Citation: Xu, Q., Gong, P., Chen, T., Michael, J. & Li, S. (2015). Modelling and characterization of NAND flash memory channels. Measurement, 70, pp. 225-231. doi: 10.1016/j.measurement.2015.04.003

Modelling and Characterization of NAND Flash Memory Channels

Quan Xu[a], Pu Gong[a], Thomas M. Chen[a,*], John Michael[b], Shancang Li[c]

[a]*School of Mathematics, Computer Science and Engineering, City University London, Northampton Square, London EC1V 0HB, UK*
[b]*iStorage Limited, 13 Alperton Lane, Perivale, Middlesex UB6 8DH, UK*
[c]*Institute for Informatics and Digital Innovation, Edinburgh Napier University, Edinburgh, EH10 5DT, UK*

Abstract

The threshold voltage distribution after ideal programming in NAND flash memory cells is usually distorted by a combination of the random telegraph noise (RTN), cell-to-cell Interference (CCI), and the retention process. To decide the original bits more accurately in this scenario, a precise channel model shall be utilized on the basis of the measured threshold voltages. This paper aims to characterize these various distortions occurring in multi-level cell (MLC) flash memories. A mathematical description of the overall distribution for the total flash channel distortion is presented. The final threshold voltage distribution for each symbol of MLC flash is also characterized, which is important for calculating the exact soft decisions of cell bits and the application of advanced flash error correction. The results of the theoretical analysis have been validated through Monte Carlo simulations of the flash channel.

Keywords: NAND flash, error correction codes, flash channel, soft decisions

---

1. Introduction

NAND flash memory is becoming essential as the storage media to a range of applications today, such as flash drive, solid state disks, mobile phone, etc.

---

[*] Corresponding author
Email addresses: `Quan.Xu.1@city.ac.uk` (Quan Xu), `Pu.Gong.1@city.ac.uk` (Pu Gong), `Tom.Chen.1@city.ac.uk` (Thomas M. Chen), `john@istorage-uk.com` (John Michael), `S.Li@napier.ac.uk` (Shancang Li)

Preprint submitted to Measurement *April 2, 2015*

Generally speaking, the original information is firstly encoded with an error correction code, and the resulted bit sequence is stored to flash memories by programming the cell threshold voltages to different levels. Measurement over the flash memory channel is then used for the decoding and decision-making of the original bits.

As program-erase (P/E) cycles and the data retention time increase, the threshold voltage distribution of sub-20 nm NAND flash memory is commonly distorted severely, making the measured channel outputs extremely unreliable. On the other hand, advanced soft decision-based error correction codes (ECC) such as low density parity check codes (LDPC) [1, 2, 3] are gradually replacing traditional ECC using hard decisions in the current flash memory design practice in order to compensate for the high raw bit error probability. Decoding of these soft decision-based algorithms depends heavily on the accuracy of the reliability information obtained by multiple memory measurements, or a channel model with exactly characterized threshold voltage distribution.

Generally, Gaussian-based channel models are widely used in the study of flash coding. For instance, pulse amplitude modulation (PAM) with Gaussian noise of same variances is used by Wang et al. to model flash cell threshold voltage levels [4]. Sun et al. approximated the flash channel with a similar model except with different variances for each level [5]. A similar model was used by Zhou et al. to explore the advantages of dynamic thresholds in flash reading [6]. All of these Gaussian-based channel models are relatively simple allowing researchers to focus more on the coding aspects. However, the fact that the real distribution is far different from the approximate model degrades the correctness of reliability information. Therefore, other researchers have sought to characterize the final voltage distribution more accurately based on these noise sources. Lee and Sung developed parameter estimation algorithms to find the means and variances of the threshold voltage, which is more exact because the parameters are estimated separately for each symbol. However, the final distributions are still approximated as Gaussian mixtures [7]. Dong et al. treated the cell-to-cell interference as the dominant distortion and mathematically derived the final voltage distribution, but the results were approximate due to the neglect of the other noises [8]. Another idea is to observe the channel from the statistical view and ignore the probability forms of distortions. The channel model based on this idea has been proposed recently by Moon et al. and successfully used to instruct the error correction coding [9, 10]. However, such statistical way is not capable to provide clear mathematical descriptions of the channel noises and final threshold voltage distribution.

