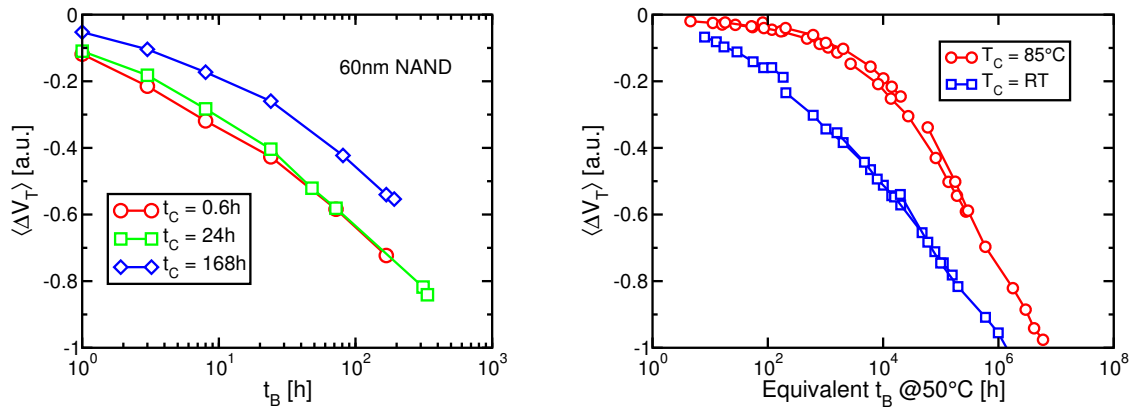
Figure 22. Dependence of ΔV_T on cycling conditions, namely total cycling time t_C (**left**) and cycling temperature T_C [227], © 2016 IEEE (**right**).

3.4.2. Models

We start this Section with the description of a compact model [221,222,226] that captures the main features of the data and can be used for some first-order extrapolations, while moving to more refined descriptions of the underlying physics later on. To this aim, we assume a log-time dependence of the V_T shift that, considering that the first read operation is performed at time t_0 after the end of the cycling phase, leads to

$$\Delta V_T(t_B) = V_T(t_0 + t_B) - V_T(t_0) = -\alpha \ln(t_0 + t_B) + \alpha \ln(t_0) = -\alpha \ln\left(1 + \frac{t_B}{t_0}\right), \qquad (9)$$

which is what we observe. The dependences of ΔV_T on N_C and probability level (see Figure 19) can then be simply inserted into the expression of α. To account for the thermal activation of detrapping during bake, which results in a rigid shift of the detrapping curve in a semilog plot (see right-hand side of Figure 21), we can now modify the expression of the term t_0, which is the zero-loss intercept of the log-time V_T curve (see Figure 21, left). This results in:

$$t_0^* = t_0\, e^{E_A(1/k_B T_B - 1/k_B T_{ref})}; \quad \Delta V_T(t_B) = -\alpha \ln\left(1 + \frac{t_B}{t_0^*}\right), \qquad (10)$$

where t_0 is now the initial wait time at the reference temperature T_{ref} and k_B is the Boltzmann constant. Following this line of thought, even data in Figure 22 can be represented by a dependence of the intercept t_0^* on cycling conditions. To do this, it is assumed that damage creation during P/E cycles and recovery between them can be described as a fast-cycling experiment (i.e., with negligible delay between cycles) generating the whole oxide damage, followed by a single recovery phase of a certain duration. Combined with the thermal activation term, this results in the final expression

$$t_B^* = t_0^* + At_c\, e^{E_A(1/k_B T_B - 1/k_B T_C)}; \quad \Delta V_T(t_B) = -\alpha \ln\left(1 + \frac{t_B}{t_B^*}\right), \qquad (11)$$

which allows to completely describe the ΔV_T dynamics as a function of the various experimental parameters once the fitting parameter A is selected. This means that the term $\ln(1 + t_B/t_B^*)$, called the *universal damage metric* or UDM, should now account for all ΔV_T transients measured on a specific technology, irrespective of cycling and bake conditions. This is demonstrated in Figure 23 (left). Figure 23 (right) shows instead the value of t_B^* for different cycling conditions.
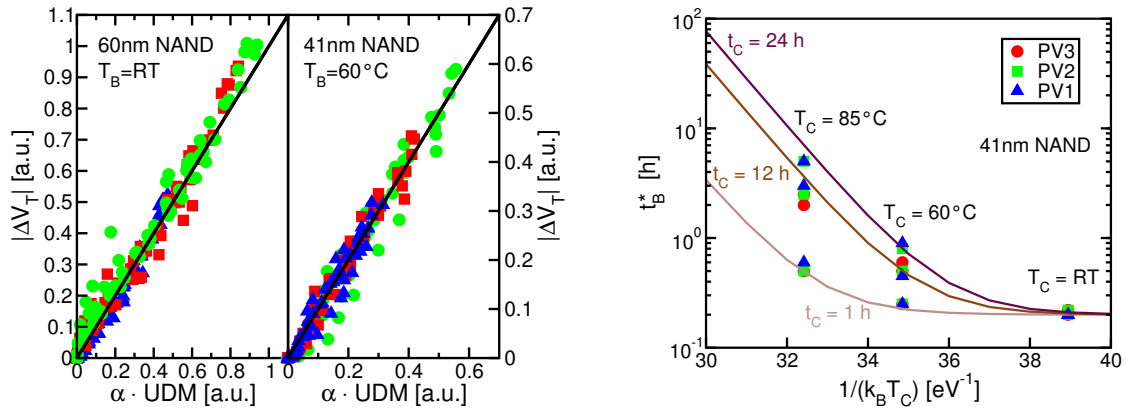
Figure 23. (**Left**) ΔV_T transients as a function of the UDM term (symbols represent the different programmed levels); (**Right**) Values of the parameter t_B^* extracted from experimental data (symbols) and Equation (11) (lines) [226], © 2011 IEEE.

The experimental data presented so far and the physical understanding leading to Equation (11) led to the recognition that traditional testing procedures based on fast cycling and long retention times were magnifying the V_T loss with respect to real usage conditions, when cycling takes place over the functional life of the device (see the effect of increasing t_C in Figure 22, left). The concept of *distributed cycling* conditions was then proposed in [221] as a way to better emulate the real array behavior by performing either a uniform cycling over a longer time or several groups of fast cycles preceded by bake times, usually at high temperature to accelerate the charge loss. However, the previous model shows its limitations when dealing with highly non-uniform cycling patterns, calling for a more comprehensive interpretation of the trapping/detrapping physics, that was pursued in [223–225,227,233]. Such a model was built upon a few phenomenological assumptions: a Poisson distribution for the number of trapped electrons n_t and a uniform distribution over a log-time axis of their detrapping time constant τ_d. This led to an expression for α [223,224]

$$\alpha \propto \langle \Delta V_T^1 \rangle \langle n_t \rangle, \tag{12}$$

which directly links the slope of the ΔV_T transient to the product of the average (indicated by the operator $\langle \cdot \rangle$) ΔV_T^1 and n_t. Via a clever analytical manipulation, an expression for the entire statistical distribution of ΔV_T can be obtained:

$$p(\Delta V_T(t)) = \mathscr{F}^{-1} \left\{ e^{\langle n_d(t) \rangle \left(\mathscr{F} \{ p(\Delta V_T^1) \} - 1 \right)} \right\}, \tag{13}$$

where $\mathscr{F}$ and $\mathscr{F}^{-1}$ are the Fourier transform and its inverse and the average number of detrapped electrons, $\langle n_d(t) \rangle$, is given by

$$\langle n_d(t) \rangle = \int \langle n_t(\log_{10} \tau_d) \rangle \left(1 - e^{-t/\tau_d} \right) d \log_{10} \tau_d, \tag{14}$$

an activation energy $E_A = 1.1$ eV being embedded into τ_d. Charge trapping is also inserted in a phenomenological manner, in agreement with the observed square-root dependence on N_C [221,222], resulting in the following expression for the average number of electrons being trapped at cycle N_C:

$$\Delta n_t(\tau_d) = \frac{\eta^2}{2 \langle n_t(\tau_d, N_C - 1) \rangle}. \tag{15}$$

In Equation (15), $\eta \approx 0.01$ is a fitting parameter, and the expression is modified for small values of n_t to avoid the divergence in Δn_t (see [224]). Typical results for the n_t distribution as a function of time (fast cycling conditions, where P/E cycles are conducted before the retention test) are shown

in Figure 24, left: at a given time t, because of the detrapping term $1 - e^{-t/\tau_d}$ in Equation (14), only electrons having $\tau_d \ll t$ are detrapped, while those with $\tau_d \gg t$ are unaffected. Nice agreement with experimental data (Figure 24, right) confirms the validity of this approach.
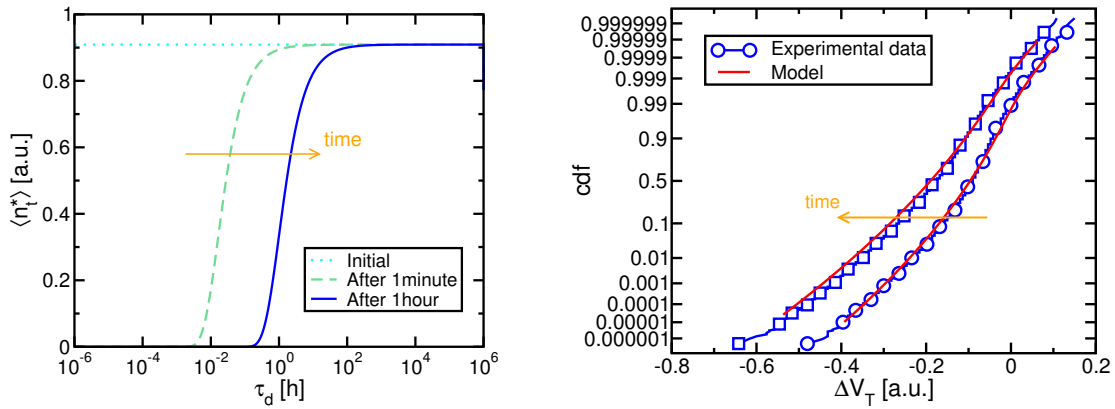

Figure 24. (**Left**) distribution of $n_t(\tau_d)$ for different bake times after fast cycling; (**Right**) comparison between experimental data and model results for ΔV_T.

The power of this model can be best appreciated when dealing with so-called *distributed cycling* schemes, where P/E cycles are spread over longer times and separated by bake phases at different temperatures, to emulate the real behavior of a device. Such schemes cannot be managed by the compact expression given by Equation (11), but are easily dealt with by the model synthesized by Equations (12)–(15). As a reference, we take three schemes that are deemed almost equivalent by the model adopted in the JEDEC standard [235] and are depicted in Figure 25: all involve fast-cycling phases at room temperature (marked as blue) separated by high temperature bakes (marked as red). In all cases, the array is subjected to $N_C = 10k$ cycles, but their distribution over time is different. Results for ΔV_T at the end of each process are reported in Figure 26 (left): different ΔV_T transients are obtained with the three schemes, featuring nonetheless the same asymptotic slope. The model nicely captures the differences and explains them in terms of the different distributions of $n_t(\tau_d)$, as a consequence of the trapping/detrapping phases. Note also that the asymptotic value of n_t is the same for all schemes, as it is due to the total number of cycles, explaining the uniqueness in the asymptotic slope (see Equation (12)).
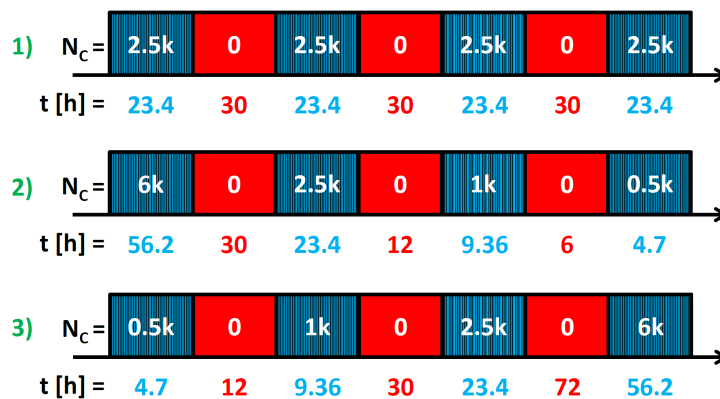

Figure 25. View of the considered distributed-cycling schemes that are almost equivalent from the viewpoint of the JEDEC standard [236].
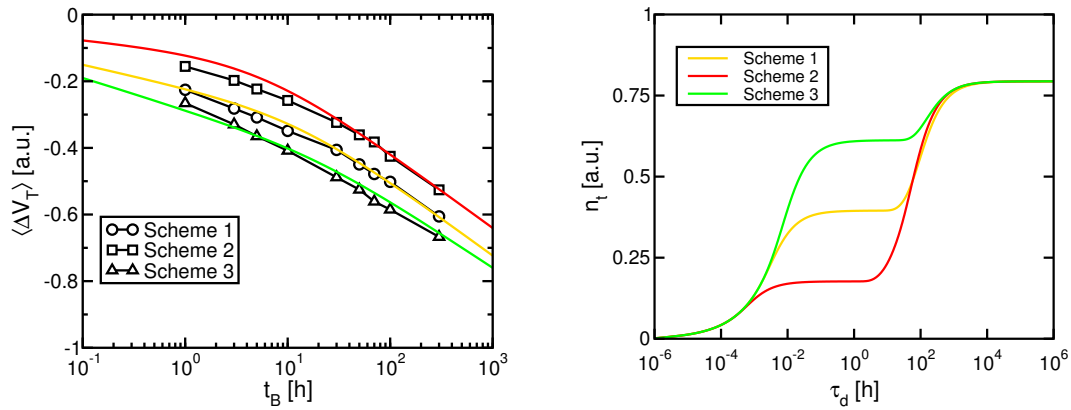
Figure 26. (**Left**) comparison between experimental data (symbols) and model results (lines) for ΔV_T after the distributed cycling schemes in Figure 25 [234], © 2014 IEEE; (**Right**) distribution of $n_t(\tau_d)$ at the end of the cycling schemes.

Although the above-mentioned model is successful in describing distributed-cycling procedures and provides a useful support for developing appropriate qualification schemes, it still misses some features, namely the programmed-level dependence of charge detrapping and α. In particular, it was noted that charge detrapping is enhanced with the oxide electric field and that little or no detrapping takes place during bake phases performed with cells in the erased state [227,237], suggesting that holes may play a non-negligible role in what is usually described as an electron-only detrapping process [238]. These developments led to a recent refinement of the above-mentioned model [225,233] that goes beyond the classical vision of pure charge detrapping, considering multiple-state defects related to structural relaxation of the oxide network. A pictorial view of the proposed defect states can be seen in Figure 27, where the lower states $D2$ and $N2$ represent the usual trapping/detrapping phenomena, for which the model discussed so far still holds (besides a minor modification to account for trapping). Now, however, the reservoir for charged defects that can exchange electrons with the substrate (state $D2$) is dictated by a structural relaxation process from an initial state $D1$ generated by P/E cycling, allowing a higher degree of flexibility in the description of the ΔV_T transients. By assigning a log-time distribution to the structural relaxation time constants while retaining nearly-constant values for the trapping and detrapping time constants (whose ratio is assumed to be affected by the oxide field), a greatly improved fitting was achieved, even over fine details of the experimental data. However, the assumptions of the model still need further investigations to be confirmed, and charge detrapping is likely to remain an active area of research in Flash reliability for quite some time to come.
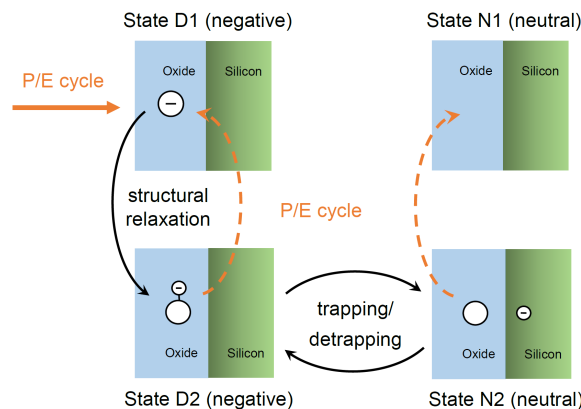

Figure 27. Pictorial view of the multiple-state defects responsible for charge trapping/detrapping according to [225,227,233].

3.4.3. Effect on V_T Distribution after Cycling

From the previous results, it is clear that charge detrapping leads to both an average charge loss and an enlargement of the V_T distribution, due to the randomness of the detrapping process. A comparative analysis of the impact of charge detrapping on the width of the V_T distribution after cycling and bake was reported in [239] and shown in Figure 28, normalized to the ISPP step V_s. The two bottom-most contributions are related to the time-0 placement of the V_T distribution—namely, the ISPP step V_s and the fluctuation in the number of injected electrons per step [158,159,240]—and are not discussed here. The remaining part is made up of a native RTN contribution plus the additional terms induced by cycling. Note that RTN is not dependent on the V_T level while detrapping increases with such level, as previously discussed, becoming a non-negligigble error source for the higher V_T levels. A similar breakdown of the distribution width was presented in [241], without accounting for detrapping. Other studies also identified charge detrapping as the dominant error source in highly-cycled NAND samples [242].
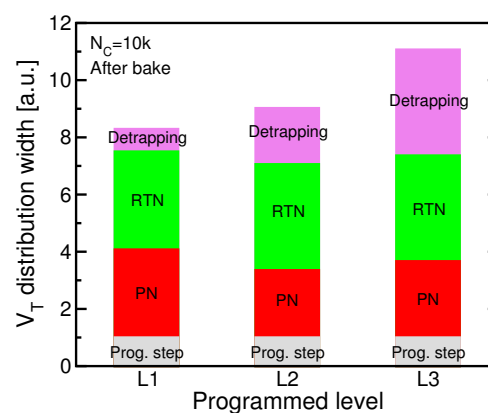

Figure 28. Contributions to the width of the V_T distribution after cycling and bake for different programmed levels [239], © 2015 IEEE.

To overcome this limitation, several solutions have been proposed, including a dummy read [243] or a moving reference [244] scheme and a *self-healing* approach [245], where a built-in heater in the proximity of the tunnel oxide is used to accelerate damage recovery and prolong the endurance. The insertion of specific block healing cycles during run time was shown to increase lifetime and performance of NAND devices [246], suggesting that distributed healing schemes could be devised to maximize their efficacy while reducing the energy dissipation of such cycles [247].

3.5. Interface State Recovery

In analogy with oxide defects, interface states are created during P/E cycles and recover in-between. Besides their effect on the endurance figure of NAND devices, discussed in Section 3.1, they also influence the V_T shift after cycling. In fact, it has been proposed that interface state recovery represents the major source of V_T instability in 90 nm NAND Flash [63,209,248]. Interface state annealing manifests itself via a reduction of the activation energy, down from the 1.1 eV associated to detrapping to about 0.7 eV [230,231] or even 0.15 eV [249], and are also characterized by a non-uniform distribution over the cell area, usually peaked at the STI edges [74,250]. However, it has been suggested that their effect is dominant for high values of the read current only [230,231], while playing a minor role with respect to charge detrapping as the read current is reduced.

3.6. Disturbs

The arrangements of Flash cells into compact arrays implies that many cells other than the selected ones are biased during read and program operations. Although such biases should not affect the

cell state, degradation in the tunnel oxide characteristics following P/E cycling can reduce the cell immunity to such disturbs. The NAND layout generates three types of disturbs, namely related to the read, program and pass operation, that are discussed hereafter.

3.6.1. Read Disturb

The bias condition of a NAND block during read has been discussed in Section 2.2 and is shown in Figure 29: a suitable read voltage is applied to the selected WLs to sense whether the cells in the page are in the on or off condition, while all other cells in the string are biased to a pass voltage higher than their maximum V_T. The pass bias is hence analogous to a program operation, albeit performed at a lower gate bias, and as such it can induce a V_T shift in the cells. This *read disturb* (RD) cumulates with the number of read operations and could eventually lead to a failure. Note that this is a recoverable error, that is nulled after the block is erased and reprogrammed.
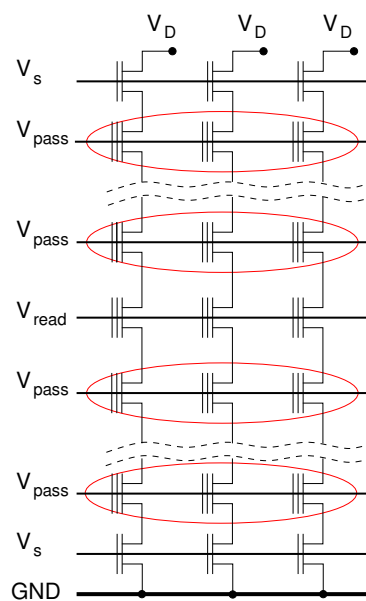

Figure 29. NAND array bias condition during read. Cells subjected to RD are circled.

The oxide thickness and bias conditions are designed to assure that no significant V_T shift is expected after RD. However, this might no longer be the case after heavy P/E cycling, when the gate oxide quality is degraded and SILC and charge detrapping arise, enhancing the conduction at low fields (see the *I–V* characteristics in Figure 18, left). RD is hence exacerbated in cycled cells [251–253] and in read-intensive applications such as database access, explaining why a lower endurance is sometimes specified in such cases.

Experimental data for the memory lifetime report a power-law dependence on N_C, with exponent ranging from −0.3 to −0.4 [251,254]. Data for the V_T shift as a function of the number of reads feature different dependences [255,256], while raw bit-error rate (RBER) data vs. the number of reads reveal a linear dependence whose slope increases with cycles [60,255,257], with a higher BER observed on SLCs than on MLCs. As with SILC, several process optimization [258,259], voltage tuning [255,260] or system-level correction [261] proposals have been put forward to alleviate the issue.

Typical RD characteristics are reported in Figure 30 [255]: the V_T shift is larger for the erased state and decreases for higher programmed V_T levels, in agreement with the reduction on the tunnel oxide field. As a consequence, more errors are expected in the logical pages corresponding to those states. It has also been found [262] that repeated RD on partially-programmed blocks leads to an increased BER when the remaining pages are programmed, suggesting that extensive RD can lead to

a degradation of the program characteristic of the cell. In any case, it has been claimed [61] that RD errors are not the prevalent ones in SSDs used in servers.
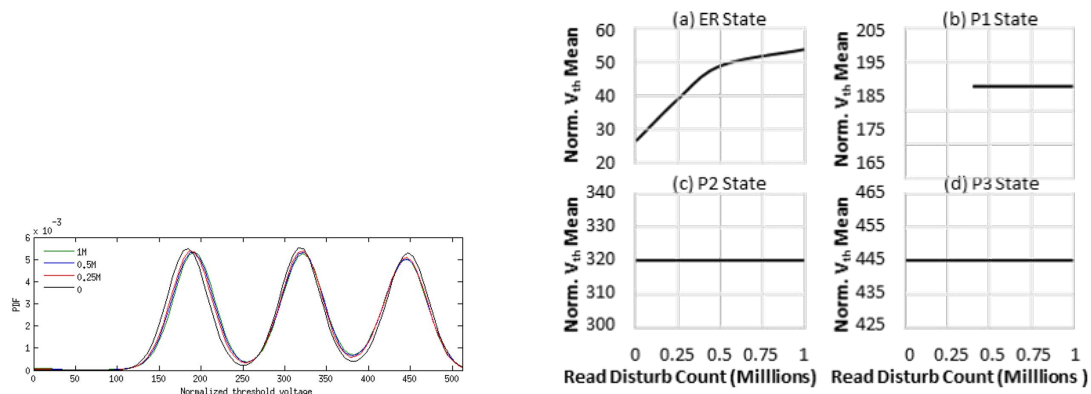

Figure 30. (**Left**) V_T distribution for increasing RD time; (**Right**) mean V_T shift of the different levels as a function of the number of reads [255], © 2015 IEEE.

3.6.2. Program and Pass Disturbs

When programming a page, we must ensure that only the desired BLs are actually programmed, while other cells on the same page that must remain in their original state are not affected by such operation. Since electrons are injected into the floating gate via FN tunneling, which is a process strongly dependent on the oxide field, we can inhibit programming by reducing such a field, i.e., the gate-to-channel voltage drop. The simpler solution is to disconnect (i.e., ground) the SSL (see Figure 1) and raise the inhibited BLs to a high voltage [263,264], while biasing the unselected WLs too, to avoid unintentional erase. However, this solution was soon abandoned because of power and time issues related to the charging of highly-capacitive BLs and because of high-voltage limitations in scaled technologies. In replacement, a *self boosting* technique was proposed [265], illustrated in Figure 31: BLs to be programmed are grounded, which turns on their DSL and maintains the string channel to ground (green line). Note that a pass voltage (different from the one used during read) is applied to the unselected WLs. Inhibited BLs, instead, are biased at V_{cc}, which precharges their channel. When the high passing voltage is applied to the WLs, however, the channel potential (red line) is lifted by capacitive coupling, turning off the DSL and reaching a value high enough to suppress electron injection into the floating gate.

This solution generates two kinds of disturbs, highlighted in Figure 32: a program disturb on the cells on the programmed WL (encircled in green), related to the potential difference between program and boosted bias, and a pass disturb on the cells located on the same (grounded) BL as the programmed ones (red circles), related to the pass voltage. The nature of such disturbs is just the same as the RD, namely a soft-programming at low voltage, with the difference that now the applied field is higher with respect to RD, while the number of cycles is significantly lower. Note also that an increase in the pass voltage will lead to an increase in the channel potential, reducing the program disturb but raising the pass one. This typical behavior is shown in Figure 33 (left). While a logarithmic dependence of V_T on the number of cycles has been reported [266], the dependence of BER is more complex, as shown on the right of Figure 33 [257]. To limit the pass disturb, improved boosting schemes have been developed, including local boosting [36,267], ramped WL voltage [268], dynamic pass voltage biasing [269] or negative WL pre-biasing [270]. Interestingly, a self-boosting scheme during read was also proposed, to alleviate read-disturb issues [271].
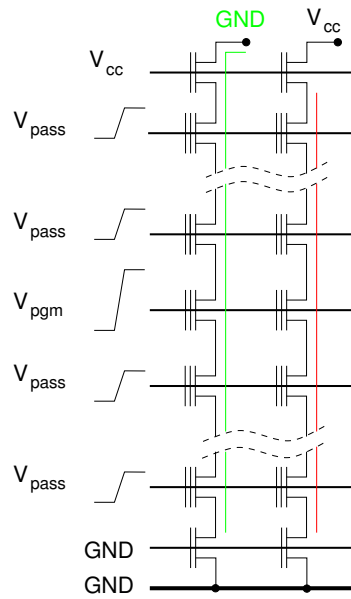
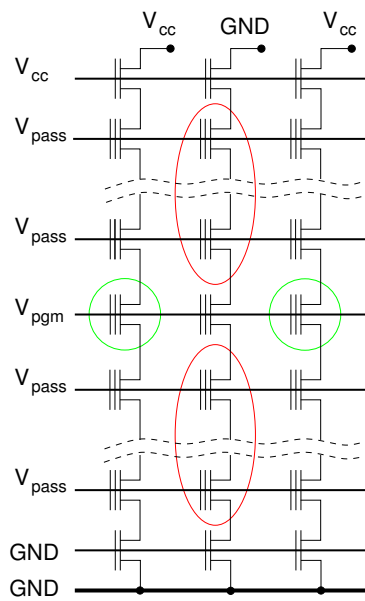**Figure 31.** NAND array bias condition during program, highlighting the self-boosting effect.

Figure 32. NAND array bias condition during program. Cells subjected to program (green) and pass (red) disturbs are circled.
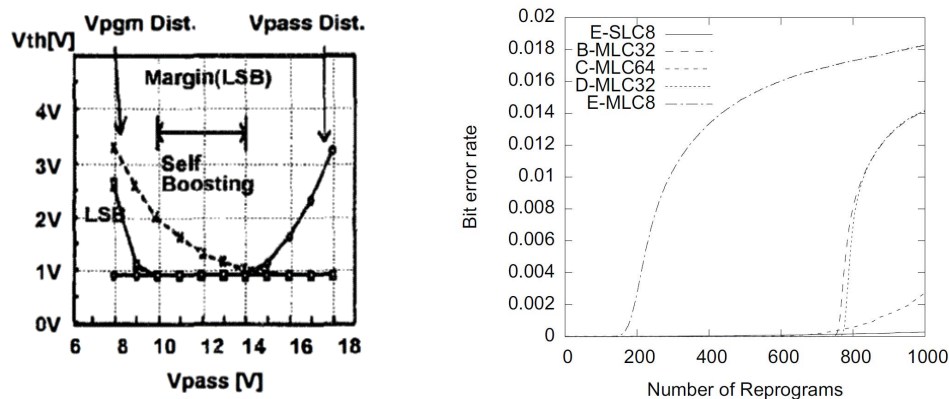
Figure 33. (**Left**) V_T shift due to program and pass disturbs as a function of the pass voltage [36], © 1996 IEEE; (**Right**) BER due to program disturb as a function of the number of P/E cycles [257], © 2009 ACM.

The boosting operation in scaled NAND devices is also responsible for additional reliability concerns. First, the string select transistors must withstand the boosted potential with minimum leakage current, as such current will discharge the channel and increase program disturb (for this reason, such devices are hard to scale). However, under program inhibit conditions, the SSL transistor is subjected to gate-induced drain leakage (GIDL) due to band-to-band tunneling, injecting hot electrons toward the first WL (see Figure 32) and into its floating gate, when erased [266]. This effect leads to an enhanced program disturb at the first WL and is usually curbed by increasing the SSL-to-WL spacing [266] or employing one [272] or two [273] dummy edge WLs. In fact, because of short-channel effects in scaled technologies, the actual electron injection point into the floating gate can move farther into the string, affecting even inner cells [274]. Moreover, such hot carriers may be trapped into the oxide and be subsequently released, giving rise to detrapping-related retention degradation [275]. The full picture is further complicated by electrostatic coupling between adjacent strings, via the STI isolation: when a string is programmed, its channel is kept at zero voltage and it will behave like the gate of a parasitic transistor whose gate oxide is the STI isolation. Adjacent inhibited cells will therefore experience large vertical GIDL current, between the (boosted) surface of the channel and its bulk [276]. Air-gap isolation as well as optimization of channel implant and STI depth uniformity [277] have been suggested as remedies. A related effect has been noted in [278], where a negative V_T shift contribution was found and explained in terms of hot-hole injection from the STI edge to the control-gate [279,280]. All this suggests that the detailed leakage mechanisms affecting the boosted channel potential must be studied with extreme care [281].

In local self-boosting schemes, where the channel of the programmed WL is isolated from the rest of the string by grounding the neighboring WLs and is boosted by the programmed Wl, other issues arise, such as hot-electron injection via drain-induced barrier lowering (DIBL) [282,283], that paved the way for proposals of hot-carrier based programming of NAND cells [284,285].

In conclusion, this overview demonstrates that many reliability issues have threatened the scaling of the planar NAND technology. In turn, its successful scaling down to the 15 nm node is a tribute to the effectiveness of the many different solutions adopted to keep the failure rate under control. NAND-based SSDs are today the main candidates for mass storage, having reached reliability performance comparable to, or in some aspects even better than, HDDs [62]. In the following, we will extend this discussion to the current leading-edge technology for NAND: 3D.

4. 3D NAND Reliability

In this Section, we will mainly focus our attention on the most promising 3D NAND architecture, i.e., the one based on vertical strings of cells and parallel planes connecting the control gates [23–25].

A schematic view of the memory array organization is shown in Figure 34 (left), but we will leave the details about the array operation to the specific literature [28]. The individual cell is a vertical-channel transistor with polysilicon body, usually shaped as a hollow cylinder filled with dielectric [286], whose cross-section is depicted on the right of Figure 34. Such a device was shown to exhibit superior performance with respect to a conventional vertical nanowire transistor, because of the high defectivity in the central region that plagued the performance of the latter structure. The gate stack of today's 3D NAND can be based on either a floating gate [23,287–291], similar to planar NAND devices, or a charge-trap stack similar to an oxide/nitride/oxide (ONO) layer, where the charge is stored in traps within the nitride layer [10,11,22,292].

As far as reliability is concerned, all mechanisms effective in planar NAND cells are also present in 3D devices, though their impact is affected by the new cell geometry and design and array operation. Unfortunately, the very limited amount of data publicly available makes it hard to conduct an exhaustive assessment of the role played by each of them. On a general standpoint, the 3D cylindrical cell footprint and channel length can be made larger than their planar counterparts [22,291], as the 3D architecture relieves the burden of scaling the cell area by placing it onto the increase in the number of stacked layers. As a consequence, a better reliability is achieved in 3D NAND cells, which resulted in the development of QLCs. However, the vertical polysilicon channel and the layer stacking give rise to new and unexpected issues, providing ever newer ground for research. In the following, we will review the reliability issues in 3D NAND within the same framework adopted so far: we will only discuss issues that become relevant with P/E cycling or repetitive data read, neglecting those associated, for example, with V_T placement, string resistance and time-0 variability.
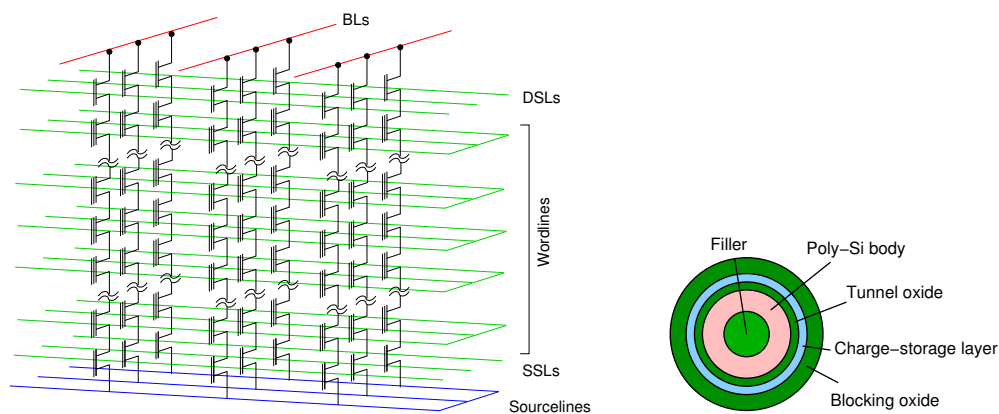
Figure 34. (**Left**) schematic view of the array organization in a 3D NAND; (**Right**) schematic cross-section of the typical NAND cell used in 3D arrays.

4.1. Endurance

Figure 35 shows the endurance performance of a 3D NAND cell as a function of N_C [286] and the V_T distribution for a NAND array after heavy cycling [22]. Because of the larger cell size, 3D NAND are reported to have better endurance properties with respect to their planar counterpart, as also claimed in [293–296], where an endurance of 5-7 k P/E cycles is reported for a TLC 3D NAND. All these cells feature an ONO stack. Retention data on a floating-gate based 3D NAND [290] show comparable results, with higher V_T shift but larger voltage window.
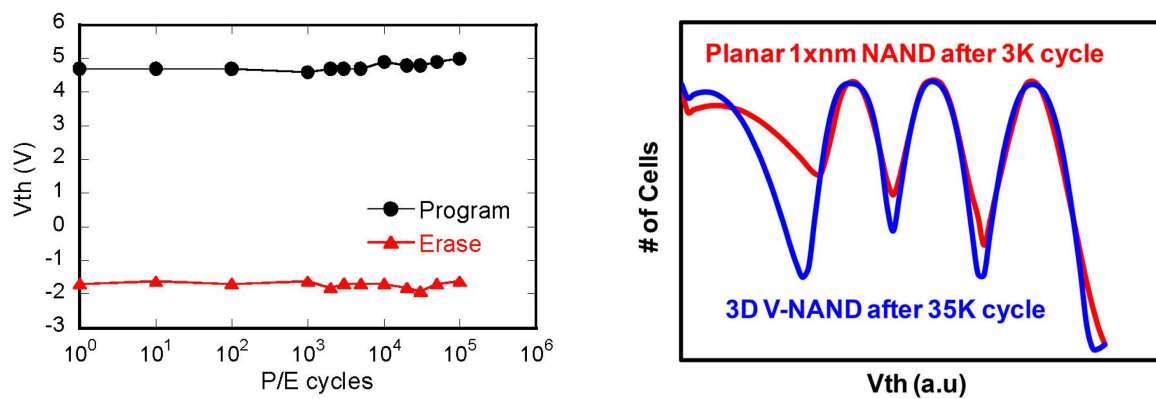
Figure 35. (**Left**) Endurance figure for a 3D NAND [286], © 2007 IEEE; (**Right**) V_T distribution after cycling for a planar and a 3D NAND TLC [22], © 2015 IEEE. Devices are charge-trap based.

4.2. Retention

Floating-gate based 3D NAND suffer from anomalous SILC after heavy P/E cycling, though this effect is not expected to be the dominant cause of retention loss. Charge-trap based memories are claimed to be insensitive to such a loss [297], due to the localized nature of storage, although it has been shown [298] that defect build-up in the tunnel oxide eventually leads to TAT mechanism enabling charge loss from the nitride layer. In any case, to our knowledge, there are no published reports of anomalous SILC in 3D NAND, so it seems safe to extend to this technology the conclusion drawn for planar arrays, where other effects are dominating retention.

A comparison between 2D and floating-gate based 3D NAND with respect to detrapping is reported in Figure 36 (left), showing a smaller ΔV_T spread for the 3D array, even if subjected to a higher number of P/E cycles [299]. This is once again the result of the larger cell in 3D technology, i.e., larger gate capacitance, minimizing the effect of single-charge detrapping onto V_T. Published single-cell data (Figure 36) suggest that a good intrinsic retention can be achieved, as expected from a floating-gate design [289].
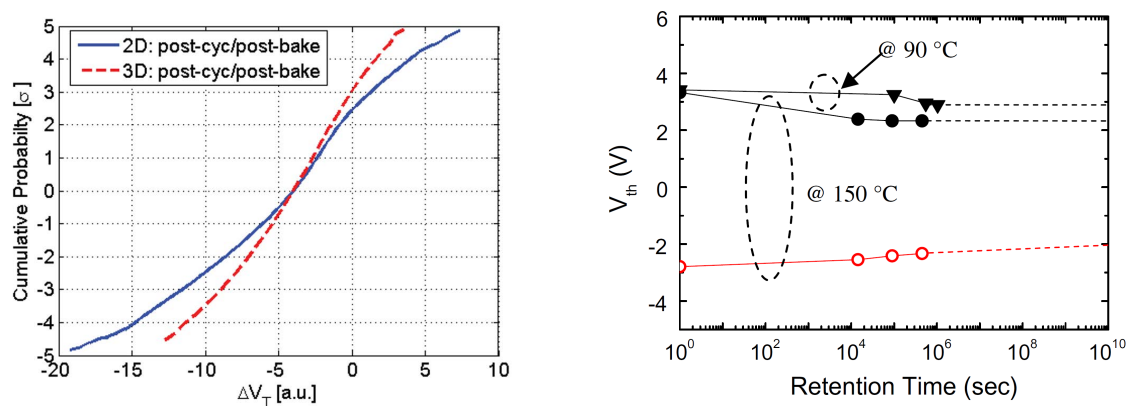
Figure 36. (**Left**) comparison between 2D and 3D cycled NAND arrays for V_T shift induced by charge detrapping [299], © 2015 IEEE; (**Right**) retention for different temperatures [289], © 2013 IEEE. All devices are floating-gate based.

The retention properties of charge-trap based planar devices have been extensively investigated as a function of temperature and bias [300], pointing out the roles of vertical charge loss through the tunnel oxide and lateral charge migration in the nitride [301–304], which also reduces the stored charge located above the channel region, that determines V_T. In 3D devices, the importance of such effects is shown in Figure 37, where lateral diffusion along the nitride is modulated by the state of

adjacent cells [305]: having all cells programmed decreases the concentration gradient, slowing down diffusion and reducing the V_T shift. Similar results obtained via simulations were also presented in [306–308] and, changing the device size, in [309]. Note also the thermal activation, accelerating both leakage mechanisms and making pattern-dependent charge spreading effects more relevant at high temperatures. The resulting multiple activation energies are discussed in [310]. An analogous charge-gain phenomenon for the low-V_T state was reported in [311]. Such a feature had been already discussed in planar nitride memories and floating-gate devices, being interpreted in terms of recovery of negative bias-temperature instability [312,313].
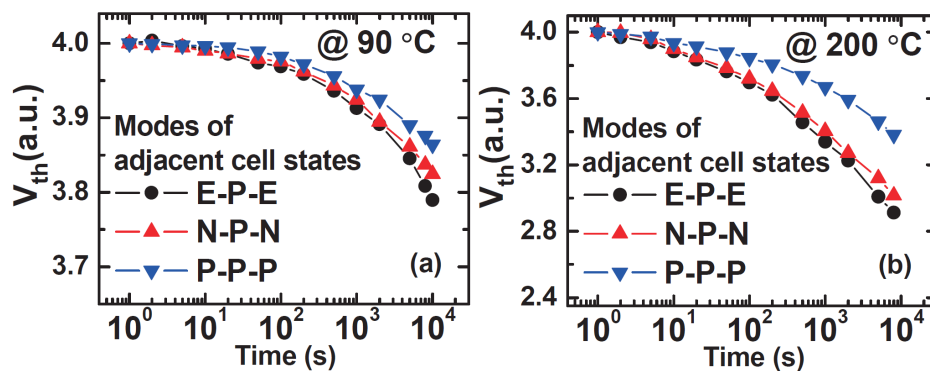
Figure 37. Charge retention characteristics for a charge-trap based 3D NAND as a function of the state of the adjacent cells and for different temperatures [305], © 2015 IEEE.