This paper derives exact probability density functions for the combined channel noises and the final cell threshold voltage that used in calculating the soft channel information. The approach here uses the characteristic function of each noise distribution to determine the cumulative influence caused by all noises. For the case of 4-level MLC flash memory, the noise distribution is calculated and compared to simulation results. The final distributions of each symbol are determined in explicit formulas providing a way to calculate the soft information. Simulation results over the channel with specified P/E cycling number and retention time demonstrate the theoretical deductions are consistent with the practical channel outputs.

This paper is organized as follows. Section 2 explains the basic operations and noise sources of NAND flash memory. In Section 3, the various noises are explained, and characteristic functions are used to investigate the channel distribution. The final distributions for each symbol in 4-level MLC are given as a case study. Simulation results are also presented in this section.

2. Noises in NAND Flash Memory

The information stored in NAND flash memory is physically expressed in the cell threshold voltages that are programmed by injecting a certain amount of electrical charge into the memory cells. An erase operation involves removing the charges before programming a memory cell, which sets its threshold voltage to the lowest voltage window. Due to operation variability, the threshold voltage distribution of erased cells tends to have a wide Gaussian-like distribution, which can be approximately modelled as

$$p_e(x) = \frac{1}{\sigma_e\sqrt{2\pi}} e^{-(x-\mu_e)^2/2\sigma_e^2} \tag{1}$$

where μ_e and σ_e are the mean and standard deviation of the erased state.

For programming, a scheme called incremental stair pulse programming (ISPP) is used to achieve a tight threshold voltage bound for the representation of each symbol (or level) of the MLC. Let V_p denote the verify voltage of the target programmed level, and ΔV_{pp} the program step voltage. Ideally, the ISPP results in an uniform distribution over $[V_p, V_p + \Delta V_{pp}]$ which has width of ΔV_{pp}. Suppose V_p and $V_p + \Delta V_{pp}$ for the k-th programmed level are denoted as $V_l^{(k)}$ and $V_r^{(k)}$, respectively. The threshold voltage distribution of the kth programmed level after ideal programming can be modelled as

$$p_p^{(k)}(x) = \begin{cases} \frac{1}{\Delta V_{pp}}, & V_l^{(k)} \le x \le V_r^{(k)} \\ 0, & else \end{cases} \tag{2}$$

Thus, the initial cell threshold voltages of K-level MLC flash memory, denoted as $V_i^{(k)}(0 \le k \le K-1)$ have the following distributions

$$V_i^{(k)} \sim p_i^{(k)}(x) = \begin{cases} p_e(x), & k = 0 \\ p_p^{(k)}(x), & 1 \le k \le K-1 \end{cases} \tag{3}$$

Nevertheless, the above ideal distribution can be significantly distorted in practice by three types of noises: RTN, CCI, and charge loss in retention. These noises are shown in the MLC flash channel model in Fig. 1. In the model, $V_f^{(k)}(0 \le k \le K-1)$ is the final cell threshold voltage, and ΔV_{RTN}, ΔV_{CCI} and ΔV_{ret} are the threshold voltage shifts caused by RTN, CCI and retention noise, respectively.

RTN is caused by the electrons capture and emission developed over P/E cycling, and it directly results in the threshold voltage shift and fluctuation. The probability density function of ΔV_{RTN} can be modeled as a symmetric

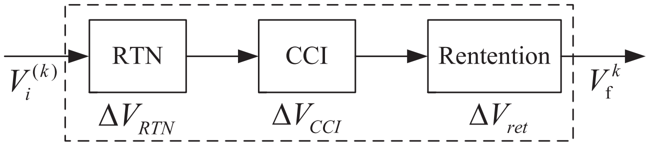

Figure 1: Noise and Interferences of the MLC flash memory

exponential function

$$\Delta V_{RTN} \sim p_r(x) = \frac{1}{2\lambda_r} e^{-|x|/\lambda_r} \tag{4}$$

where the parameter λ_r scales with the P/E cycling number N in an approximate power-law fashion, i.e., λ_r is approximately proportional to N^{α}.

Due to the parasitic capacitor-coupling effect, the threshold voltage shift of one cell affects the neighboring cells, resulting in CCI. The threshold voltage shift of a victim cell caused by CCI, denoted ΔV_{CCI}, can be estimated as

$$\Delta V_{CCI} = \sum_l \delta V^{(l)} \gamma^{(l)} \tag{5}$$

where $\delta V^{(l)}$ represents the threshold voltage shift of one interfering cell which is programmed after the victim cell, and γ^l is the coupling ratio subject to the parasitic capacitance between the interfering cell and the victim cell.