A related effect that was first observed in planar devices [314] is a fast charge loss taking place on a μs timescale [315]. Currently, there is no consensus on the root cause of this phenomenon, ascribed to either electron detrapping from shallow traps in the nitride [314] or in the tunnel oxide [316], or to lateral electron migration enhanced by holes injected during erase that are not fully compensated by electrons [315]. The effect is however dependent on the nature of the gate stack, being smaller in conventional ONO stacks with respect to band-engineered ones, and its magnitude increases with cycling.

4.3. Cell Current Fluctuations

Cell current fluctuations in 3D cells, related to capture and emission of single carriers, present some distinct features that differentiate them from what is usually seen in planar devices. These can be ascribed to the trap location, either at one of the oxide interfaces (tunnel oxide and filler) or within the polysilicon itself, because of the polycrystalline nature of the conductive channel, and to 3D percolation as well as the conduction mechanisms of electrons in the polysilicon. These factors are discussed hereafter.

4.3.1. Random Telegraph Noise

RTN in silicon nanowire structures has received attention in the past and single-trap characterization in vertical structures for 3D NAND has been also presented [317–319]. Statistical data for RTN are shown in [320–322] and are presented in Figure 38 (left) for different temperatures. Note that the usual exponential distribution of amplitude appears, whose slope is greatly reduced with respect to planar technologies [299,323,324]. This is due not only to the larger cell size, but also to the different doping of the channel and nature of the electron conduction process in the subthreshold region: a uniform conduction over the channel area is expected in 3D cells [325], as opposed to a surface conduction in planar devices. The effect of the improved RTN figure on the V_T distribution width is shown on the right of Figure 38, making clear that a tighter V_T distribution can be achieved with 3D

NAND. This is confirmed by data in Figure 39, where the V_T distribution width in 2D and 3D NAND is compared: note the improvement in both RTN and detrapping contributions.
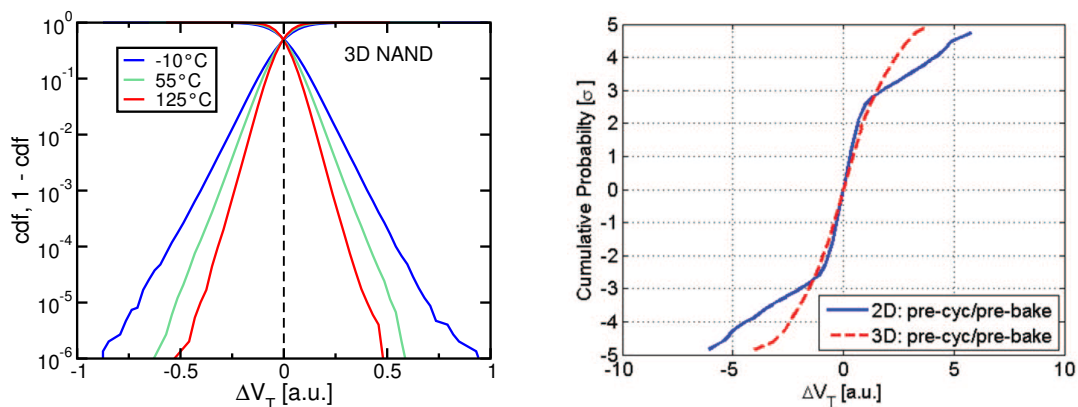
Figure 38. (**Left**) RTN cdf after cycling for a planar and a 3D NAND as a function of temperature [324]; (**Right**) ΔV_T cdf for a planar and a 3D NAND [299], © 2015 IEEE. Devices are floating-gate based.
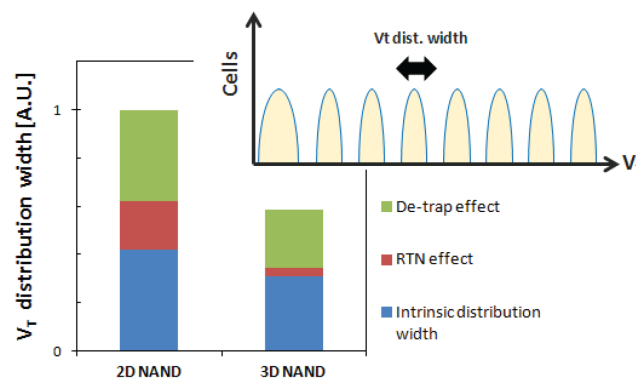
Figure 39. Contributions to the V_T distribution width in planar and 3D NAND [299], © 2015 IEEE.

4.3.2. Polysilicon Grains

The temperature dependence of RTN in 3D NAND that can be seen in Figure 38 (left) is quite different with respect to what is reported for planar devices (see Figure 11): while RTN in planar NAND is temperature-independent, its slope decreases at higher temperatures in 3D NAND (similar dependences were also reported by [326,327]) and affect all cells in the array (see the unique exponential distribution in 3D NAND compared to the exponential tails departing from a central Gaussian distribution in the 2D case). The explanation for the exponential distribution of the RTN ΔV_T that was put forward in Section 3.2.3 ascribes such fluctuations to current percolation around the charged traps, meaning that different conduction mechanisms and/or trap locations should now be considered and reminding us that RTN remains a powerful tool for probing electron conduction in microscopic devices. In 3D NAND devices, in particular, the vertical string is usually made of polycrystalline silicon, because of the fabrication process. Such a material is composed of many single-crystal "grains" having different orientations, separated by grain boundaries that are highly-defective regions [328]. As a matter of fact, since its early use in analog resistors or thin-film transistors, it was recognized that polysilicon characteristics such as low-frequency noise are dominated by capture/emission at the grain boundary interfaces [329,330]. More recently, the effect of polysilicon grain traps charging and discharging on the transient BL current has been addressed [331,332].

A characterization of electron conduction in vertical poly-Si channels was presented in [333], comparing different technological processes and concluding that charge capture/emission from traps within the poly-Si is the main source of current fluctuations, as opposed to traps at the silicon/oxide interfaces. RTN improves for larger cells and smaller average grain size, as shown in Figure 40 [326,334] (where it was measured above the threshold). Other works also reported poor reliability and higher variability in 3D NAND for large poly-Si grain size [317,335]. This is because the larger number of grains associated with their smaller size results in a self-averaging effect, as discussed in [336] for the case of a polysilicon gate in a planar transistor.

These features call for a thorough analysis of the conduction mechanism in poly-Si channels, accounting for grain boundaries and related traps. The current understanding of the phenomenon is that defects at grain boundaries can trap electrons and act as percolation centers, while causing a degradation in the carrier mobility with respect to its bulk value [337,338]. Charge transport is actually limited by the energy barriers at the grain boundaries (due to the trapped charge) rather than by transport within the grains, except for very large grain size [339]. Several works have exploited numerical simulations to investigate the effect of grain boundaries on variability in nanowires [340–345] and 3D NAND devices [346,347]. However, several important features of such models still have to be assessed, such as the grain size [348,349], the density and energy distribution of grain boundary traps [350,351], and the mobility degradation and conduction process at the grain boundaries [352,353].
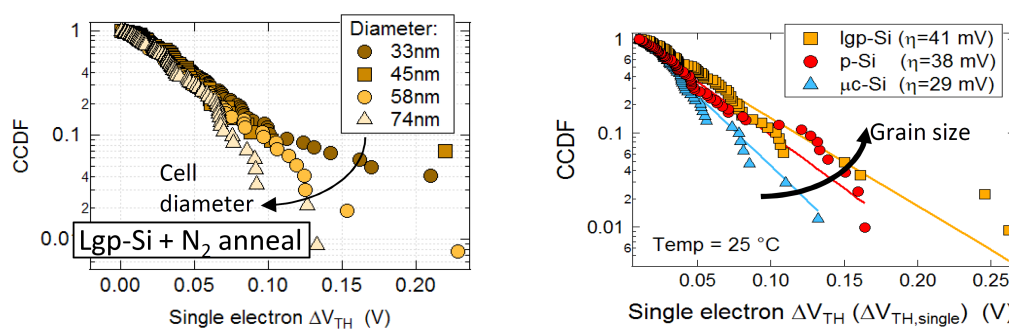
Figure 40. RTN complementary cdf for vertical poly-Si channel transistors having different channel diameter (**left**) and poly-Si grain size (**right**) [334], © 2012 IEEE.

4.4. Disturbs

In the 3D NAND architecture of Figure 34, strings are vertically aligned while control gates run over parallel planes [28], meaning that read and program/pass disturbs are much more complex with respect to a planar layout, as different pages now share the same WL. Moreover, the onset of disturbs is strongly related to the actual array architecture and internal subdivision as well as operation mode; therefore, extreme care should be taken before generalizing the following results. A schematic of the different types of program disturb is reported in Figure 41 [294,354]: an *x* mode for cells on the same selector line but different BLs, analogous to conventional 2D arrays, a *y* mode for cells on the same BL, and a mixed *xy* mode for all other cells sharing the programmed WL. The modes differ for the BL bias (selected cell and *y* mode disturb = 0; *x* and *xy* modes = high) and the drain selector bias (selected cell and *x* mode disturb = high; *y* and *xy* modes = 0), meaning that different boosting levels are achieved. As a matter of fact, the channel is first precharged to a high voltage in *x*-mode disturb before being boosted, while it is boosted from ground level during *y* mode, resulting in a lower potential (i.e., increased disturbs) in the latter case, as shown in Figure 42 (left) [355].
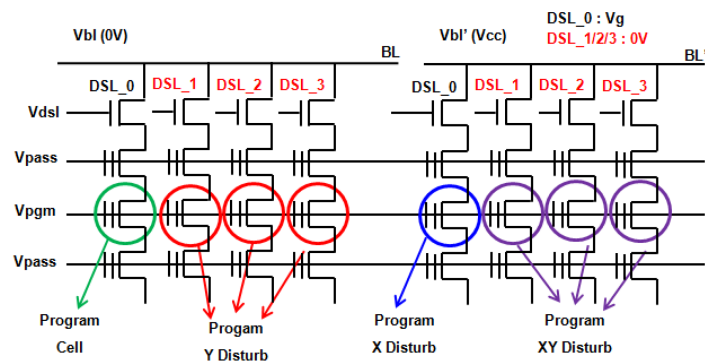
Figure 41. Schematic view of 3D NAND program disturb modes [354], © 2012 IEEE.

The boosting efficiency of 3D NAND is affected by different factors, such as the coupling coefficient and the leakage current. A high coupling coefficient is expected in 3D NAND because of the floating channel lacking a silicon substrate [356]. On the other hand, the worse subthreshold slope of 3D polysilicon devices with respect to planar devices increases the leakage current that discharges the boosted channel potential [294,354]. Clever device optimization, including selector design and insertion of dummy cells to relieve edge WL disturbs, eventually leads to much alleviated program disturbs, as shown in Figure 42 (right) [354]. Other optimization schemes such as an initial channel precharge phase [357] and selector doping [358] have been also proposed. In analogy with planar devices, suitable management schemes for 3D NAND have been also proposed [359,360]. Similar considerations apply to vertical-gate structures, in which channel arrays are horizontally stacked with vertical WLs [361,362].
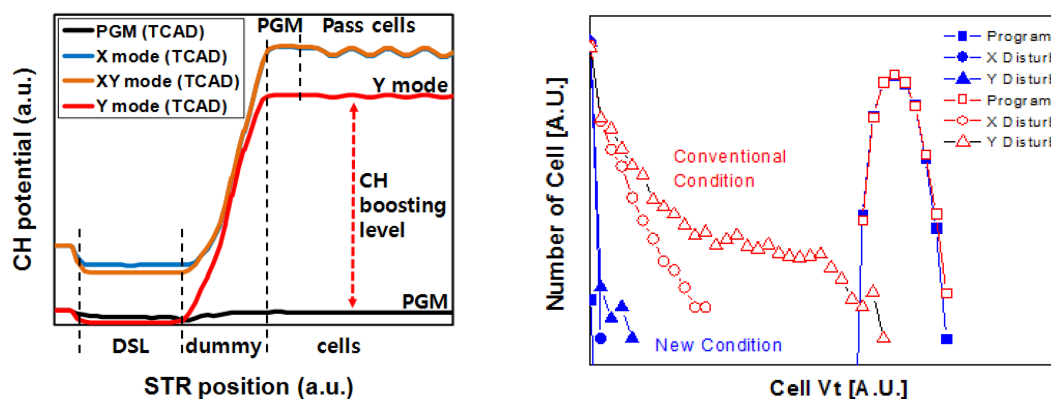
Figure 42. (**Left**) Channel potential for different cell modes during boosting in a 3D NAND [355], © 2013 IEEE; (**Right**) V_T pdf for programmed and disturbed cells in a 3D NAND [354], © 2012 IEEE. Devices are charge-trap based.

A study of read disturb in 3D NAND was presented in [363], highlighting an enhanced disturb on unselected cells due to channel boosting. A related effect was discussed in [364], where it is shown that the trailing edge of the WL bias at the end of a read phase can draw the channel potential negative, as it becomes floated when high-V_T cells along the string are turned off. This can become an issue if the read operation is followed by a program one, as in program-verify steps: the channel of program-inhibited cells will now reach a lower boosted potential, resulting in increased program disturbs. Suitable program schemes must therefore be adopted to avoid this issue.

5. Conclusions

NAND reliability limitations arise from the high-field operation needed to store and remove electrons to/from the floating gate and its undesirable effect on the tunnel oxide, whose insulating

properties degrade with usage. This results in a series of issues affecting endurance, retention and read accuracy which were thoroughly reviewed in this paper, with an eye on their physical origin and dependence on cell/array parameters. Physical understanding of those issues led to clever solutions in terms of array organization and operation, boosted by process/technology advancements, that have pushed the planar technology to the 15 nm node, after which the conventional scaling path could no longer be followed.

The revolutionary move to 3D NAND resulted in a larger cell size, relieving the impact of traditional reliability issues, but gave rise to new and subtle reliability concerns, owed to the floating body and polysilicon channel of 3D NAND cells. These reliability issues are reviewed in this paper, comparing their relevance with what is obtained on planar devices. It is clear that 3D NAND allowed the cell architecture to be engineered and allowed the array operation to increase the overall reliability of the cell, resulting in improved performance in terms of storage density, as testified by the impressive development of QLCs, and read/write throughput. A key understanding of the physical mechanisms at the basis of the cell operation and degradation with usage remains of utmost importance to support the development of ever-improving solid-state storage solutions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Harari, E. Electrically Erasable Non-Volatile Semiconductor Memory. U.S. Patent 4,115,914, 26 September 1978.

2. Masuoka, F.; Asano, M.; Iwahashi, H.; Komuro, T.; Tanaka, S. A new Flash E^2PROM cell using triple polysilicon technology. In Proceedings of the 1984 International Electron Devices Meeting, San Francisco, CA, USA, 9–12 December 1984; pp. 464–467.

3. Masuoka, F.; Momodomi, M.; Iwata, Y.; Shirota, R. New ultra high density EPROM and Flash EEPROM with NAND structure cell. In Proceedings of the 1987 International Electron Devices Meeting, New York, NY, USA, 6–9 December 1987; pp. 552–555.

4. Micheloni, R.; Marelli, A.; Eshghi, K. (Eds.) *Inside Solid State Drives (SSDs)*; Springer: Heidelberg/Berlin, Germany, 2013.

5. Fontana, R.E.; Decad, G.M.; Hetzler, S.R. Volumetric density trends (TB/in.^{3}) for storage components: Tape, hard disk drives, NAND, and blu-ray. *J. Appl. Phys.* **2015**, *117*, 17E301-1–17E301-4.

6. Tokutomi, T.; Doi, M.; Hachiya, S.; Kobayashi, A.; Tanakamaru, S.; Takeuchi, K. Enterprise-grade 6x fast read and 5x highly reliable SSD with TLC NAND-Flash memory for Big-Data storage. In Proceedings of the 2015 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 22–26 February 2015; pp. 140–142.

7. Kim, H.J.; Lim, J.D.; Lee, J.W.; Na, D.H.; Shin, J.H.; Kim, C.H.; Yu, S.W.; Shin, J.Y.; Lee, S.K.; Rajagopal, D.; et al. 1 GB/s 2 Tb NAND Flash multi-chip package with frequency-boosting interface chip. In Proceedings of the 2015 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 22–26 February 2015; pp. 138–140.

8. Choi, S.; Kim, D.; Choi, S.; Kim, B.; Jung, S.; Chun, K.; Kim, N.; Lee, W.; Shin, T.; Jin, H.; et al. A 93.4 mm^2 64 Gb MLC NAND-Flash memory with 16 nm CMOS technology. In Proceedings of the 2014 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 9–13 February 2014; pp. 328–329.

9. Helm, M.; Park, J.K.; Ghalam, A.; Guo, J.; wan Ha, C.; Hu, C.; Kim, H.; Kavalipurapu, K.; Lee, E.; Mohammadzadeh, A.; et al. A 128 Gb MLC NAND-Flash device using 16 nm planar cell. In Proceedings of the 2014 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 9–13 February 2014; pp. 326–327.

10. Lee, S.; Lee, J.-Y.; Park, I.-H.; Park, J.; Yun, S.-W.; Kim, M.-S.; Lee, J.-H.; Kim, M.; Lee, K.; Kim, T.; et al. A 128 Gb 2b/cell NAND Flash memory in 14nm technology with $t_{PROG} = 640$ μs and 800 MB/s I/O rate. In Proceedings of the 2014 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 31 January–4 February 2016; pp. 138–139.

11. Sako, M.; Watanabe, Y.; Nakajima, T.; Sato, J.; Muraoka, K.; Fujiu, M.; Kouno, F.; Nakagawa, M.; Masuda, M.; Kato, K.; et al. A low-power 64 Gb MLC NAND-Flash memory in 15 nm CMOS technology. In Proceedings of the 2015 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 22–26 February 2015; pp. 128–129.

12. Prall, K.; Parat, K. 25 nm 64 Gb MLC NAND technology and scaling challenges. In Proceedings of the 2010 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 6–8 December 2010; pp. 102–105.

13. Goda, A.; Parat, K. Scaling directions for 2D and 3D NAND cells. In Proceedings of the 2012 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 10–13 December 2012; pp. 2.1.1–2.1.4.

14. Lee, S. Scaling challenges in NAND Flash device toward 10 nm technology. In Proceedings of the 2012 4th IEEE International Memory Workshop (IMW), Milan, Italy, 20–23 May 2012; pp. 1–4.

15. Monzio Compagnoni, C.; Goda, A.; Spinelli, A.S.; Feeley, P.; Lacaita, A.L.; Visconti, A. Reviewing the evolution of the NAND Flash technology. *Proc. IEEE* **2017**, doi:10.1109/JPROC.2017.2665781.

16. Endoh, T.; Kinoshita, K.; Tanigami, T.; Wada, Y.; Sato, K.; Yamada, K.; Yokoyama, T.; Takeuchi, N.; Tanaka, K.; Awaya, N.; et al. Novel ultra high density Flash memory with a stacked-surrounding gate transistor (S-SGT) structured cell. In Proceedings of the 2001 International Electron Devices Meeting (IEDM'01), Washington, DC, USA, 2–5 December 2001; pp. 33–36.

17. Jung, S.M.; Jang, J.; Cho, W.; Cho, H.; Jeong, J.; Chang, Y.; Kim, J.; Rah, Y.; Son, Y.; Park, J.; et al. Three dimensionally stacked NAND Flash memory technology using stacking single crystal Si layers on ILD and TANOS structure for beyond 30 nm node. In Proceedings of the 2006 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 11–13 December 2006; pp. 37–40.

18. Lai, E.K.; Lue, H.T.; Hsiao, Y.H.; Hsieh, J.Y.; Lu, C.P.; Wang, S.Y.; Yang, L.W.; Yang, T.; Chen, K.C.; Gong, J.; et al. A multi-layer stackable thin-film transistor (TFT) NAND-type Flash memory. In Proceedings of the 2006 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 11–13 December 2006; pp. 41–44.

19. Park, K.T.; Kim, D.; Hwang, S.; Kang, M.; Cho, H.; Jeong, Y.; Seo, Y.I.; Jang, J.; Kim, H.S.; Jung, S.M.; et al. A 45 nm 4Gb 3-dimensional double-stacked multi-Level NAND Flash memory with shared bitline structure. In Proceedings of the 2008 IEEE International Solid-State Circuits Conference—Digest of Technical Papers, San Francisco, CA, USA, 3–7 February 2008; pp. 510–511.

20. Tanaka, H.; Kido, M.; Yahashi, K.; Oomura, M.; Katsumata, R.; Kito, M.; Fukuzumi, Y.; Sato, M.; Nagata, Y.; Matsuoka, Y.; et al. Bit cost scalable technology with punch and plug process for ultra high density Flash memory. In Proceedings of the 2007 IEEE Symposium on VLSI Technology, Kyoto, Japan, 12–14 June 2007; pp. 14–15.

21. Park, K.T.; Han, J.M.; Kim, D.; Nam, S.; Choi, K.; Kim, M.-S.; Kwak, P.; Lee, D.; Choi, Y.H.; Kang, K.M.; et al. Three-dimensional 128Gb MLC vertical NAND Flash-memory with 24-WL stacked layers and 50MB/s high-speed programming. In Proceedings of the 2014 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 9–13 February 2014; pp. 334–335.

22. Park, K.T.; Nam, S.; Kim, D.; Kwak, P.; Lee, D.; Choi, Y.H.; Choi, M.H.; Kwak, D.H.; Kim, D.H.; Kim, M.S.; et al. Three-dimensional 128 Gb MLC vertical NAND Flash memory with 24-WL stacked layers and 50 MB/s high-speed programming. *IEEE J. Solid-State Circuits* **2015**, *50*, 204–213.

23. Tanaka, T.; Helm, M.; Vali, T.; Ghodsi, R.; Kawai, K.; Park, J.K.; Yamada, S.; Pan, F.; Einaga, Y.; Ghalam, A.; et al. A 768Gb 3b/cell 3D-floating-gate NAND Flash memory. In Proceedings of the 2014 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 31 January–4 February 2016; pp. 142–143.

24. Yamashita, R.; Magia, S.; Higuchi, T.; Yoneya, K.; Yamamura, T.; Mizukoshi, H.; Zaitsu, S.; Yamashita, M.; Toyama, S.; Kamae, N.; et al. A 512Gb 3b/cell Flash memory on 64-word-line-layer BiCS technology. In Proceedings of the 2017 IEEE International Solid- State Circuits Conference (ISSCC), San Francisco, CA, USA, 5–9 February 2017; pp. 196–198.

25. Kim, C.; Cho, J.H.; Jeong, W.; Park, I.H.; Park, H.W.; Kim, D.H.; Kang, D.; Lee, S.; Lee, J.S.; Kim, W.; et al. A 512 Gb 3b/cell 64-stacked WL 3D V-NAND Flash memory. In Proceedings of the 2017 IEEE International Solid- State Circuits Conference (ISSCC), San Francisco, CA, USA, 5–9 February 2017; pp. 202–204.

26. Shirota, R. Developments in 3D-NAND Flash technology. In *Advances in Non-Volatile Memory and Storage Technology*; Nishi, Y., Ed.; Elsevier: Amsterdam, The Netherlands, 2014; Chapter 2, pp. 27–74.

27. Aritome, S. *NAND Flash Memory Technologies*; John Wiley & Sons: Hoboken, NJ, USA, 2015.

28. Micheloni, R. (Ed.) *3D Flash Memories*; Springer: Heidelberg/Berlin, Germany, 2016.

29. Lee, J.; Jang, J.; Lim, J.; Shin, Y.G.; Lee, K.; Jung, E. A new ruler on the market: 3D-NAND flash for high-density memory and its technology evolution and challenges on the future. In Proceedings of the 2016 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 3–7 December 2016; pp. 11.2.1–11.2.4.

30. Brewer, J.E.; Gill, M. (Eds.) *Nonvolatile Memory Technology with Emphasis on Flash*; Wiley-IEEE Press: Hoboken, NJ, USA, 2008.

31. Micheloni, R.; Crippa, L.; Marelli, A. (Eds.) *Inside NAND Flash Memories*; Springer: Heidelberg/Berlin, Germany, 2010.

32. Pavan, P.; Bez, R.; Olivo, P.; Zanoni, E. Flash memory cells—An overview. *Proc. IEEE* **1997**, *85*, 1248–1271.

33. Aritome, S.; Satoh, S.; Maruyama, T.; Watanabe, H.; Shuto, S.; Hemink, G.J.; Shirota, R.; Watanabe, S.; Masuoka, F. A 0.67 μm^2 self-aligned shallow trench isolation cell (SA-STI cell) for 3 V-only 256 Mbit NAND EEPROMs. In Proceedings of the 1994 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 11–14 December 1994; pp. 61–64.

34. Naso, G.; Botticchio, L.; Castelli, M.; Cerafogli, C.; Cichocki, M.; Conenna, P.; D'Alessandro, A.; Santis, L.D.; Di Cicco, D.; Di Francesco, W.; et al. A 128 Gb 3 b/cell NAND Flash design using 20 nm planar-cell technology. In Proceedings of the 2013 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 17–21 February 2013; pp. 218–219.

35. Bauer, M.; Alexis, R.; Atwood, G.; Baltar, B.; Fazio, A.; Frary, K.; Hensel, M.; Ishac, M.; Javanifard, J.; Landgraf, M.; et al. A Multilevel-cell 32 Mb Flash memory. In Proceedings of the 1995 IEEE International Solid-State Circuits Conference, San Francisco, CA, USA, 15–17 February 1995; pp. 132–133.

36. Jung, T.S.; Choi, Y.J.; Suh, K.D.; Suh, B.H.; Kim, J.K.; Lim, Y.H.; Koh, Y.N.; Park, J.W.; Lee, K.J.; Park, J.H.; et al. A 117-mm^2 3.3-V only 128-Mb multilevel NAND Flash memory for mass storage applications. *IEEE J. Solid-State Circuits* **1996**, *31*, 1575–1583.

37. Shibata, N.; Maejima, H.; Isobe, K.; Iwasa, K.; Nakagawa, M.; Fujiu, M.; Shimizu, T.; Honma, M.; Hoshi, S.; Kawaai, T.; et al. A 70 nm 16 Gb 16-level-cell NAND flash memory. *IEEE J. Solid-State Circuits* **2008**, *43*, 929–937.

38. Micheloni, R.; Crippa, L. Multi-bit NAND Flash memories for ultra-high density storage devices. In *Advances in Non-Volatile Memory and Storage Technology*; Nishi, Y., Ed.; Woodhead Publishing: Amsterdam, The Netherlands, 2014; Chapter 3, pp. 75–119.

39. Aritome, S.; Shirota, R.; Kirisawa, R.; Endoh, T.; Nakayama, R.; Sakui, K.; Masuoka, F. A reliable bi-polarity write/erase technology in Flash EEPROMs. In Proceedings of the 1990 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 9–12 December 1990; pp. 111–114.

40. Hemink, G.J.; Tanaka, T.; Endoh, T.; Aritome, S.; Shirota, R. Fast and accurate programming method for multi-level NAND EEPROMs. In Proceedings of the 1995 Symposium on VLSI Technology—Digest of Technical Papers, Kyoto, Japan, 6–8 June 1995; pp. 129–130.

41. Micheloni, R.; Marelli, A.; Ravasio, R. *Error-Correction Codes for Non-Volatile Memories*; Springer: Heidelberg/Berlin, Germany, 2008.

42. Zhao, K.; Zhao, W.; Sun, H.; Zhang, T.; Zhang, X.; Zheng, N. LDPC-in-SSD: Making advanced error correction codes work effectively in solid state drives. In Proceedings of the 11th USENIX Conference on File and Storage Technologies, San Jose, CA, USA, 12–15 February 2013; pp. 243–256.

43. Wang, J.; Vakilinia, K.; Chen, T.Y.; Courtade, T.; Dong, G.; Zhang, T.; Shankar, H.; Wesel, R. Enhanced precision through multiple reads for LDPC decoding in Flash memories. *IEEE J. Sel. Areas Commun.* **2014**, *32*, 880–891.

44. Lin, W.; Yen, S.W.; Hsu, Y.C.; Lin, Y.H.; Liang, L.C.; Wang, T.C.; Shih, P.Y.; Lai, K.H.; Cheng, K.Y.; Chang, C.Y. A low power and ultra high reliability LDPC error correction engine with digital signal processing for embedded NAND Flash controller in 40nm CMOS. In Proceedings of the 2014 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers, Honolulu, HI, USA, 9–12 June 2014; pp. 101–102.

45. Haymaker, K.; Kelley, C.A. Structured bit-interleaved LDPC codes for MLC Flash memory. *IEEE J. Sel. Areas Commun.* **2014**, *32*, 870–879.

46. Cho, S.-G.; Kim, D.; Choi, J.; Ha, J. Block-wise concatenated BCH codes for NAND Flash memories. *IEEE Trans. Commun.* **2014**, *62*, 1164–1177.

47. Kim, D.; Ha, J. Quasi-primitive block-wise concatenated BCH codes for NAND Flash memories. In Proceedings of the 2014 IEEE Information Theory Workshop (ITW), Hobart, Australia, 2–5 November 2014; pp. 611–615.

48. Spinner, J.; Rajab, M.; Freudenberger, J. Construction of high-rate generalized concatenated codes for applications in non-volatile Flash memories. In Proceedings of the 2016 IEEE Information Theory Workshop (ITW), Cambridge, UK, 11–14 September 2016.

49. Spinner, J.; Freudenberger, J.; Shavgulidze, S. A soft input decoding algorithm for generalized concatenated codes. *IEEE Trans. Commun.* **2016**, *64*, 3585–3595.

50. Garvin, P.K.; Stanard, H.D. Method and System for Managing Bad Areas in Flash Memories. U.S. Patent 6,260,156 B1, 10 July 2001.

51. Verma, G.; Mielke, N. Reliability performance of ETOX based Flash memories. In Proceedings of the 26th Annual Proceedings, International Reliability Physics Symposium, Monterey, CA, USA, 12–14 April 1988; pp. 158–166.

52. Aritome, S.; Shirota, R.; Hemink, G.; Endoh, T.; Masuoka, F. Reliability issues of Flash memory cells. *Proc. IEEE* **1993**, *81*, 776–788.

53. Cappelletti, P.; Modelli, A. Flash memory reliability. In *Flash Memories*; Cappelletti, P., Golla, C., Olivo, P., Zanoni, E., Eds.; Kluwer: Dordrecht, The Netherlands, 1999.

54. Tao, G.; Valk, H.; Scarpa, A.; van Marwijk, L.; Kuper, F.G.; Dormans, G.J.M.; Verhaar, R.D.J. Reliability aspects of advanced embedded floating-gate non-volatile memories with uniform channel FN tunneling for both program and erase. In Proceedings of the Non-Volatile Semiconductor Memory Workshop, Monterey, CA, USA, 12–15 August 2001; pp. 130–132.

55. Cappelletti, P.; Bez, R.; Modelli, A.; Visconti, A. What we have learned on Flash memory reliability in the last ten years. In Proceedings of the 2004 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 13–15 December 2004; pp. 489–492.

56. Ielmini, D.; Spinelli, A.S.; Lacaita, A.L. Recent developments on Flash memory reliability. *Microelectron. Eng.* **2005**, *80C*, 321–328.

57. Monzio Compagnoni, C.; Spinelli, A.S.; Lacaita, A.L.; Ghetti, A.; Visconti, A. Emerging constraints on NAND Flash memory reliability. In *Nonvolatile Memories: Materials, Devices and Applications*; American Scientific Publishers: Valencia, CA, USA, 2012; pp. 267–293.

58. Zambelli, C.; Olivo, P. SSD reliability. In *Inside Solid-State Drives (SSD)s*; Micheloni, R., Marelli, A., Eshghi, K., Eds.; Springer: Heidelberg/Berlin, Germany, 2013; Chapter 8, pp. 203–231.

59. Park, Y.; Lee, J.; Cho, S.S.; Jin, G.; Jung, E. Scaling and reliability of NAND Flash devices. In Proceedings of the 2014 IEEE International Reliability Physics Symposium (IRPS), Waikoloa, HI, USA, 1–5 June 2014; pp. 2E.1.1–2E.1.4.

60. Mielke, N.; Marquart, T.; Wu, N.; Kessenich, J.; Belgal, H.; Schares, E.; Trivedi, F.; Goodness, E.; Nevill, L.R. Bit error rate in NAND Flash memories. In Proceedings of the IEEE International Reliability Physics Symposium, Phoenix, AZ, USA, 27 April–1 May 1 2008; pp. 9–19.

61. Meza, J.; Wu, Q.; Kumar, S.; Mutlu, O. A large-scale study of Flash memory failures in the field. In Proceedings of the ACM SIGMETRICS—International Conference on Measurement and Modeling of Computer Systems, Portland, OR, USA, 15–19 June 2015; pp. 177–190.

62. Schroeder, B.; Lagisetty, R.; Merchant, A. Flash reliability in production: the expected and the unexpected. In Proceedings of the 14th USENIX Conference on File and Storage Technologies (FAST'16), Santa Clara, CA, USA, 22–25 February 2016; pp. 67–80.

63. Lee, J.D.; Choi, J.H.; Park, D.; Kim, K. Effects of interface trap generation and annihilation on the data retention characteristics of Flash memory cells. *IEEE Trans. Device Mater. Reliab.* **2004**, *4*, 110–117.

64. Itsumi, M. Positive and negative charging of thermally-grown SiO_2 induced by Fowler-Nordheim emission. *J. Appl. Phys.* **1981**, *52*, 3491–3497.

65. Maserjian, J.; Zamani, N. Behavior of the Si/SiO_2 interface observed by Fowler-Nordheim tunneling. *J. Appl. Phys.* **1982**, *53*, 559–567.

66. Maserjian, J.; Zamani, N. Observation of positively charged state generation near the Si/SiO_2 interface during Fowler-Nordheim tunneling. *J. Vac. Sci. Technol.* **1982**, *20*, 743–746.

67. Lai, S.K. Interface trap generation in silicon dioxide when electrons are captured by trapped holes. *J. Appl. Phys.* **1983**, *54*, 2540–2546.

68. Olivo, P.; Riccò, B.; Sangiorgi, E. Electron trapping/detrapping within thin SiO_2 films in the high field tunneling regime. *J. Appl. Phys.* **1983**, *54*, 5267–5276.

69. Chen, I.C.; Holland, S.; Hu, C. Electron-trap generation by recombination of electrons and holes in SiO_2. *J. Appl. Phys.* **1987**, *61*, 4544–4548.

70. Euzent, B.; Jeng, C.; Lee, J.; Boruta, N. Reliability aspects of a floating gate E^2PROM. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Orlando, FL, USA, 7–9 April 1981; pp. 11–16.

71. Witters, J.S.; Groeseneken, G.; Maes, H.E. Degradation of tunnel-oxide floating-gate EEPROM devices and the correlation with high field-current-induced degradation of thin gate oxides. *IEEE Trans. Electron Devices* **1989**, *36*, 1663–1682.

72. Park, Y.B.; Schroder, D.K. Degradation of thin tunnel gate oxide under constant Fowler-Nordheim current stress for a Flash EEPROM. *IEEE Trans. Electron Devices* **1998**, *45*, 1361–1368.

73. Oyama, K.; Shirai, H.; Kodama, N.; Kanamori, K.; Saitoh, K.; Hisamune, Y.S.; Okazawa, T. A novel erasing technology for 3.3 V Flash memory with 64 Mb capacity and beyond. In Proceedings of the 1992 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 13–16 December 1992; pp. 607–610.

74. Fayrushin, A.; Seol, K.; Na, J.; Hur, S.; Choi, J.; Kim, K. The new program/erase cycling degradation mechanism of NAND Flash memory devices. In Proceedings of the 2009 International Electron Devices Meeting (IEDM), Baltimore, MD, USA, 7–9 December 2009; pp. 823–826.

75. Fayrushin, A.; Lee, C.H.; Park, Y.; Choi, J.H.; Chung, C. Unified endurance degradation model of floating gate NAND Flash memory. *IEEE Trans. Electron Devices* **2013**, *60*, 2031–2037.

76. Shirota, R.; Yang, B.J.; Chiu, Y.Y.; Chen, H.T.; Ng, S.F.; Wang, P.Y.; Chang, J.H.; Kurachi, I. New method to analyze the shift of floating gate charge and generated tunnel oxide trapped charge profile in NAND Flash memory by program/erase endurance. *IEEE Trans. Electron Devices* **2015**, *62*, 114–120.

77. Chang, Y.H.; Chang, L.P. Efficient wear leveling in NAND Flash memory. In *Inside Solid State Drives (SSDs)*; Micheloni, R., Marelli, A., Eshghi, K., Eds.; Springer: Heidelberg/Berlin, Germany, 2013; Chapter 9, pp. 233–257.

78. Ralls, K.S.; Skocpol, W.J.; Jackel, L.D.; Howard, R.E.; Fetter, L.A.; Epworth, R.W.; Tennant, D.M. Discrete resistance switching in submicrometer silicon inversion layers: individual interface traps and low-frequency (1/f?) noise. *Phys. Rev. Lett.* **1984**, *52*, 228–231.

79. Kirton, M.J.; Uren, M.J. Noise in solid-state microstructures: a new perspective on individual defects, interface states and low-frequency (1/f) noise. *Adv. Phys.* **1989**, *38*, 367–468.

80. Hung, K.K.; Ko, P.K.; Hu, C.; Cheng, Y.C. Flicker noise characterization of advanced MOS technologies. In Proceedings of the 1988 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 11–14 December 1988; pp. 34–37.

81. Hung, K.K.; Ko, P.K.; Hu, C.; Cheng, Y.C. Random telegraph noise of deep-submicrometer MOSFET's. *IEEE Electron Device Lett.* **1990**, *11*, 90–92.

82. Ohata, A.; Toriumi, A.; Iwase, M.; Natori, K. Observation of random telegraph signals: anomalous nature of defects at the Si/SiO_2 interface. *J. Appl. Phys.* **1990**, *68*, 200–204.

83. Shi, Z.M.; Mieville, J.P.; Barrier, J.; Dutoit, M. Low frequency noise and quantum transport in 0.1 μm n-MOSFET's. In Proceedings of the 1991 International Electron Devices Meeting (IEDM), Washington, DC, USA, 8–11 December 1991; pp. 363–366.

84. Fang, P.; Hung, K.K.; Ko, P.K.; Hu, C. Hot-electron-induced traps studied through the random telegraph noise. *IEEE Electron Device Lett.* **1991**, *12*, 273–275.

85. Simoen, E.; Diericks, B.; Claeys, C.L.; Declerck, G.J. Explaining the amplitude of RTS noise in submicrometer MOSFET's. *IEEE Trans. Electron Devices* **1992**, *39*, 422–429.

86. Tsai, M.H.; Ma, T.P.; Hook, T.B. Channel length dependence of random telegraph signal in sub-micron MOSFET's. *IEEE Electron Device Lett.* **1994**, *15*, 504–506.

87. Uren, M.J.; Kirton, M.J.; Collins, S. Anomalous telegraph noise in small-area silicon metal-oxide-semiconductor field-effect transistors. *Phys. Rev. B* **1988**, *37*, 8346–8350.

88. Kirton, M.J.; Uren, M.J.; Collins, S.; Schultz, M.; Karmann, A.; Scheffer, K. Individual defects at the $Si : SiO_2$ interface. *Semicond. Sci. Technol.* **1989**, *4*, 1116–1126.

89. Nakamura, H.; Yasuda, N.; Taniguchi, K.; Hamaguchi, C.; Toriumi, A. Existence of double-charged oxide traps in submicron MOSFET's. *Jpn. J. Appl. Phys.* **1989**, *28*, L2057–L2060.

90. Tega, N.; Miki, H.; Osabe, T.; Kotabe, A.; Otsuga, K.; Kurata, H.; Kamohara, S.; Tokami, K.; Ikeda, Y.; Yamada, R. Anomalously large threshold voltage fluctuation by complex random telegraph signal in floating gate Flash memory. In Proceedings of the 2006 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 11–13 December 2006; pp. 491–494.

91. Fantini, P.; Ghetti, A.; Marinoni, A.; Ghidini, G.; Visconti, A.; Marmiroli, A. Giant random telegraph signals in nanoscale floating-gate devices. *IEEE Electron Device Lett.* **2007**, *28*, 1114–1116.

92. Tega, N.; Miki, H.; Yamaoka, M.; Kume, H.; Mine, T.; Ishida, T.; Mori, Y.; Yamada, R.; Torii, K. Impact of threshold voltage fluctuation due to random telegraph noise on scaled-down SRAM. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Phoenix, AZ, USA, 27 April–1 May 1 2008; pp. 541–546.

93. Toh, S.O.; Tsukamoto, Y.; Guo, Z.; Jones, L.; King Liu, T.J.; Nikolic, B. Impact of random telegraph signals on V_{min} in 45 nm SRAM. In Proceedings of the 2009 International Electron Devices Meeting (IEDM), Baltimore, MD, USA, 7–9 December 2009; pp. 767–770.