In the NAND all-bit-line structure, a victim cell is affected by three surrounding cells that are programmed after it, one along the vertical direction with γ_y while two along the two diagonal directions with γ_{xy}. Here we ignore the CCI from the two diagonal directions and treat γ_y as a constant to simplify the mathematical derivations in Section 3.

In addition to RTN and CCI, the cell threshold voltage may be reduced by interface trap recovery and electron detrapping. This is referred to as the data retention limitation. It has been demonstrated that ΔV_{ret} can be approximately modelled as a Gaussian distribution $\mathcal{N}(\mu_d, \sigma_d^2)$, where both μ_d and σ_d^2 scale to N in an approximate power-law fashion, and scale to the retention time t in a logarithmic fashion. Besides, the significance of μ_d and σ_d^2 is also proportional to the initial threshold voltage $V_i^{(k)}$.

3. Characterization of Flash Memory Channel

As mentioned earlier, the goal of this work is to characterize the flash channel exactly in terms of the probability density functions of noises and cell threshold voltages. A mathematical formulation for the overall distributions of the noises shown in the channel model presented in Section 2 will be derived. Then the final threshold voltage distribution of each level of the memory will be found.

3.1. Distributions of Flash Channel Noises

Let ΔV denote the threshold voltage shift induced by noises, and $\widetilde{p}(x)$ the probability density function of ΔV. In the channel model shown in Fig. 1, ΔV is the sum of the three shifts:

$$\begin{aligned} \Delta V &= \Delta V_{RTN} + \Delta V_{CCI} - \Delta V_{ret} \\ &= \Delta V_{RTN} + \gamma_y \cdot \delta V - \Delta V_{ret} \end{aligned} \tag{6}$$

Let $\Delta V_{(n)}$ represent the threshold voltage shift when the interfering cell in the vertical direction is being programmed to the nth state, and let $\widetilde{p}_{(n)}(x)$ represent the probability density function of $\Delta V_{(n)}$. We have

$$\Delta V_{(n)} = \Delta V_{RTN} + \gamma_y (V_{(n)} - V_{(0)}) - \Delta V_{ret} \tag{7}$$

Assuming that the interfering cells have the same probability to be programmed to each symbol, the overall distribution of total noises induced threshold voltage shifts can be approximated as

$$\Delta V \sim \widetilde{p}(x) = \frac{1}{K} \sum_{n=0}^{K-1} \widetilde{p}_{(n)}(x) \tag{8}$$

Next, we first derive $\widetilde{p}_{(n)}(x)$ and use it to find $\widetilde{p}(x)$ for the overall distribution according to (8). Considering all K possible levels for n, (7) can be rewritten as

$$\Delta V_{(n)} = \begin{cases} \Delta V_{RTN} - \Delta V_{ret}, & n = 0 \\ \Delta V_{RTN} + \gamma_y V_i^{(n)} - \gamma_y V_i^{(0)} - \Delta V_{ret}, & 1 \leq n \leq K-1 \end{cases} \tag{9}$$

In order to obtain $\widetilde{p}_{(n)}(x)$, we need to calculate the convolution of probability density functions of all components in the equation above, but multiple convolutions is computationally difficult. As a simpler alternative, we use characteristic functions instead of convolutions. The characteristic functions of the major components in (9) are the following:

$$\Delta V_{RTN} \rightarrow c_{RTN}(\xi) = \frac{1}{1 + \lambda_r^2 \xi^2} \tag{10}$$

$$-\Delta V_{ret} \rightarrow c_{ret}(\xi) = e^{-\frac{1}{2}\sigma_d^2 \xi^2} \cdot e^{-i\xi\mu_d} \tag{11}$$

$$-\gamma_y V_i^{(0)} \rightarrow c_e(\xi) = e^{-\frac{1}{2}\sigma_e^2 \gamma_y^2 \xi^2} \cdot e^{-i\xi\mu_e\gamma_y} \tag{12}$$

$$\begin{aligned} \gamma_y V_i^{(n)} \rightarrow c_n(\xi) = & \frac{1}{i\xi\gamma_y \Delta V_{pp}} \left(e^{i\xi\gamma_y V_r^{(n)}} - e^{i\xi\gamma_y V_l^{(n)}} \right), \\ & 1 \leq n \leq K-1 \end{aligned} \tag{13}$$

For $1 \leq n \leq K-1$, we split the right side of (9) into two parts: $\Delta V_{(n)} = V_1 + V_2$ with $V_1 = \Delta V_{RTN} + \gamma_y V_{(n)}$ and $V_2 = -\gamma_y V_{(0)} - \Delta V_{ret}$. Next consider the probability density function of V_1. The characteristic function of V_1, represented as $c_1(\xi)$, is given by

$$\begin{aligned} c_1(\xi) &= c_{RTN}(\xi) \cdot c_n(\xi) \\ &= \frac{1}{i\gamma_y \Delta V_{pp}} \cdot \frac{1}{(1 + \lambda_r^2 \xi^2)\xi} \left[e^{i\gamma_y V_r^{(n)} \xi} - e^{i\gamma_y V_l^{(n)} \xi} \right] \end{aligned} \tag{14}$$

Therefore, the probability density function of V_1, denoted as $f_1(x)$, can be calculated as

$$\begin{aligned} f_1(x) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-i\xi x} c_1(\xi) d\xi \\ &= \frac{1}{2\gamma_y \Delta V_{pp}} \left[q(t_1(x)) - q(t_2(x))\right] \end{aligned} \tag{15}$$

where

$$q(t) = \frac{1}{\pi i}\int_{-\infty}^{+\infty} \frac{1}{(1+\lambda_r^2\xi^2)\xi} e^{i\xi t} d\xi \tag{16}$$

and

$$\begin{bmatrix} t_1(x) \\ t_2(x) \end{bmatrix} = \begin{bmatrix} \gamma_y V_r^{(n)} - x \\ \gamma_y V_l^{(n)} - x \end{bmatrix} \tag{17}$$

Define

$$Q(z) = \frac{1}{\lambda_r^2((1/\lambda_r^2) + z^2)z} e^{izt} \tag{18}$$

which have three simple poles in the real axis: $z_0 = 0$, $z_1 = i/\lambda_r$ and $z_2 = -i/\lambda_r$. Based on residue theory, we get the result below for the integral in (16):

$$q(t) = \begin{cases} 1 - e^{-\frac{1}{\lambda_r}t} & t \geq 0 \\ -1 + e^{\frac{1}{\lambda_r}t} & t < 0 \end{cases} \tag{19}$$

Next we consider the probability density function of V_2. The characteristic function of V_2, represented as $c_2(\xi)$, is given by

$$\begin{aligned} c_2(\xi) &= c_e(\xi) \cdot c_{ret}(\xi) \\ &= e^{-\frac{1}{2}(\sigma_e^2\gamma_y^2 + \sigma_d^2)\xi^2} \cdot e^{i(-\mu_d - \mu_e\gamma_y)\xi} \end{aligned} \tag{20}$$

The probability density function of V_2, denoted as $f_2(x)$, can be further obtained by

$$\begin{aligned} f_2(x) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-i\xi x} c_2(\xi) d\xi \\ &= \frac{1}{\sqrt{2\pi(\sigma_e^2\gamma_y^2+\sigma_d^2)}} e^{\frac{-(x+\mu_d+\mu_e\gamma_y)^2}{2(\sigma_e^2\gamma_y^2+\sigma_d^2)}} \end{aligned} \tag{21}$$

With the results above, $\widetilde{p}_{(n)}(x)$ $(1 \leq n \leq K-1\ )$ can be determined by convolving $f_1(x)$ and $f_2(x)$:

$$\begin{aligned} \widetilde{p}_{(n)}(x) &= f_1(x) \otimes f_2(x) \\ &= \int_{-\infty}^{+\infty} f_1(\tau) f_2(x-\tau) d\tau \end{aligned} \tag{22}$$

Let $B = -\mu_d - \mu_e\gamma_y$, $C = \sqrt{\sigma_e^2\gamma_y^2 + \sigma_d^2}$, and define

$$\zeta(x,\omega,\upsilon) = erf\left(\frac{-x+\omega}{\sqrt{2}\upsilon}\right) \tag{23}$$

The final expression of $\widetilde{p}_{(n)}(x)$ $(1 \leq n \leq K-1\ )$ is given in (24).