94. Tanizawa, M.; Ohbayashi, S.; Okagaki, T.; Sonoda, K.; Eikyu, K.; Hirano, Y.; Ishikawa, K.; Tsuchiya, O.; Inoue, Y. Application of a statistical compact model for random telegraph noise to scaled-SRAM Vmin analysis. In Proceedings of the 2010 Symposium on VLSI Technology (VLSI-Technology), Honolulu, HI, USA, 15–17 June 2010; pp. 95–96.

95. Kurata, H.; Otsuga, K.; Kotabe, A.; Kajiyama, S.; Osabe, T.; Sasago, Y.; Narumi, S.; Tokami, K.; Kamohara, S.; Tsuchiya, O. The impact of random telegraph signals on the scaling of multilevel Flash memories. In Proceedings of the 2006 Symposium on VLSI Technology (VLSI-Technology), Honolulu, HI, USA, 13–15 June 2006; pp. 112–113.

96. Gusmeroli, R.; Monzio Compagnoni, C.; Riva, A.; Spinelli, A.S.; Lacaita, A.L.; Bonanomi, M.; Visconti, A. Defects spectroscopy in SiO_2 by statistical random telegraph noise analysis. In Proceedings of the 2006 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 11–13 December 2006; pp. 483–486.

97. Kurata, H.; Otsuga, K.; Kotabe, A.; Kajiyama, S.; Osabe, T.; Sasago, Y.; Narumi, S.; Tokami, K.; Kamohara, S.; Tsuchiya, O. Random telegraph signal in Flash memory: Its impact on scaling of multilevel Flash memory beyond the 90-nm node. *IEEE J. Solid-State Circuits* **2007**, *42*, 1362–1369.

98. Fukuda, K.; Shimizu, Y.; Amemiya, K.; Kamoshida, M.; Hu, C. Random telegraph noise in Flash memories—Model and technology scaling. In Proceedings of the 2007 International Electron Devices Meeting (IEDM), Washington, DC, USA, 10–12 December 2007; pp. 169–172.

99. Miki, H.; Osabe, T.; Tega, N.; Kotabe, A.; Kurata, H.; Tokami, K.; Ikeda, Y.; Kamohara, S.; Yamada, R. Quantitative analysis of random telegraph signals as fluctuations of threshold voltages in scaled Flash memory cells. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Phoenix, AZ, USA, 15–19 April 2007; pp. 29–35.

100. Kandiah, K.; Deighton, M.O.; Whiting, F.B. A physical model for random telegraph signal currents in semiconductor devices. *J. Appl. Phys.* **1989**, *62*, 937–948.

101. Mueller, H.H.; Schulz, M. Conductance modulation of submicrometer metal-oxide-semiconductor field effect transistors by single-electron trapping. *J. Appl. Phys.* **1996**, *79*, 4178–4180.

102. Lust, L.M.; Kakalios, J. Dynamic percolation model of conductance fluctuations in hydrogenated amorphous silicon. *Phys. Rev. Lett.* **1995**, *75*, 2192–2195.

103. Brews, J.R. Theory of the carrier-density fluctuations in an IGFET near threshold. *J. Appl. Phys.* **1975**, *46*, 2181–2192.

104. Arnold, E. Conduction mechanisms in bandtails at the Si-SiO_2 interface. *Surf. Sci.* **1976**, *58*, 60–70.

105. Wong, H.S.; Taur, Y. Three-dimensional "atomistic" simulation of discrete random dopant distribution effects in sub-0.1 μm MOSFET's. In Proceedings of the 1993 International Electron Devices Meeting (IEDM), Washington, DC, USA, 5–8 December 1993; pp. 705–708.

106. Asenov, A. Random dopant induced threshold voltage lowering and fluctuations in sub-0.1 μm MOSFET's: A 3-D "atomistic" simulation study. *IEEE Trans. Electron Devices* **1998**, *45*, 2505–2513.

107. Asenov, A.; Brown, A.R.; Davies, J.H.; Saini, S. Hierarchical approach to "atomistic" 3-D MOSFET simulation. *IEEE Trans. Comput.-Aided Des.* **1999**, *18*, 1558–1565.

108. Asenov, A.; Saini, S. Suppression of random dopant-induced threshold voltage fluctuations in sub-0.1 μm MOSFET's with epitaxial and δ-doped channels. *IEEE Trans. Electron Devices* **1999**, *46*, 1718–1724.

109. Vandamme, L.K.J.; Sodini, D.; Gingl, Z. On the anomalous behavior of the relative amplitude of RTS noise. *Solid-State Electron.* **1998**, *42*, 901–905.

110. Asenov, A.; Balasubramaniam, R.; Brown, A.R.; Davies, J.H.; Saini, S. Random telegraph signal amplitudes in sub 100 nm (decanano) MOSFETs: A 3D "atomistic" simulation study. In Proceedings of the 2000 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 10–13 December 2000; pp. 279–282.

111. Asenov, A.; Balasubramaniam, R.; Brown, A.R.; Davies, J.H. RTS amplitudes in decananometer MOSFETs: 3-D simulation study. *IEEE Trans. Electron Devices* **2003**, *50*, 839–845.

112. Ghetti, A.; Bonanomi, M.; Monzio Compagnoni, C.; Spinelli, A.S.; Lacaita, A.L.; Visconti, A. Physical modeling of single-trap RTS statistical distribution in Flash memories. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Phoenix, AZ, USA, 27 April–1 May 1 2008; pp. 610–615.

113. Ghetti, A.; Monzio Compagnoni, C.; Spinelli, A.S.; Visconti, A. Comprehensive analysis of random telegraph noise instability and its scaling in deca-nanometer Flash memories. *IEEE Trans. Electron Devices* **2009**, *56*, 1746–1752.

114. Alexander, C.L.; Brown, A.R.; Watling, J.R.; Asenov, A. Impact of single charge trapping nano-MOSFETs—Electrostatics versus transport effects. *IEEE Trans. Nanotechnol.* **2005**, *4*, 339–344.

115. Alexander, C.; Brown, A.R.; Watling, J.R.; Asenov, A. Impact of scattering in "atomistic" device simulations. *Solid-State Electron.* **2005**, *49*, 733–739.

116. Kovac, U.; Alexander, C.; Roy, G.; Riddet, C.; Cheng, B.; Asenov, A. Hierarchical simulation of statistical variability: from 3-D MC with "ab initio" ionized impurity scattering to statistical compact models. *IEEE Trans. Electron Devices* **2010**, *57*, 2418–2426.

117. Kovac, U.; Alexander, C.; Asenov, A. Statistical estimation of electrostatic and transport contributions to device parameter variation. In Proceedings of the International Workshop on Computational Electronics (IWCE), Pisa, Italy, 27–29 October 2010; pp. 235–238.

118. Ancona, M.G.; Tiersten, H.F. Macroscopic physics of the silicon inversion layer. *Phys. Rev. B* **1987**, *35*, 7959–7965.

119. Ancona, M.G.; Yu, Z.; Dutton, R.W.; Voorde, P.J.V.; Cao, M.; Vook, D. Density-gradient analysis of MOS tunneling. *IEEE Trans. Electron Devices* **2000**, *47*, 2310–2319.

120. Asenov, A.; Slavcheva, G.; Brown, A.R.; Davies, J.H.; Saini, S. Increase in the random dopant induced threshold fluctuations and lowering in sub-100 nm MOSFETs due to quantum effects: A 3-D density-gradient simulation study. *IEEE Trans. Electron Devices* **2001**, *48*, 722–729.

121. Ghetti, A.; Monzio Compagnoni, C.; Biancardi, F.; Lacaita, A.L.; Beltrami, S.; Chiavarone, L.; Spinelli, A.S.; Visconti, A. Scaling trends for random telegraph noise in deca-nanometer Flash memories. In Proceedings of the 2008 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 15–17 December 2008; pp. 835–838.

122. Monzio Compagnoni, C.; Gusmeroli, R.; Spinelli, A.S.; Lacaita, A.L.; Bonanomi, M.; Visconti, A. Statistical model for random telegraph noise in Flash memories. *IEEE Trans. Electron Devices* **2008**, *55*, 388–395.

123. Monzio Compagnoni, C.; Spinelli, A.S.; Beltrami, S.; Bonanomi, M.; Visconti, A. Cycling effect on the random telegraph noise instabilities of NOR and NAND Flash arrays. *IEEE Electron Device Lett.* **2008**, *29*, 941–943.

124. Spinelli, A.S.; Monzio Compagnoni, C.; Gusmeroli, R.; Ghidotti, M.; Visconti, A. Investigation of the random telegraph noise instability in scaled Flash memory arrays. *Jpn. J. Appl. Phys.* **2008**, *47*, 2598–2601.

125. Seidel, K.; Hoffmann, R.; Löhr, D.A.; Melde, T.; Czernohorsky, M.; Paul, J.; Beug, M.F.; Beyer, V. Analysis of trap mechanisms responsible for random telegraph noise and erratic programming on sub-50 nm floating gate flash memories. In Proceedings of the 2009 10th Annual Non-Volatile Memory Technology Symposium (NVMTS), Portland, OR, USA, 25–28 October 2009; pp. 67–71.

126. Kwon, W.; Song, Y.H.; Cai, Y.; Ryu, W.; Jang, Y.; Shin, S.; Jun, J.; Lee, S.A.; Park, C.K.; Lee, W.S. 3-D channel structure Flash having short channel effect immunity and low random telegraph noise. In Proceedings of the 2008 10th Annual Non-Volatile Memory Technology Symposium (NVMTS), Pacific Grove, CA, USA, 11–14 November 2008; pp. 20–21.

127. Wang, R.V.; Lee, Y.H.; Lu, Y.L.R.; McMahon, W.; Hu, S.; Ghetti, A. Shallow trench isolation edge effect on random telegraph signal noise and implications for Flash memory. *IEEE Trans. Electron Devices* **2009**, *56*, 2107–2113.

128. Lee, C.H.; Yang, I.C.; Lee, C.; Cheng, C.H.; Chong, L.H.; Chen, K.F.; Huang, J.S.; Ku, S.H.; Zous, N.K.; Huang, I.J.; et al. Junction optimization for Reliability issues in floating gate NAND Flash cells. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Monterey, CA, USA, 10–14 April 2011; pp. 6B.3.1–6B.3.5.

129. Kim, T.; He, D.; Morinville, K.; Sarpatwari, K.; Millemon, B.; Goda, A.; Kessenich, J. Tunnel oxide nitridation effect on the evolution of V_t instabilities (RTS/QED) and defect characterization for sub-40 nm Flash memory. *IEEE Electron Device Lett.* **2011**, *32*, 999–1001.

130. Kim, T.; Franklin, N.; Srinivasan, C.; Kalavade, P.; Goda, A. Extreme short-channel effect on RTS and inverse scaling behavior: source–drain implantation effect in 25-nm NAND Flash memory. *IEEE Electron Dev. Lett.* **2011**, *32*, 1185–1187.

131. Ghetti, A.; Amoroso, S.M.; Mauri, A.; Monzio Compagnoni, C. Impact of nonuniform doping on random telegraph noise in Flash memory devices. *IEEE Trans. Electron Devices* **2012**, *59*, 309–315.

132. Amoroso, S.M.; Ghetti, A.; Brown, A.R.; Mauri, A.; Monzio Compagnoni, C.; Asenov, A. Impact of cell shape on random telegraph noise in decananometer Flash memories. *IEEE Trans. Electron Devices* **2012**, *59*, 2774–2779.

133. Kim, S.; Lee, M.; Choi, G.B.; Lee, J.; Lee, Y.; Cho, M.; Ahn, K.O.; Kim, J. RTS noise reduction of 1Y-nm floating gate NAND Flash memory using process optimization. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Monterey, CA, USA, 19–23 April 2015; pp. MY.8.1–MY.8.4.

134. Tomita, T.; Miyaji, K. Channel doping concentration and cell program state dependence on random telegraph noise spatial amd statistical distribution in 30 nm NAND flash memory. *Jap. J. Appl. Phys.* **2015**, *54*, 04DD02-1–04DD02-6.

135. Joe, S.M.; Yi, J.H.; Park, S.K.; Kwon, H.I.; Lee, J.H. Position-dependent threshold-voltage variation by random telegraph noise in NAND Flash memory strings. *IEEE Electron Device Lett.* **2010**, *31*, 635–637.

136. Joe, S.M.; Yi, J.H.; Park, S.K.; Shin, H.; Park, B.G.; Park, Y.J.; Lee, J.H. Threshold voltage fluctuation by random telegraph noise in floatin gate NAND Flash memory string. *IEEE Trans. Electron Devices* **2011**, *58*, 67–73.

137. Joe, S.M.; Jeong, M.K.; Jo, B.S.; Han, K.R.; Park, S.K.; Lee, J.H. The effect of adjacent bit-line cell interference on random telegraph noise in NAND Flash memory cell strings. *IEEE Trans. Electron Devices* **2012**, *59*, 3568–3573.

138. Miccoli, C.; Paolucci, G.M.; Monzio Compagnoni, C.; Spinelli, A.S.; Goda, A. Cycling pattern and read/bake conditions dependence of random telegraph noise in decananometer NAND Flash arrays. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Monterey, CA, USA, 19–23 April 2015; pp. MY.9.1–MY.9.6.

139. Miki, H.; Tega, N.; Yamaoka, M.; Frank, D.J.; Bansal, A.; Kobayashi, M.; Cheng, K.; D'Emic, C.P.; Ren, Z.; Wu, S.; et al. Statistical measurement of random telegraph noise and its impact in scaled-down high-k/metal-gate MOSFETs. In Proceedings of the 2012 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 10–13 December 2012; pp. 450–453.

140. Grasser, T. Stochastic charge trapping in oxides: From random telegraph noise to bias temperature instabilities. *Microelectron. Reliab.* **2012**, *52*, 39–70.

141. Machlup, S. Noise in semiconductors: spectrum of a two-parameter random signal. *J. Appl. Phys.* **1954**, *25*, 341–343.

142. Yuzhelevski, Y.; Yuzhelevski, M.; Jung, G. Random telegraph noise analysis in time domain. *Rev. Sci. Instrum.* **2000**, *71*, 1681–1688.

143. Campbell, J.P.; Qin, J.; Cheung, K.P.; Yu, L.C.; Suehle, J.S.; Oates, A.; Sheng, K. Random telegraph noise in highly scaled nMOSFETs. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Montreal, QC, Canada, 26–30 April 2009; pp. 382–388.

144. Campbell, J.P.; Yu, L.C.; Cheung, K.P.; Qin, J.; Suehle, J.S.; Oates, A.; Sheng, K. Large random telegraph noise in sub-threshold operation of nano-scale nMOSFETs. In Proceedings of the 2009 IEEE International Conference on IC Design & Technology (ICICDT), Austin, TX, USA, 18–20 May 2009; pp. 17–20.

145. Castellani, N.; Monzio Compagnoni, C.; Mauri, A.; Spinelli, A.S.; Lacaita, A.L. Three-dimensional electrostatics- and atomistic doping-induced variability of RTN time constants in nanoscale MOS devices—Part I: Physical investigation. *IEEE Trans. Electron Devices* **2012**, *59*, 2488–2494.

146. Monzio Compagnoni, C.; Castellani, N.; Mauri, A.; Spinelli, A.S.; Lacaita, A.L. Three-dimensional electrostatics- and atomistic doping-induced variability of RTN time constants in nanoscale MOS devices—Part II: Spectroscopic implications. *IEEE Trans. Electron Devices* **2012**, *59*, 2495–2500.

147. Adamu-Lema, F.; Monzio Compagnoni, C.; Amoroso, S.M.; Castellani, N.; Gerrer, L.; Markov, S.; Spinelli, A.S.; Lacaita, A.L.; Asenov, A. Accuracy and issues of the spectroscopic analysis of RTN traps in nanoscale MOSFETs. *IEEE Trans. Electron Devices* **2013**, *60*, 833–839.

148. Grasser, T.; Kaczer, B.; Goes, W.; Reisinger, H.; Aichinger, T.; Hehenberger, P.; Wagner, P.J.; Schanovsky, F.; Franco, J.; Luque, M.T.; et al. The paradigm shift in understanding the bias temperature instability: From reaction-diffusion to switching oxide traps. *IEEE Trans. Electron Devices* **2011**, *58*, 3652–3666.

149. Nagumo, T.; Takeuchi, K.; Hase, T.; Hayashi, Y. Statistical characterization of trap position, energy, amplitude and time constants by RTN measurement of multiple individual traps. In Proceedings of the 2010 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 6–8 December 2010; pp. 628–631.

150. Abe, K.; Teramoto, A.; Sugawa, S.; Ohmi, T. Understanding of traps causing random telegraph noise based on experimentally extracted time constants and amplitude. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Monterey, CA, USA, 10–14 April 2011; pp. 381–386.

151. Yonezawa, A.; Teramoto, A.; Obara, T.; Kuroda, R.; Sugawa, S.; Ohmi, T. The study of time constant analysis in random telegraph noise at the subthreshold voltage region. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Monterey, CA, USA, 14–18 April 2013; pp. XT.11.1–XT.11.6.

152. Guo, S.; Ren, P.; Wang, R.; Yu, Z.; Luo, M.; Zhang, X.; Huang, R. A new efficient method for characterizing time constants of switching oxide traps. In Proceedings of the 2014 IEEE International Reliability Physics Symposium (IRPS), Waikoloa, HI, USA, 1–5 June 2014; pp. XT.14.1–XT.14.4.

153. Grasser, T.; Rott, K.; Reisinger, H.; Waltl, M.; Franco, J.; Kaczer, B. A unified perspective of RTN and BTI. In Proceedings of the 2014 IEEE International Reliability Physics Symposium (IRPS), Waikoloa, HI, USA, 1–5 June 2014; pp. 4A.5.1–4A.5.7.

154. Chen, J.; Higashi, Y.; Kato, K.; Mitani, Y. Further understandings on random telegraph signal noise through comprehensive studies on large time constant variation and its strong correlations to thermal activation energies. In Proceedings of the 2014 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers, Honolulu, HI, USA, 9–12 June 2014; pp. 164–165.

155. Chen, M.J.; Tu, K.C.; Wang, H.H.; Chen, C.L.; Lai, S.Y.; Liu, Y.S. A statistical model for the headed and tail distributions of random telegraph signal magnitudes in nanoscale MOSFETs. *IEEE Trans. Electron Devices* **2014**, *61*, 2495–2502.

156. Chen, M.J.; Tu, K.C.; Chuang, L.Y.; Wang, H.H. Graphically transforming Mueller-Schulz percolation criteria to random telegraph signal magnitudes in scaled FETs. *IEEE Electron Dev. Lett.* **2015**, *36*, 217–219.

157. Monzio Compagnoni, C.; Ghidotti, M.; Lacaita, A.L.; Spinelli, A.S.; Visconti, A. Random telegraph noise effect on the programmed threshold-voltage distribution of Flash memories. *IEEE Electron Dev. Lett.* **2009**, *30*, 984–986.

158. Monzio Compagnoni, C.; Spinelli, A.S.; Gusmeroli, R.; Beltrami, S.; Ghetti, A.; Visconti, A. Ultimate accuracy for the NAND Flash program algorithm due to the electron injection statistics. *IEEE Trans. Electron Devices* **2008**, *55*, 2695–2702.

159. Monzio Compagnoni, C.; Gusmeroli, R.; Spinelli, A.S.; Visconti, A. Analytical model for the electron-injection statistics during programming of nanoscale NAND Flash memories. *IEEE Trans. Electron Devices* **2008**, *55*, 3192–3199.

160. Alexander, R.M. Accelerated testing in FAMOS deviecs—8K EPROM. In Proceedings of the 16th Annual IEEE International Reliability Physics Symposium, San Diego, CA, USA, 18–20 April 1978; pp. 229–232.

161. Horiuchi, M.; Katto, H. FCAT—A low-voltage high-speed alterable *n*-channel nonvolatile memory device. *IEEE Trans. Electron Devices* **1979**, *26*, 914–918.

162. Wang, S.T. Charge retention of floating-gate transistors under applied bias conditions. *IEEE Trans. Electron Devices* **1980**, *27*, 297–299.

163. Bhattacharyya, A. Modelling of write/erase and charge retention characteristics of floating gate EEPROM devices. *Solid-State Electron.* **1984**, *26*, 899–906.

164. Papadas, C.; Pananakakis, G.; Ghibaudo, G.; Riva, C.; Pio, F.; Ghezzi, P. Modeling of the intrinc retention characteristics of FLOTOX EEPROM cells under elevated temperature conditions. *IEEE Trans. Electron Devices* **1995**, *42*, 678–681.

165. Watanabe, H.; Ishihara, T.; Matsunaga, Y.; Matsuzawa, K.; Matsushita, D.; Muraoka, K. Numerical study of data retention due to direct tunneling for nonvolatile memory cell. *IEEE Trans. Electron Devices* **2005**, *52*, 955–961.

166. Dockerty, R.C. Degradation mechanisms in rewritable n-channel FAMOS devices. In Proceedings of the 13th Annual IEEE International Reliability Physics Symposium, Downey, CA, USA, 1–3 April 1975; pp. 6–9.

167. Baglee, D.A.; Smayling, M.C. The effects of write/erase cycling on data loss in EEPROMs. In Proceedings of the 1985 IEEE International Electron Devices Meeting (IEDM), New York, NY, USA, 1–4 December 1985; pp. 624–626.

168. Olivo, P.; Nguyen, T.N.; Riccò, B. High-field-induced degradation in ultra-thin SiO_2 films. *IEEE Trans. Electron Devices* **1988**, *35*, 2259–2267.

169. Naruke, K.; Taguchi, S.; Wada, M. Stress induced leakage current limiting to scale down EEPROM tunnel oxide thickness. In Proceedings of the 1988 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 11–14 December 1988; pp. 424–427.

170. Moazzami, R.; Hu, C. Stress-induced current in thin silicon dioxide films. In Proceedings of the 1992 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 13–16 December 1992; pp. 139–142.

171. Ghetti, A.; Alam, M.A.; Bude, J.; Venturi, F. Assessment of quantum yield experiments via full band Monte Carlo simulations. In Proceedings of the 1997 International Electron Devices Meeting (IEDM), Washington, DC, USA, 7–10 December 1997; pp. 873–876.

172. Sakakibara, K.; Ajika, N.; Eikyu, K.; Ishikawa, K.; Miyoshi, H. A quantitative analysis of time-decay reproducible stress-induced leakage current in SiO_2 films. *IEEE Trans. Electron Devices* **1997**, *44*, 1002–1007.

173. Chou, A.I.; Lai, K.; Kumar, K.; Chowdhury, P.; Lee, J.C. Modeling of stress-induced leakage current in ultrathin oxides with the trap-assisted tunneling mechanism. *Appl. Phys. Lett.* **1997**, *70*, 3407–3409.

174. Kamohara, S.; Park, D.; Hu, C. Deep-trap SILC (Stress-Induced Leakage Current) model for nominal and weak oxides. In Proceedings of the IEEE International Reliability Physics Symposium, Reno, NV, USA, 31 March–2 April 1998; pp. 57–61.

175. Riccò, B.; Torelli, G.; Lanzoni, M.; Manstretta, A.; Maes, H.E.; Montanari, D.; Modelli, A. Nonvolatile multilevel memories for digital applications. *Proc. IEEE* **1998**, *86*, 2399–2421.

176. Larcher, L.; Paccagnella, A.; Scarpa, A.; Ghidini, G. A new model of tunneling current and SILC in ultra-thin oxides. In Proceedings of the 1988 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 11–14 December 1988; pp. 901–904.

177. Takagi, S.; Yasuda, N.; Toriumi, A. A new *I–V* model for stress-induced leakage current including inelastic tunneling. *IEEE Trans. Electron Devices* **1999**, *46*, 348–354.

178. Ielmini, D.; Spinelli, A.S.; Lacaita, A.L.; Ghidini, G. Impact ionization and stress-induced leakage current in thin gate oxides. In Proceedings of the First European Workshop on Ultimate Integration of Silicon, Grenoble, France, 20–21 January 2000; pp. 85–88.

179. Ielmini, D.; Spinelli, A.S.; Lacaita, A.L. Experimental evidence for recombination-assisted leakage in thin oxides. *Appl. Phys. Lett.* **2000**, *76*, 1719–1721.

180. Satake, H.; Toriumi, A. Common origin for stress-induced leakage current and electron trap generation in SiO_2. *Appl. Phys. Lett.* **1995**, *67*, 3489–3490.

181. Alers, G.B.; Weir, B.E.; Alam, M.A.; Timp, G.L.; Sorch, T. Trap assisted tunneling as a mechanism of degradation and noise. In Proceedings of the IEEE International Reliability Physics Symposium, Reno, NV, USA, 31 March–2 April 1998; pp. 76–79.

182. James, D. Recent advances in memory technology. In Proceedings of the 2013 24th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC), Saratoga Springs, NY, USA, 14–16 May 2013; pp. 386–395.

183. Yamada, S.; Amemiya, K.; Yamane, T.; Hazama, H.; Hashimoto, K. Non-uniform current flow through thin oxide after Fowler-Nordheim current stress. In Proceedings of the IEEE International Reliability Physics Symposium, Dalas, TX, USA, 30 April 30–2 May 1996; pp. 108–112.

184. Arai, F.; Maruyama, T.; Shirota, R. Extended data retention process technology for highly reliable Flash EEPROMs of 10^6 to 10^7 W/E cycles. In Proceedings of the IEEE International Reliability Physics Symposium, Reno, NV, USA, 31 March–2 April 1998; pp. 378–382.

185. Sakura, T.; Utsunomiya, H.; Kamakura, Y.; Taniguchi, K. A detailed study of soft- and pre-soft-breakdowns in small geometry MOS structures. In Proceedings of the 1988 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 11–14 December 1988; pp. 183–186.

186. Tsuji, N.; Sakakibara, K.; Ajika, N.; Miyoshi, H. Microscopic and statistical approach to SILC characteristics. In Proceedings of the 1998 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers, Honolulu, HI, USA, 9–11 June 1998; pp. 196–197.

187. Kameyama, H.; Okuyama, Y.; Kamohara, S.; Kubota, K.; Kume, H.; Okuyama, K.; Manabe, Y.; Nozoe, A.; Uchida, H.; Hidaka, M.; et al. A new data retention mechanism after endurance stress on Flash memory. In Proceedings of the IEEE International Reliability Physics Symposium, San Jose, CA, USA, 10–13 April 2000; pp. 194–199.

188. Kumagai, Y.; Teramoto, A.; Inatsuka, T.; Kuroda, R.; Suwa, T.; Sugawa, S.; Ohmi, T. Evaluation for anomalous stress-induced leakage current of gate SiO_2 fimls using array test pattern. *IEEE Trans. Electron Devices* **2011**, *58*, 3307–3313.

189. Modelli, A.; Gilardoni, F.; Ielmini, D.; Spinelli, A.S. A new conduction mechanism for the anomalous bits in thin oxide Flash EEPROMs. In Proceedings of the IEEE International Reliability Physics Symposium, Orlando, FL, USA, 30 April–3 May 2001; pp. 61–66.

190. Inatsuka, T.; Kumagai, Y.; Kuroda, R.; Teramoto, A.; Sugawa, S.; Ohmi, T. Recovery characteristics of anomalous stress induced leakage current of 5.6nm oxide films. *Jpn. J. Appl. Phys.* **2012**, *51*, 04DC02.

191. Ielmini, D.; Spinelli, A.S.; Lacaita, A.L.; Modelli, A. A new two-trap tunneling model for the anomalous SILC in flash memories. *Microelectron. Eng.* **2001**, *59*, 189–195.

192. Ielmini, D.; Spinelli, A.S.; Lacaita, A.L.; Modelli, A. Modeling of anomalous SILC in flash memories based on tunneling at multiple defects. *Solid-State Electron.* **2002**, *46*, 1749–1756.

193. Ielmini, D.; Spinelli, A.S.; Lacaita, A.L.; Modelli, A. Statistical modeling of reliability and scaling projections for Flash memories. In Proceedings of the 2001 International Electron Devices Meeting (IEDM'01), Washington, DC, USA, 2–5 December 2001; pp. 703–706.

194. Ielmini, D.; Spinelli, A.S.; Lacaita, A.L.; Modelli, A. A statistical model for SILC in Flash memories. *IEEE Trans. Electron Devices* **2002**, *49*, 1955–1961.

195. Kuhn, P.J.; Hoefler, A.; Harp, T.; Hornung, B.; Paulsen, R.; Burnett, D.; Higman, J.M. A reliability methodology for low temperature data retention in floating gate non-volatile memories. In Proceedings of the IEEE International Reliability Physics Symposium, Orlando, FL, USA, 30 April–3 May 2001; pp. 266–270.

196. Hoefler, A.; Higman, J.M.; Harp, T.; Kuhn, P.J. Statistical modeling of the program/erase cycling acceleration of low temperature data retention in floating gate nonvolatile memories. In Proceedings of the IEEE International Reliability Physics Symposium, Dallas, TX, USA, 7–11 April 2002; pp. 21–25.

197. Schuler, F.; Degraeve, R.; Hendrickx, P.; Wellekens, D. Physical description of anomalous charge loss in floating gate based NVM's and identification of its dominant parameter. In Proceedings of the IEEE International Reliability Physics Symposium, Dallas, TX, USA, 7–11 April 2002; pp. 26–33.

198. Degraeve, R.; Schuler, F.; Lorenzini, M.; Wellekens, D.; Hendrickx, P.; Houdt, J.V.; Haspeslagh, L.; Groeseneken, G.; Tempel, G. Analytical model for failure rate prediction due to anomalous charge loss of flash memories. In Proceedings of the 2001 International Electron Devices Meeting (IEDM'01), Washington, DC, USA, 2–5 December 2001; pp. 699–702.

199. Belgal, H.P.; Righos, N.; Kalastirsky, I.; Peterson, J.J.; Shiner, R.; Mielke, N. A new reliability model for post-cycling charge retention of flash memories. In Proceedings of the IEEE International Reliability Physics Symposium, Dallas, TX, USA, 7–11 April 2002; pp. 7–20.

200. Larcher, L.; Pavan, P. Statistical simulations to inspect and predict data retention and program disturbs in Flash memories. In Proceedings of the 2003 International Electron Devices Meeting (IEDM), Washington, DC, USA, 8–10 December 2003; pp. 165–168.

201. Shim, B.S.; Jin, S.; Park, Y.J.; Min, H.S. A new statistical model for SILC distribution of Flash memory and the effect of spatial trap distribution. In Proceedings of the International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Monterey, CA, USA, 6–8 September 2006; pp. 326–329.

202. Serov, A.; Shin, D.; Kim, D.S.; Kim, T.; Lee, K.H.; Park, Y.K.; Yoo, M.H.; Kim, T.; Sung, S.K.; Lee, C.H. Statistical retention modeling in floating-gate cell: ONO scaling. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Montreal, QC, Canada, 26–30 April 2009; pp. 887–890.

203. Amoroso, S.; Gerrer, L.; Adamu-Lema, F.; Markov, S.; Asenov, A. Impact of statistical variability and 3D electrostatics on post-cycling anomalous charge loss in nanoscale Flash memories. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Monterey, CA, USA, 14–18 April 2013; pp. 3B.4.1–3B.4.6.

204. Kubota, T.; Ando, K.; Muramatsu, S. The effect of the floating gate / tunnel SiO_2 interface on FLASH memory data retention reliability. In Proceedings of the IEEE International Reliability Physics Symposium, Dalas, TX, USA, 30 April 30–2 May 1996; pp. 12–16.

205. Park, M.; Suh, K.; Kim, K.; Hur, S.H.; Kim, K.; Lee, W.S. The effect of trapped charge distributions on data retention characteristics of NAND Flash memory cells. *IEEE Electron Device Lett.* **2007**, *28*, 750–752.

206. Park, M.; Ahn, E.; Cho, E.; Kim, K.; Lee, W.S. The effect of negative V_{TH} of NAND Flash memory cells on data retention characteristics. *IEEE Electron Device Lett.* **2009**, *30*, 155–157.

207. Lee, M.Y.; Hsiao, W.H.; Chen, R.P.; Kuo, L.K.; Dai, S.Q.; Ting, R.C.; Chen, C.L.; Chu, D.G.; Lu, C.Y. Anomalous single bit retention induced by asymmetric STI-corner-thinning for floating gate Flash memories. In Proceedings of the 19th IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits, Singapore, 2–6 July 2012; pp. 263–267.

208. Kang, D.; Lee, K.; Seo, S.; Kim, S.; Lee, J.S.; Bae, D.S.; Li, D.H.; Hwang, Y.; Shin, H. Generation dependence of retention characteristics in extremely scaled NAND Flash memory. *IEEE Trans. Electron Devices* **2013**, *34*, 1139–1141.

209. Yang, H.; Kim, H.; Park, S.-I.; Kim, J.; Lee, S.H.; Choi, J.K.; Hwang, D.; Kim, C.; Park, M.; Lee, K.H.; et al. Reliability issues and models of sub-90nm NAND Flash memory cells. In Proceedings of the 8th International Conference on Solid-State and Integrated Circuit Technology Proceedings, Shanghai, China, 23–26 October 2006; pp. 760–762.

210. Dumin, D.J.; Maddux, J.R. Correlation of stress-induced leakage current in thin oxides with trap generation inside the oxides. *IEEE Trans. Electron Devices* **1993**, *40*, 986–993.

211. Scott, R.S.; Dumin, D.J. The transient nature of excess low-level leakage currents in thin oxides. *J. Electrochem. Soc.* **1995**, *142*, 586–590.

212. Scott, R.S.; Dumin, D.J. The superposition of transient low-level leakage currents in stressed silicon oxides. *Solid-State Electron.* **1995**, *38*, 1325–1328.

213. Scott, R.S.; Dumin, N.A.; Hughes, T.W.; Dumin, D.J.; Moore, B.T. Properties of high-voltage stress generated traps in thin silicon oxide. *IEEE Trans. Electron Devices* **1996**, *43*, 1133–1142.

214. Scott, R.S.; Dumin, D.J. The charging and discharging of high-voltage stress-generated traps in thin silicon oxide. *IEEE Trans. Electron Devices* **1996**, *43*, 130–136.

215. Spinelli, A.S.; Lacaita, A.L.; Rigamonti, M.; Ghidini, G. Experimental method for the determination of the energy distribution of stress-induced oxide traps. *IEEE Electron Device Lett.* **1999**, *20*, 106–108.

216. Spinelli, A.S.; Lacaita, A.L.; Rigamonti, M.A.; Ielmini, D.; Ghidini, G. Separation of electron and hole traps by transient current analysis. *Microelectron. Eng.* **1999**, *48*, 151–154.

217. Monzio Compagnoni, C.; Spinelli, A.S.; Bianchini, A.; Lacaita, A.L.; Spiga, S.; Fanciulli, M. Characterization of transient currents in HfO_2 capacitors in the short timescale. *Microelectron. Eng.* **2006**, *83*, 1927–1930.

218. Manzini, S.; Modelli, A. Tunneling discharge of trapped holes in silicon dioxide. In *Insulating Films on Semiconductors*; Elsevier Science: Amsterdam, The Netherlands, 1983; pp. 112–115.

219. Yamada, R.; Mori, Y.; Okuyama, Y.; Yugami, J.; Nishimoto, T.; Kume, H. Analysis of detrap current due to oxide traps to improve Flash memory retention. In Proceedings of the IEEE International Reliability Physics Symposium, San Jose, CA, USA, 10–13 April 2000; pp. 200–204.

220. Yamada, R.; King, T. Variable stress-induced leakage current and analysis of anomalous charge loss for flash memory application. In Proceedings of the IEEE International Reliability Physics Symposium, Dallas, TX, USA, 30 March–4 April 2003; pp. 491–496.

221. Mielke, N.; Belgal, H.; Kalastirsky, I.; Kalavade, P.; Kurtz, A.; Meng, Q.; Righos, N.; Wu, J. Flash EEPROM threshold instabilities due to charge trapping during program/erase cycling. *IEEE Trans. Device Mater. Reliab.* **2004**, *4*, 335–344.

222. Mielke, N.; Belgal, H.; Fazio, A.; Meng, Q.; Righos, N. Recovery effects in the distributed cycling of Flash memories. In Proceedings of the IEEE International Reliability Physics Symposium, San Jose, CA, USA, 26–30 March 2006; pp. 29–35.

223. Paolucci, G.M.; Monzio Compagnoni, C.; Miccoli, C.; Spinelli, A.S.; Lacaita, A.L.; Visconti, A. Revisiting charge trapping/detrapping in Flash memories from a discrete and statistical standpoint—Part I: V_T instabilities. *IEEE Trans. Electron Devices* **2014**, *61*, 2802–2810.

224. Paolucci, G.M.; Monzio Compagnoni, C.; Miccoli, C.; Spinelli, A.S.; Lacaita, A.L.; Visconti, A. Revisiting charge trapping/detrapping in Flash memories from a discrete and statistical standpoint—Part II: On-field operation and distributed-cycling effects. *IEEE Trans. Electron Devices* **2014**, *61*, 2811–2819.

225. Resnati, D.; Monzio Compagnoni, C.; Paolucci, G.M.; Miccoli, C.; Barber, J.; Bertuccio, M.; Beltrami, S.; Lacaita, A.L.; Spinelli, A.S.; Visconti, A. A step ahead toward a new microscopic picture for charge trapping/detrapping in Flash memories. In Proceedings of the IEEE International Reliability Physics Symposium, Pasadena, CA, USA, 17–21 April 2016; pp. 6C-3-1–6C-3-7.

226. Miccoli, C.; Monzio Compagnoni, C.; Beltrami, S.; Spinelli, A.S.; Visconti, A. Threshold-voltage instability due to damage recovery in nanoscale NAND Flash memories. *IEEE Trans. Electron Devices* **2011**, *58*, 2406–2414.

227. Resnati, D.; Nicosia, G.; Paolucci, G.M.; Visconti, A.; Monzio Compagnoni, C. Cycling-induced charge trapping/detrapping in Flash memories—Part I: Experimental evidence. *IEEE Trans. Electron Devices* **2016**, *63*, 4753–4760.

228. Ghidini, G.; Sebastiani, A.; Brazzelli, D. Stress induced leakage current and bulk oxide trapping: Temperature evolution. In Proceedings of the IEEE International Reliability Physics Symposium, Dallas, TX, USA, 7–11 April 2002; pp. 415–416.

229. Lee, K.; Kang, M.; Seo, S.; Kang, D.; Kim, S.; Li, D.H.; Shin, H. Activation energies (E_a) of failure mechanisms in advanced NAND Flash cells for different generations and cycling. *IEEE Trans. Electron Devices* **2013**, *60*, 1099–1107.

230. Paolucci, G.M.; Miccoli, C.; Monzio Compagnoni, C.; Crespi, L.; Spinelli, A.S.; Lacaita, A.L. Investigation of cycling-induced V_T instabilities in NAND Flash cells via compact modeling. In Proceedings of the 2012 4th IEEE International Memory Workshop (IMW), Milan, Italy, 20–23 May 2012; pp. 194–197.

231. Paolucci, G.M.; Miccoli, C.; Monzio Compagnoni, C.; Spinelli, A.S.; Lacaita, A.L. String current in decananometer NAND Flash arrays: A compact-modeling investigation. *IEEE Trans. Electron Devices* **2012**, *59*, 2331–2337.

232. Miccoli, C.; Monzio Compagnoni, C.; Chiavarone, L.; Beltrami, S.; Lacaita, A.L.; Spinelli, A.S.; Visconti, A. Reliability characterization issues for nanoscale Flash memories: A case study on 45-nm NOR devices. *IEEE Trans. Device Mater. Reliab.* **2013**, *13*, 362–369.

233. Resnati, D.; Nicosia, G.; Paolucci, G.M.; Visconti, A.; Monzio Compagnoni, C. Cycling-induced charge trapping/detrapping in Flash memories—Part II: Modeling. *IEEE Trans. Electron Devices* **2016**, *63*, 4761–4768.

234. Paolucci, G.M; Monzio Compagnoni, C.; Miccoli, C.; Bertuccio, M.; Beltrami, S.; Barber, J.; Kessenich, J.; Lacaita, A.L.; Spinelli, A.S.; Visconti, A. A new spectral method to modeling charge trapping/detrapping in NAND Flash memories. In Proceedings of the IEEE International Reliability Physics Symposium, Waikoloa, HI, USA, 1–5 June 2014; pp. 2E-2-1–2E-2-6.

235. JEDEC Standard JESD22-A117C. *Electrically Erasable Programmable ROM (EEPROM) Program/Erase Endurance and Data Retention Stress Test*; Technical Report; JEDEC Solid State Technology Association: Arlington, VA, USA, 2011.

236. JEDEC Standard JESD22-A117A. *Electrically Erasable Programmable ROM (EEPROM) Program/Erase Endurance and Data Retention Stress Test*; Technical Report; JEDEC Solid State Technology Association: Arlington, VA, USA, 2006.

237. Lee, W.H.; Hur, C.H.; Lee, H.M.; Yoo, H.; Lee, S.E.; Lee, B.Y.; Park, C.; Kim, K. Post-cycling data retention failure in multilevel NOR Flash memory with nitrided tunnel-oxide. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Montreal, QC, Canada, 26–30 April 2009; pp. 907–908.

238. Lakshmanna, V.; Vengurlekar, A.S. Logarithmic detrapping response for holes injected into SiO_2 and the influence of thermal activation and electric fields. *J. Appl. Phys.* **1988**, *63*, 4548–4554.

239. Paolucci, G.M.; Monzio Compagnoni, C.; Spinelli, A.S.; Lacaita, A.L.; Goda, A. Fitting cells into a narrow V_T interval: Physical constraints along the lifetime of an extremely scaled NAND Flash memory array. *IEEE Trans. Electron Devices* **2015**, *62*, 1491–1497.

240. Monzio Compagnoni, C.; Chiavarone, L.; Calabrese, M.; Gusmeroli, R.; Ghidotti, M.; Lacaita, A.L.; Spinelli, A.S.; Visconti, A. Granular electron injection and random telegraph noise impact on the programming accuracy of NOR Flash memories. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Montreal, QC, Canada, 26–30 April 2009; pp. 274–279.

241. Aritome, S.; Kikkawa, T. Scaling challenge of self-aligned STI cell (SA-STI cell) for NAND Flash memories. *Solid-State Electron.* **2013**, *82*, 54–62.

242. Lee, K.; Kang, D.; Shin, H.; Kwon, S.; Kim, S.; Hwang, Y. Analysis of failure mechanisms in erased state of sub 20-nm NAND Flash memory. In Proceedings of the 2014 44th European Solid State Device Research Conference (ESSDERC), Venezia, Italy, 22–26 September 2014; pp. 58–61.

243. Wu, D.Y.; Chen, S.F.; Lin, C.J.; King, Y.C. Dummy read scheme for lifetime improvement of MLC NAND Flash memories. *IEEE Trans. Device Mater. Reliab.* **2016**, *16*, 583–587.

244. Lee, C.; Lee, S.K.; Ahn, S.; Lee, J.; Park, W.; Cho, Y.; Jang, C.; Yang, C.; Chung, S.; Yun, I.S.; et al. A 32-Gb MLC NAND Flash memory with Vth endurance enhancing schemes in 32 nm CMOS. *IEEE J. Solid-State Circuits* **2011**, *46*, 97–106.

245. Lue, H.T.; Du, P.Y.; Chen, C.P.; Chen, W.C.; Hsieh, C.C.; Hsiao, Y.H.; Shih, Y.H.; Lu, C.Y. Radically extending the cycling endurance of Flash memory (to > 100M cycles) by using built-in thermal annealing to self-heal the stress-induced damage. In Proceedings of the 2012 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 10–13 December 2012; pp. 199–202.

246. Chang, Y.M.; Chang, Y.H.; Chen, J.J.; Kuo, T.W.; Li, H.P.; Lue, H.T. On trading wear-leveling with heal-leveling. In Proceedings of the 2014 51st ACM/EDAC/IEEE Design Automation Conference (DAC), San Francisco, CA, USA, 1–5 June 2014; pp. 478–483.

247. Chen, R.; Wang, Y.; Liu, D.; Shao, Z.; Jiang, S. Heating dispersal for self-healing NAND Flash memory. *IEEE Trans. Comput.* **2017**, *66*, 361–367.

248. Lee, J.D.; Choi, J.H.; Park, D.; Kim, K. Degradation of tunnel oxide by FN current stress and its effects on data retention characteristics of 90 nm NAND Flash memory cells. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Dallas, TX, USA, 30 March–4 April 2003; pp. 497–501.

249. Lee, K.; Kang, M.; Seo, S.; Li, D.H.; Kim, J.; Shin, H. Analysis of failure mechanisms and extraction of activation energies (E_a) in 21-nm NAND Flash cells. *IEEE Electron Device Lett.* **2013**, *34*, 48–50.

250. Yan, C.R.; Chen, J.F.; Lee, Y.J.; Liao, Y.J.; Lin, C.Y.; Chen, C.Y.; Lin, Y.C.; Chen, H.H. Extraction and analysis of interface states in 50-nm NAND Flash devices. *IEEE Trans. Electron Devices* **2013**, *60*, 992–997.

251. Dunn, C.; Kaya, C.; Lewis, T.; Strauss, T.; Schreck, J.; Hefley, P.; Middennorf, M.; San, T. Flash EEPROM disturb mechanism. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), San Jose, CA, USA, 11–14 April 1994; pp. 299–308.

252. Kato, M.; Miyamoto, N.; Kume, H.; Satoh, A.; Adachi, T.; Ushiyama, M.; Kimura, K. Read-disturb degradation mechanism due to electron trapping in the tunnel oxide for low-voltage flash memories. In Proceedings of the 1994 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 11–14 December 1994; pp. 45–48.

253. Endoh, T.; Shimizu, K.; Iizuka, H.; Masuoka, F. A new write/erase method to improve the read disturb characteristics based on the decay phenomena of stress leakage current for Flash memories. *IEEE Trans. Electron Devices* **1998**, *45*, 98–104.

254. Wada, T. Acceleration method for gate-disturb degradation on embedded flash EEPROM. *Microelectron. Reliab.* **2000**, *40*, 1279–1283.

255. Cai, Y.; Luo, Y.; Ghose, S.; Mutlu, O. Read disturb errors in MLC NAND Flash memory: characterization, mitigation, and recovery. In Proceedings of the 28th International Conference on Defects in Semiconductors, Matsue City, Japan, 27–31 July 2015; pp. 438–449.

256. Kang, D.; Lee, K.; Kwon, S.; Kim, S.; Hwang, Y.; Shin, H. Analysis of read disturbance mechanism in retention of sub-20 nm NAND flash memory. *Jpn. J. Appl. Phys.* **2015**, *54*, 04DD03-1–04DD03-4.

257. Grupp, L.M.; Caulfield, A.M.; Coburn, J.; Swanson, S.; Yaakobi, E.; Wolf, P.H.S.J.K. Characterizing Flash memory: Anomalies, observations, and applications. In Proceedings of the 2009 42nd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO), New York, NY, USA, 12–16 December 2009; pp. 24–33.

258. Goda, A.; Moriyama, W.; Hazama, H.; Iizuka, H.; Shimizu, K.; Aritome, S.; Shirota, R. A novel surface-oxidized barrier-SiN cell technology to improve endurance and read-disturb characteristics for gigabit NAND Flash memories. In Proceedings of the 2000 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 10–13 December 2000; pp. 771–774.

259. Lee, Y.H.; McMahon, W.; Lu, Y.L.R.; Tewg, J.Y.J.; Ma, S.T. On the scaling of Flash cell spacer for gate disturb and charge retention otimization. *IEEE Trans. Electron Devices* **2009**, *56*, 1959–1965.

260. Joe, S.M.; Jeong, M.K.; Kang, M.; Han, K.R.; Park, S.K.; Lee, J.H. New read schemes using boosted channel potential of adjacent bit-line Strings in NAND Flash memory. *IEEE Electron Device Lett.* **2012**, *33*, 1198–1200.

261. Deguchi, Y.; Tokutomi, T.; Takeuchi, K. System-level error correction by read-disturb error model of 1Xnm TLC NAND Flash memory for read-intensive enterprise solid-state drives (SSDs). In Proceedings of the IEEE International Reliability Physics Symposium, Pasadena, CA, USA, 17–21 April 2016; pp. MY-6-1–MY-6-4.

262. Papandreou, N.; Parnell, T.; Mittelholzer, T.; Pozidis, H.; Griffin, T.; Tressler, G.; Fisher, T.; Camp, C. Effect of read disturb on incomplete blocks in MLC NAND Flash arrays. In Proceedings of the IEEE International Memory Workshop (IMW), Paris, France, 15–18 May 2016; pp. 117–120.

263. Tanaka, T.; Momodomi, M.; Iwata, Y.; Tanaka, Y.; Oodaira, H.; Itoh, Y.; Shirota, R.; Ohuchi, K.; Masuoka, F. A 4-Mbit NAND-EEPROM with tight programmed V_T distribution. In Proceedings of the 1990 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers, Honolulu, HI, USA, 4–7 June 1990; pp. 105–106.

264. Tanaka, T.; Tanaka, Y.; Nakamura, H.; Oodaira, H.; Aritome, S.; Shirota, R.; Masuoka, F. A quick intelligent program architecture for 3 V-only NAND-EEPROMs. In Proceedings of the 1992 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers, Seattle, WA, USA, 2–4 June 1992; pp. 20–21.

265. Suh, K.D.; Suh, B.H.; Lim, Y.H.; Kim, J.K.; Choi, Y.J.; Koh, Y.N.; Lee, S.S.; Kwon, S.C.; Choi, B.S.; Yum, J.S.; et al. A 3.3 V 32 Mb NAND Flash memory with incremental step pulse programming scheme. *IEEE J. Solid-State Circuits* **1995**, *30*, 1149–1156.

266. Lee, J.D.; Lee, C.K.; Lee, M.W.; Kim, H.S.; Park, K.C.; Lee, W.S. A new programming disturbance phenomenon in NAND Flash memory by source/drain hot-electrons generated by GIDL current. In Proceedings of the 21st IEEE Non-Volatile Semiconductor Memory Workshop, Monterey, CA, USA, 12–16 February 2006; pp. 31–33.

267. Oh, D.; Lee, C.; Lee, S.; Kim, T.K.; Song, J.; Choi, J. A new self-boosting phenomenon by source/drain depletion cut-off in NAND Flash memory. In Proceedings of the 22nd IEEE Non-Volatile Semiconductor Memory Workshop, Monterey, CA, USA, 26–30 August 2007; pp. 39–41.

268. Cho, T.; Lee, Y.T.; Kim, E.; Lee, J.; Choi, S.; Lee, S.; Kim, D.H.; Han, W.K.; Lim, Y.H.; Lee, J.D.; et al. A 3.3 V 1 Gb multi-level NAND Flash memory with non-uniform threshold voltage distribution. In Proceedings of the 1995 IEEE International Solid-State Circuits Conference, San Francisco, CA, USA, 5–7 February 2001; pp. 1–2.

269. Park, K.T.; Song, Y.; Kang, M.; Lee, S.; Lim, Y.H.; Suh, K.D.; Chung, C. Dynamic Vpass controlled program scheme and optimized erase Vth control for high program inhibition in MLC NAND Flash memories. *IEEE J. Solid-State Circuits* **2010**, *45*, 2165–2172.

270. Shirota, R.; Huang, C.H.; Nagai, S.; Sakamoto, Y.; Li, F.H.; Mitiukhina, N.; Arakawa, H. A new programming scheme for the improvement of program disturb characteristics in scaled NAND Flash memory. *IEEE Trans. Electron Devices* **2012**, *59*, 2767–2773.

271. Kang, M.; Park, K.T.; Song, Y.; Hwang, S.; Choi, B.Y.; Song, Y.; Lee, Y.T.; Kim, C. Improving read disturb characteristics by self-boosting read scheme for multilevel NAND Flash memories. *Jpn. J. Appl. Phys.* **2009**, *48*, 04C062-1–04C062-6.

272. Park, K.T.; Lee, S.; Sel, J.S.; Choi, J.; Kim, K. Scalable wordline shielding using dummy cell beyond 40 nm NAND Flash memory for eliminating abnormal disturb of edge memory cell. *Jpn. J. Appl. Phys.* **2007**, *46*, 2188–2192.

273. Kanda, K.; Koyanagi, M.; Yamamura, T.; Hosono, K.; Yoshihara, M.; Miwa, T.; Kato, Y.; Mak, A.; Chan, S.L.; Tsai, F.; et al. A 120 mm^2 16 Gb 4-MLC NAND Flash memory with 43 nm CMOS technology. In Proceedings of the 2008 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 3–7 February 2008; pp. 430–431.

274. Park, I.H.; Hahn, W.G.; Song, K.W.; Choi, K.H.; Choi, H.K.; Lee, S.B.; Lee, C.S.; Song, J.H.; Han, J.M.; Kyoung, K.H.; et al. A new GIDL phenomenon by field effect of neighboring cell transistors and its control solutions in sub-30 nm NAND Flash devices. In Proceedings of the 2012 Symposium on VLSI Technology (VLSI-Technology), Honolulu, HI, USA, 12–14 June 2012; pp. 23–24.

275. Park, M.; Park, S.; Cho, S.; Lee, D.K.; Jeong, Y.; Hong, C.; Lee, H.S.; Cho, M.K.; Ahn, K.O.; Koh, Y. NAND Flash reliability degradation induced by HCI in boosted channel potential. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Garden Grove, CA, USA, 2–6 May 2010; pp. 975–976.

276. Lee, C.; Hwang, J.; Fayrushin, A.; Kim, H.; Son, B.; Park, Y.; Jin, G.; Jung, E. Channel coupling phenomenon as scaling barrier of NAND Flash memory beyond 20nm node. In Proceedings of the 2013 5th IEEE International Memory Workshop (IMW), Monterey, CA, USA, 26–29 May 2013; pp. 72–75.

277. Chandrasekaran, S.; Venkatesan, S.; Eagle, O.H.; Iyengar, V.V.; Reyes, A.B.; Gowda, S.S. The effect of shallow trench isolation improvement on program disturb response in 20 nm NAND Flash technology. In Proceedings of the 2014 IEEE Workshop On Microelectronics And Electron Devices (WMED), Boise, ID, USA, 18 April 2014; pp. 1–4.

278. Seo, S.; Kim, H.; Park, S.; Lee, S.; Aritome, S.; Hong, S. Novel negative Vt shift program disturb phenomena in 2X 3X NAND Flash memory cells. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Monterey, CA, USA, 10–14 April 2011; pp. 641–644.

279. Aritome, S.; Seo, S.; Kim, H.S.; Park, S.K.; Lee, S.K.; Hong, S. Novel negative V_T shift phenomenon of program-inhibit cell in $2X - 3X$-nm self-aligned STI NAND Flash memory. *IEEE Trans. Electron Devices* **2012**, *59*, 2950–2955.

280. Zambelli, C.; Andrian, F.; Aritome, S.; Olivo, P. Compact modeling of negative V_T shift disturb in NAND Flash memories. *IEEE Trans. Electron Devices* **2016**, *63*, 1516–1523.

281. Torsi, A.; Zhao, Y.; Liu, H.; Tanzawa, T.; Goda, A.; Kalavade, P.; Parat, K. A program disturb model and channel leakage current study for sub-20 nm NAND Flash cells. *IEEE Trans. Electron Devices* **2011**, *58*, 11–16.

282. Oh, D.; Lee, S.; Lee, C.; Song, J.; Lee, W.; Choi, J. Program disturb phenomenon by DIBL in MLC NAND Flash device. In Proceedings of the 2008 Joint Non-Volatile Semiconductor Memory Workshop (NVSMW) & International Conference on Memory Technology and Design (ICMTD), Opio, France, 18–21 May 2008; pp. 5–7.

283. Kang, M.; Hahn, W.; Park, I.H.; Park, J.; Song, Y.; Lee, H.; Eun, C.; Ju, S.; Choi, K.; Lim, Y.; et al. DIBL-induced program disturb characteristics in 32-nm NAND Flash memory array. *IEEE Trans. Electron Devices* **2011**, *58*, 3626–3629.

284. Tsai, W.J.; Huang, J.S.; Tsai, P.H.; Yan, S.G.; Cheng, C.H.; Cheng, C.C.; Chen, Y.J.; Lee, C.H.; Han, T.T.; Lu, T.C.; et al. A novel low-voltage low-power programming method for NAND Flash cell by utilizing self-boosting channel potential for carrier heating. *IEEE Trans. Electron Devices* **2011**, *58*, 1620–1627.

285. Ahn, S.T.; Mun, K.; Lee, K.W.; Cho, G.; Park, S.K.; Lee, S.; Hong, S. Advanced hot-carrier injection programming scheme for sub 20nm NAND Flash cell and beyond. In Proceedings of the 2012 4th IEEE International Memory Workshop (IMW), Milan, Italy, 20–23 May 2012; pp. 29–32.

286. Fukuzumi, Y.; Katsumata, R.; Kito, M.; Kido, M.; Sato, M.; Tanaka, H.; Nagata, Y.; Matsuoka, Y.; Iwata, Y.; Aochi, H.; et al. Optimal integration and characteristics of vertical array devices for ultra-high density, bit-cost scalable Flash memory. In Proceedings of the 2007 International Electron Devices Meeting (IEDM), Washington, DC, USA, 10–12 December 2007; pp. 449–452.

287. Whang, S.; Lee, K.; Shin, D.; Kim, B.; Kim, M.; Bin, J.; Han, J.; Kim, S.; Lee, B.; Jung, Y.; et al. Novel 3-Dimensional dual control-gate with surrounding floating-gate (DC-SF) NAND Flash cell for 1Tb file storage application. In Proceedings of the 2010 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 6–8 December 2010; pp. 668–671.

288. Noh, Y.; Ahn, Y.; Yoo, H.; Han, B.; Chung, S.; Shim, K.; Lee, K.; Kwak, S.; Shin, S.; Choi, I.; et al. A new metal control gate last process (MCGL process) for high performance DC-SF (dual control gate with surrounding floating gate) 3D NAND Flash memory. In Proceedings of the 2012 Symposium on VLSI Technology (VLSI-Technology), Honolulu, HI, USA, 12–14 June 2012; pp. 19–20.

289. Aritome, S.; Whang, S.; Lee, K.; Shin, D.; Kim, B.; Kim, M.; Bin, J.; Han, J.; Kim, S.; Lee, B.; et al. A novel three-dimensional dual control-gate with surrounding floating-gate (DC-SF) NAND flash cell. *Solid-State Electron.* **2013**, *79*, 166–171.

290. Aritome, S.; Noh, Y.; Yoo, H.; Choi, E.S.; Joo, H.S.; Ahn, Y.; Han, B.; Chung, S.; Shim, K.; Lee, K.; et al. Advanced DC-SF cell technology for 3-D NAND Flash. *IEEE Trans. Electron Devices* **2013**, *60*, 1327–1333.

291. Parat, K.; Dennison, C. A floating gate based 3-D NAND technology with CMOS under array. In Proceedings of the 2015 International Electron Devices Meeting (IEDM), Washington, DC, USA, 7–9 December 2015; pp. 48–51.

292. Kang, D.; Jeong, W.; Kim, C.; Kim, D.H.; Cho, Y.S.; Kang, K.T.; Ryu, J.; Kang, K.M.; Lee, S.; Kim, W.; et al. 256 Gb 3 b/cell V-NAND Flash memory with 48 stacked WL layers. In Proceedings of the 2014 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 31 January–4 February 2016; pp. 130–131.

293. Choi, J.; Seol, K.S. 3D approaches for non-volatile memory. In Proceedings of the 2011 Symposium on VLSI Technology (VLSI-Technology), Kyoto, Japan, 14–16 June 2011; pp. 178–179.

294. Choi, E.S.; Park, S.K. Device considerations for high density and highly reliable 3D NAND Flash cell in near future. In Proceedings of the 2012 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 10–13 December 2012; pp. 211–214.

295. Im, J.W.; Jeong, W.P.; Kim, D.H.; Nam, S.W.; Shim, D.K.; Choi, M.H.; Yoon, H.J.; Kim, D.H.; Kim, Y.S.; Park, H.W.; et al. A 128Gb 3b/cell V-NAND Flash memory with 1Gb/s I/O rate. In Proceedings of the 2015 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 22–26 February 2015; pp. 130–132.

296. Jeong, W.; Im, J.W.; Kim, D.H.; Nam, S.W.; Shim, D.K.; Choi, M.H.; Yoon, H.J.; Kim, D.H.; Kim, Y.S.; Park, H.W.; et al. A 128 Gb 3 b/cell V-NAND Flash memory with 1 Gb/s I/O rate. *IEEE J. Solid-State Circuits* **2016**, *51*, 204–212.

297. Janai, M.; Eitan, B. The kinetics of degradation of data retention of post-cycled NROM non-volatile memory products. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), San Jose, CA, USA, 17–21 April 2003; pp. 175–180.

298. Lee, M.C.; Wong, H.Y. Investigation on the origin of the anomalous tail bits on nitrided charge trap flash memory. *Microelectron. Reliab.* **2015**, *55*, 337–341.

299. Goda, A.; Miccoli, C.; Monzio Compagnoni, C. Time dependent threshold-voltage fluctuations in NAND Flash memories: From basic physics to impact on array operation. In Proceedings of the 2015 International Electron Devices Meeting (IEDM), Washington, DC, USA, 7–9 December 2015; pp. 374–377.

300. Monzio Compagnoni, C.; Spinelli, A.S.; Lacaita, A.L. Experimental study of data retention in nitride memories by temperature and field acceleration. *IEEE Electron Device Lett.* **2007**, *28*, 628–630.

301. Shappir, A.; Levy, D.; Shacham-Diamand, Y.; Lusky, E.; Bloom, I.; Eitan, B. Spatial characterization of localized charge trapping and charge redistribution in the NROM device. *Solid-State Electron.* **2004**, *48*, 1489–1495.

302. Kang, C.; Choi, J.; Sim, J.; Lee, C.; Shin, Y.; Park, J.; Sel, J.; Jeon, S.; Park, Y.; Kim, K. Effects of lateral charge spreading on the reliability of TANOS (TaN/AlO/SiN/Oxide/Si) NAND Flash memory. In Proceedings of the IEEE International Reliability Physics Symposium, Phoenix, AZ, USA, 15–19 April 2007; pp. 167–170.

303. Maconi, A.; Arreghini, A.; Monzio Compagnoni, C.; Van den bosch, G.; Spinelli, A.S.; Van Houdt, J.; Lacaita, A.L. Comprehensive investigation of the impact of lateral charge migration on retention performance of planar and 3D SONOS devices. *Solid-State Electron.* **2012**, *74*, 64–70.

304. Liu, Y.H.; Jiang, C.M.; Chen, W.C.; Wang, T.; Tsai, W.J.; Lu, T.C.; Chen, K.C.; Lu, C.Y. Electric field induced nitride trapped charge lateral migration in a SONOS Flash memory. *IEEE Electron Device Lett.* **2017**, *38*, 48–51.

305. Kang, H.J.; Choi, N.; Joe, S.M.; Seo, J.H.; Choi, E.; Park, S.K.; Park, B.G.; Lee, J.H. Comprehensive analysis of retention characteristics in 3-D NAND Flash memory cells with tube-type poly-Si channel structure. In Proceedings of the 2015 Symposium on VLSI Technology (VLSI-Technology), Kyoto, Japan, 15–18 June 2015; pp. 182–183.

306. Lun, Z.; Liu, S.; He, Y.; Hou, Y.; Zhao, K.; Du, G.; Liu, X.; Wang, Y. Investigation of retention behavior for 3D charge trapping NAND Flash memory by 2D self-consistent simulation. In Proceedings of the 2014 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Yokohama, Japan, 9–11 September 2014; pp. 141–144.

307. Li, X.; Huo, Z.; Jin, L.; Wang, Y.; Liu, J.; Jiang, D.; Yang, X.; Liu, M. Investigation of charge loss mechanisms in 3D TANOS cylindrical junction-less charge trapping memory. In Proceedings of the IEEE 12th International Conference on Solid -State and Integrated Circuit Technology (ICSICT), Guilin, China, 28–31 October 2014; pp. 720–722.

308. Oh, D.; Lee, B.; Kwon, E.; Kim, S.; Cho, G.; Park, S.; Lee, S.; Hong, S. TCAD simulation of data retention characteristics of charge trap device for 3-D NAND Flash memory. In Proceedings of the IEEE International Memory Workshop (IMW), Monterey, CA, USA, 17–20 May 2015; pp. 113–116.

309. Kim, S.H.; Yang, S.D.; Kim, J.S.; Jeong, J.K.; Lee, H.D.; Lee, G.W. Charge spreading effect of stored charge on retention characteristics in SONOS NAND Flash memory devices. *Trans. Electr. Electron. Mater.* **2015**, *16*, 183–186.

310. Park, S.; Choi, S.; Jun, K.S.; Kim, H.; Rhee, S.; Park, Y.J. Investigation on multiple activation energy of retention in charge trapping memory using self-consistent simulation. In Proceedings of the 2014 44th European Solid State Device Research Conference (ESSDERC), Venezia, Italy, 22–26 September 2014; pp. 50–53.

311. Jeong, M.K.; Kang, H.J.; Joe, S.M.; Park, S.K.; Park, B.G.; Lee, J.H. Trap profiling in nitride storage layer in 3-D NAND Flash memory using retention characteristics and AC-g_m method. *IEEE Electron Device Lett.* **2015**, *36*, 561–563.

312. Janai, M.; Bloom, I. Charge gain, NBTI, and random telegraph noise in EEPROM Flash memory devices. *IEEE Electron Device Lett.* **2010**, *31*, 1038–1040.

313. Janai, M.; Bloom, I.; Shur, Y. Charge gain, NBTI recovery and random telegraph noise in localized-trapping NVM devices. In Proceedings of the IEEE International Reliability Physics Symposium (IRPS), Monterey, CA, USA, 10–14 April 2011; pp. 819–823.

314. Chen, C.P.; Lue, H.T.; Hsieh, C.C.; Chang, K.P.; Hsieh, K.Y.; Lu, C.Y. Study of fast initial charge loss and its impact on the programmed states Vt distribution of charge-trapping NAND Flash. In Proceedings of the 2010 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 6–8 December 2010; pp. 118–121.

315. Choi, B.; Jang, S.H.; Yoon, J.; Lee, J.; Jeon, M.; Lee, Y.; Han, J.; Lee, J.; Kim, D.M.; Kim, D.H.; et al. Comprehensive evaluation of early retention (fast charge loss within a few seconds) characteristics in tube-type 3-D NAND Flash memory. In Proceedings of the 2016 Symposium on VLSI Technology (VLSI-Technology), Honolulu, HI, USA, 14–16 June 2016; pp. 61–62.

316. Subirats, A.; Arreghini, A.; Degraeve, R.; Linten, D.; den bosch, G.V.; Furnemont, A. In depth analysis of post-program V_T instability after electrical stress in 3D SONOS memories. In Proceedings of the IEEE International Memory Workshop (IMW), Paris, France, 15–18 May 2016; pp. 84–87.

317. Jeong, M.K.; Joe, S.M.; Seo, C.S.; Han, K.R.; Choi, E.; Park, S.K.; Lee, J.H. Analysis of random telegraph noise and low frequency noise properties in 3-D stacked NAND Flash memory with tube-type poly-Si channel structure. In Proceedings of the 2012 Symposium on VLSI Technology (VLSI-Technology), Honolulu, HI, USA, 12–14 June 2012; pp. 55–56.

318. Jeong, M.K.; Joe, S.M.; Jo, B.S.; Kang, H.J.; Bae, J.H.; Han, K.R.; Choi, E.; Cho, G.; Park, S.K.; Park, B.G.; et al. Characterization of traps in 3-D stacked NAND Flash memory devices with tube-type poly-Si channel structure. In Proceedings of the 2012 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 10–13 December 2012; pp. 207–210.

319. Kang, D.; Lee, C.; Hur, S.; Song, D.; Choi, J.H. A new approach for trap analysis of vertical NAND Flash cell using RTN characteristics. In Proceedings of the 2014 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 15–17 December 2014; pp. 367–370.

320. Nowak, E.; Kim, J.H.; Kwon, H.; Kim, Y.G.; Sim, J.S.; Lim, S.H.; Kim, D.S.; Lee, K.H.; Park, Y.K.; Choi, J.H.; et al. Intrinsic fluctuations in vertical NAND Flash memories. In Proceedings of the 2012 Symposium on VLSI Technology (VLSI-Technology), Honolulu, HI, USA, 12–14 June 2012; pp. 21–22.

321. Chou, Y.L.; Wang, T.; Lin, M.; Chang, Y.W.; Liu, L.; Huang, S.W.; Tsai, W.J.; Lu, T.C.; Chen, K.C.; Lu, C.Y. Poly-silicon trap position and pass voltage effects on RTN amplitude in a vertical NAND Flash cell string. *IEEE Electron Device Lett.* **2016**, *37*, 998–1001.

322. Hsieh, C.C.; Lue, H.T.; Hsu, T.H.; Du, P.Y.; Chiang, K.H.; Lu, C.Y. A Monte Carlo simulation method to predict large-density NAND product memory window from small-array test element group (TEG) verified on a 3D NAND Flash test chip. In Proceedings of the 2016 Symposium on VLSI Technology (VLSI-Technology), Honolulu, HI, USA, 14–16 June 2016; pp. 63–64.

323. Cappelletti, P. Non volatile memory evolution and revolution. In Proceedings of the 2015 International Electron Devices Meeting (IEDM), Washington, DC, USA, 7–9 December 2015; pp. 241–244.

324. Resnati, D.; Goda, A.; Spinelli, A.S.; Monzio Compagnoni, C. Thermal effects in NAND Flash memories: A comparison between 2-D and 3-D arrays. *IEEE Electron Device Lett.* **2017**, submitted.

325. Paolucci, G.M.; Spinelli, A.S.; Monzio Compagnoni, C.; Tessariol, P. A semi-analytical model for macaroni MOSFETs with application to vertical Flash memories. *IEEE Trans. Electron Devices* **2016**, *63*, 1871–1876.

326. Toledano-Luque, M.; Degraeve, R.; Roussel, P.J.; Luong, V.; Tang, B.; Lisoni, J.G.; Tan, C.L.; Arreghini, A.; Van den bosch, G.; Groeseneken, G.; et al. Statistical spectroscopy of switching traps in deeply scaled vertical poly-Si channel for 3D memories. In Proceedings of the 2013 International Electron Devices Meeting (IEDM), Washington, DC, USA, 9–11 December 2013; pp. 562–565.

327. Lee, C.M.; Tsui, B.Y. Random telegraph signal noise arising from grain boundary traps in nano-scale poly-Si nanowire thin-film transistors. In Proceedings of the 2013 Symposium on VLSI Technology (VLSI-Technology), Kyoto, Japan, 11–13 June 2013; pp. 47–48.

328. Seto, J.Y.W. The electrical properties of polycrystalline silicon films. *J. Appl. Phys.* **1975**, *46*, 5247–5254.

329. Hashimoto, T.; Aoki, M.; Yamanaka, T.; Kamigaki, Y.; Nishida, T. $1/f$ noise in poly-Si MOSFET and strong correlation with oxide traps and grain boundaries. In Proceedings of the 1994 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers, Honolulu, HI, USA, 7–9 June 1994; pp. 87–88.

330. Brederlow, R.; Weber, W.; Dahl, C.; Schmitt-Landsiedel, D.; Thewes, R. A physically based model for low-frequency noise of poly-silicon resistors. In Proceedings of the 1988 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 11–14 December 1988; pp. 89–92.

331. Kang, H.J.; Jeong, M.K.; Joe, S.M.; Seo, J.H.; Park, S.K.; Jin, S.H.; Park, B.G.; Lee, J.H. Effect of traps on transient bit-line current behavior in word-line stacked NAND Flash memory with poly-Si body. In Proceedings of the 2014 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers, Honolulu, HI, USA, 9–12 June 2014; pp. 24–25.

332. Tsai, W.J.; Lin, W.L.; Cheng, C.C.; Ku, S.H.; Chou, Y.L.; Liu, L.; Hwang, S.W.; Lu, T.C.; Chen, K.C.; Wang, T.; et al. Polycrystalline-silicon channel trap induced transient read instability in a 3D NAND Flash cell string. In Proceedings of the 2016 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 3–7 December 2016; pp. 11.3.1–11.3.4.

333. Degraeve, R.; Toledano-Luque, M.; Suhane, A.; den bosch, G.V.; Arreghini, A.; Tang, B.; Kaczer, B.; Roussel, P.; Kar, G.S.; Houdt, J.V.; Groeseneken, G. Statistical characterization of current paths in narrow poly-Si channels. In Proceedings of the 2011 International Electron Devices Meeting (IEDM), Washington, DC, USA, 5–7 December 2011; pp. 287–290.

334. Toledano-Luque, M.; Degraeve, R.; Kaczer, B.; Tang, B.; Roussel, P.J.; Weckx, P.; Franco, J.; Arreghini, A.; Suhane, A.; Kar, G.S.; et al. Quantitative and predictive model of reading current variability in deeply scaled vertical poly-Si channel for 3D memories. In Proceedings of the 2012 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 10–13 December 2012; pp. 203–206.

335. Kim, S.Y.; Park, J.K.; Hwang, W.S.; Lee, S.J.; Lee, K.H.; Pyi, S.H.; Cho, B.J. Dependence of grain size on the performance of a polysilicon channel TFT for 3D NAND Flash memory. *J. Nanosci. Nanotechnol.* **2016**, *16*, 5044–5048.

336. Brown, A.R.; Roy, G.; Asenov, A. Poly-Si-gate-related variability in decananometer MOSFETs with conventional architecture. *IEEE Trans. Electron Devices* **2007**, *54*, 3056–3063.

337. Baccarani, G.; Riccó, B.; Spadini, G. Transport properties of polycrystalline silicon films. *J. Appl. Phys.* **1978**, *49*, 5565–5570.

338. Lu, N.C.C.; Gerzberg, L.; Lu, C.Y.; Meindl, J.D. Modeling and optimization of monolithic polycrystalline silicon resistors. *IEEE Trans. Electron Devices* **1981**, *28*, 818–830.

339. Farmakis, F.V.; Brini, J.; Kamarinos, G.; Angelis, C.T.; Dimitriadis, C.A.; Miyasaka, M.; Ouisse, T. Grain and grain-boundary control of the transfer characteristics of large-grain polycrystalline silicon thin-film transistors. *Solid-State Electron.* **2000**, *44*, 913–916.

340. Hsiao, Y.H.; Lue, H.T.; Chen, W.C.; Chen, C.P.; Chang, K.P.; Shih, Y.H.; Tsui, B.Y.; Lu, C.Y. Modeling the variability caused by random grain boundary and trap-location induced asymmetrical read behavior for a tight-pitch vertical gate 3D NAND Flash memory using double-gate thin-film transistor (TFT) device. In Proceedings of the 2012 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 10–13 December 2012; pp. 609–612.

341. Hsiao, Y.H.; Lue, H.T.; Chen, W.C.; Chang, K.P.; Shih, Y.H.; Tsui, B.Y.; Hsieh, K.Y.; Lu, C.Y. Modeling the impact of random grain boundary traps on the electrical behavior of vertical gate 3-D NAND Flash memory devices. *IEEE Trans. Electron Devices* **2014**, *61*, 2064–2070.

342. Yang, C.W.; Su, P. Simulation and investigation of random grain-boundary-induced variabilities for stackable NAND Flash using 3-D Voronoi grain patterns. *IEEE Trans. Electron Devices* **2014**, *61*, 1211–1214.

343. Kim, J.; Lee, J.; Oh, H.; Rim, T.; Baek, C.K.; Meyyappan, M.; Lee, J.S. The variability due to random discrete dopant and grain boundary in 3D NAND unit cell. In Proceedings of the 12th International Naval Engineering Conference and Exhibition, Amsterdam, The Netherlands, 20–22 May 2014; pp. 66–68.

344. Wang, P.Y.; Tsui, B.Y. A novel approach using discrete grain-boundary traps to study the variability of 3-D vertical-gate NAND Flash memory cells. *IEEE Trans. Electron Devices* **2015**, *62*, 2488–2493.

345. Degraeve, R.; Clima, S.; Putcha, V.; Kaczer, B.; Roussel, P.; Linten, D.; Groeseneken, G.; Arreghini, A.; Karner, M.; Kernstock, C.; et al. Statistical poly-Si grain boundary model with discrete charging defects and its 2D and 3D implementation for vertical 3D NAND channels. In Proceedings of the 2015 International Electron Devices Meeting (IEDM), Washington, DC, USA, 7–9 December 2015; pp. 121–124.

346. Lun, Z.; Shen, L.; Cong, Y.; Du, G.; Liu, X.; Wang, Y. Investigation of the impact of grain boundary on threshold voltage of 3-D MLC NAND Flash memory. In Proceedings of the 2015 Silicon Nanoelectronics Workshop (SNW), Kyoto, Japan, 14–15 June 2015; pp. 35–37.

347. Oh, H.; Kim, J.; Lee, J.; Rim, T.; Baek, C.K.; Lee, J.S. Effects of single grain boundary and random interface traps on electrical variations of sub-30 nm polysilicon nanowire structures. *Microelectron. Eng.* **2016**, *149*, 113–116.

348. Degraeve, R.; Toledano-Luque, M.; Arreghini, A.; Tang, B.; Capogreco, E.; Lisoni, J.; Roussel, P.; Kaczer, B.; Van den bosch, G.; Groeseneken, G.; et al. Characterizing grain size and defect energy distribution in vertical SONOS poly-Si channels by means of a resistive network model. In Proceedings of the 2013 International Electron Devices Meeting (IEDM), Washington, DC, USA, 9–11 December 2013; pp. 558–561.

349. Yoo, W.S.; Ishigaki, T.; Ueda, T.; Kang, K.; Kwak, N.Y.; Sheen, D.S.; Kim, S.S.; Ko, M.S.; Shin, W.S.; Lee, B.S.; et al. Grain size monitoring of 3D Flash memory channel poly-Si using multiwavelength Raman spectroscopy. In Proceedings of the 14th Annual Non-Volatile Memory Technology Symposium (NVMTS), Jeju, Korea, 27–29 October 2014; pp. 44–47.

350. Amit, I.; Englander, D.; Horvitz, D.; Sasson, Y.; Rosenwaks, Y. Density and energy distribution of interface states in the grain boundaries of polysilicon nanowire. *Nano Lett.* **2014**, *14*, 6190–6194.

351. Shamir, A.; Amit, I.; Englander, D.; Horvitz, D.; Rosenwaks, Y. Potential barrier height at the grain boundaries of a poly-silicon nanowire. *Nanotechnology* **2015**, *26*, 355201-1–355201-6.

352. He, D.; Okada, N.; Fortmann, C.M.; Shimizu, I. Carrier transport in polycrystalline silicon films deposited by a layer-by-layer technique. *J. Appl. Phys.* **1994**, *76*, 4728–4733.

353. Walker, A.J.; Herner, S.B.; Kumar, T.; Chen, E.H. On the conduction mechanism in polycrystalline silicon thin-film transistors. *IEEE Trans. Electron Devices* **2004**, *51*, 1856–1866.

354. Shim, K.S.; Choi, E.S.; Jung, S.W.; Kim, S.H.; Yoo, H.S.; Jeon, K.S.; Joo, H.S.; Oh, J.S.; Jang, Y.S.; Park, K.J.; et al. Inherent issues and challenges of program disturbance of 3D NAND Flash cell. In Proceedings of the 2012 4th IEEE International Memory Workshop (IMW), Milan, Italy, 20–23 May 2012; pp. 95–98.

355. Yoo, H.; Choi, E.; Oh, J.; Park, K.; Jung, S.; Kim, S.; Shim, K.; Joo, H.; Jeon, K.; Seo, M.; et al. Modeling and optimization of the chip level program disturbance of 3D NAND Flash memory. In Proceedings of the 2013 5th IEEE International Memory Workshop (IMW), Monterey, CA, USA, 26–29 May 2013; pp. 147–150.

356. Komori, Y.; Kido, M.; Kito, M.; Katsumata, R.; Fukuzumi, Y.; Tanaka, H.; Nagata, Y.; Ishiduki, M.; Aochi, H.; Nitayama, A. Disturbless Flash memory due to high boost efficiency on BiCS structure and optimal memory film stack for ultra high density storage device. In Proceedings of the 2008 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 15–17 December 2008; pp. 847–850.

357. Seo, J.Y.; Kim, Y.; Park, B.G. New program inhibition scheme for high boosting efficiency in three-dimensional NAND array. *Jpn. J. Appl. Phys.* **2014**, *53*, 070304-1–070304-3.

358. Choe, B.I.; Park, B.G.; Lee, J.H. Body doping profile of select device to minimize program disturbance in three-dimensional stack NAND Flash. *Jpn. J. Appl. Phys.* **2013**, *52*, 06GE02-1–06GE02-6.

359. Chang, Y.M.; Chang, Y.H.; Kuo, T.W.; Li, H.P.; Li, Y.C. A disturb-alleviation scheme for 3D Flash memory. In Proceedings of the 2013 IEEE/ACM International Conference on Computer-Aided Design (ICCAD), San Jose, CA, USA, 18–21 November 2013; pp. 421–428.

360. Chang, Y.M.; Chang, Y.H.; Kuo, T.W.; Li, Y.C.; Li, H.P. Disturbance relaxation for 3D Flash memory. *IEEE Trans. Comput.* **2016**, *65*, 1467–1483.

361. Chang, K.P.; Lue, H.T.; Chen, C.P.; Chen, C.F.; Chen, Y.R.; Hsiao, Y.H.; Hsieh, C.C.; Shih, Y.H.; Yang, T.; Chen, K.C.; et al. Memory architecture of 3D vertical gate (3DVG) NAND Flash using plural island-gate SSL decoding method and study of its program inhibit characteristics. In Proceedings of the 2012 4th IEEE International Memory Workshop (IMW), Milan, Italy, 20–23 May 2012; pp. 25–28.

362. Yeh, T.H.; Chen, W.C.; Hsu, T.H.; Du, P.Y.; Hsieh, C.C.; Lue, H.T.; Shih, Y.H.; King, Y.C.; Lu, C.Y. Z-interference and Z-disturbance in vertical gate-type 3-D NAND. *IEEE Trans. Electron Devices* **2016**, *63*, 1047–1053.

363. Choe, B.I.; Lee, J.K.; Park, B.G.; Lee, J.H. Suppression of read disturb fail caused by boosting hot carrier injection effect for 3-D stack NAND Flash memories. *IEEE Electron Device Lett.* **2014**, *35*, 42–44.

364. Kim, Y.; Kang, M. Down-coupling phenomenon of floating channel in 3D NAND Flash memory. *IEEE Electron Device Lett.* **2016**, *37*, 1566–1569.