$$\begin{aligned} \widetilde{p}_{(n)}(x) =& \frac{1}{2\gamma_y\Delta V_{pp}}\left[q(t_1(x)) \otimes f_2(x) - q(t_2(x)) \otimes f_2(x)\right] \\ =& \frac{1}{2\gamma_y\Delta V_{pp}} \cdot \left[\zeta(x, B+\gamma_y V_r^{(n)}, C) - \zeta(x, B+\gamma_y V_l^{(n)}, C)\right] + \\ & \frac{e^{\frac{C^2}{2\lambda_r^2}}}{4\gamma_y\Delta V_{pp}} \cdot \left[e^{\frac{1}{\lambda_r}\left(-x+B+\gamma_y V_r^{(n)}\right)}\left(1-\zeta(x, B+\gamma_y V_r^{(n)}+\frac{C^2}{\lambda_r}, C)\right)\right. \\ & + e^{-\frac{1}{\lambda_r}\left(-x+B+\gamma_y V_l^{(n)}\right)}\left(1+\zeta(x, B+\gamma_y V_l^{(n)}-\frac{C^2}{\lambda_r}, C)\right) \\ & - e^{-\frac{1}{\lambda_r}\left(-x+B+\gamma_y V_r^{(n)}\right)}\left(1+\zeta(x, B+\gamma_y V_r^{(n)}-\frac{C^2}{\lambda_r}, C)\right) \\ & \left. - e^{\frac{1}{\lambda_r}\left(-x+B+\gamma_y V_l^{(n)}\right)}\left(1-\zeta(x, B+\gamma_y V_l^{(n)}+\frac{C^2}{\lambda_r}, C)\right)\right], (1 \leq n \leq K-1) \end{aligned} \tag{24}$$

For $n = 0$, the probability density function $\widetilde{p}_{(n)}(x)$ can be calculated by convolving $p_r(x)$ and $p_{ret}(x)$ as below:

$$\begin{aligned}\widetilde{p}_{(0)}(x) &= \int_{-\infty}^{+\infty} p_r(x+\tau)p_{ret}(\tau)d\tau \\ &= \frac{1}{2\lambda_r}\cdot\frac{1}{\sqrt{2\pi}\sigma_d}\int_{-\infty}^{+\infty} e^{-\frac{|\tau|}{\lambda_r}}\cdot e^{-\frac{(x-\tau+\mu_d)^2}{2\sigma_d^2}}d\tau\end{aligned} \tag{25}$$

The final expression of $\widetilde{p}_{(0)}(x)$ is given in (26).

$$\begin{aligned}\widetilde{p}_{(0)}(x) = \frac{1}{4\lambda_r}\cdot e^{\frac{\sigma_d^2}{2\lambda_r^2}}\cdot \Bigg[& e^{\frac{1}{\lambda_r}(x+\mu_d)}\left(1+\zeta(x,-\mu_d-\frac{1}{\lambda_r}\sigma_d^2,\sigma_d)\right) + \\ & e^{-\frac{1}{\lambda_r}(x+\mu_d)}\left(1-\zeta(x,-\mu_d+\frac{1}{\lambda_r}\sigma_d^2,\sigma_d)\right)\Bigg]\end{aligned} \tag{26}$$

According to (8), the overall distribution of total noises-induced threshold voltage shifts, i.e., $\widetilde{p}(x)$ is given in (27).

$$\begin{aligned}\widetilde{p}(x) =& \frac{1}{K}\times\sum_{n=0}^{K-1}\widetilde{p}_{(n)}(x) \\ =& \frac{1}{2\gamma_y\Delta V_{pp}K}\cdot\sum_{n=1}^{K-1}\left[\zeta(x,B+\gamma_y V_r^{(n)},C)-\zeta(x,B+\gamma_y V_l^{(n)},C)\right] + \\ & \frac{e^{\frac{C^2}{2\lambda_r^2}}}{4\gamma_y\Delta V_{pp}K}\cdot\sum_{n=1}^{K-1}\Bigg[e^{\frac{1}{\lambda_r}\left(-x+B+\gamma_y V_r^{(n)}\right)}\left(1-\zeta(x,B+\gamma_y V_r^{(n)}+\frac{C^2}{\lambda_r},C)\right) \\ & + e^{-\frac{1}{\lambda_r}\left(-x+B+\gamma_y V_l^{(n)}\right)}\left(1+\zeta(x,B+\gamma_y V_l^{(n)}-\frac{C^2}{\lambda_r},C)\right) \\ & - e^{-\frac{1}{\lambda_r}\left(-x+B+\gamma_y V_r^{(n)}\right)}\left(1+\zeta(x,B+\gamma_y V_r^{(n)}-\frac{C^2}{\lambda_r},C)\right) \\ & - e^{\frac{1}{\lambda_r}\left(-x+B+\gamma_y V_l^{(n)}\right)}\left(1-\zeta(x,B+\gamma_y V_l^{(n)}+\frac{C^2}{\lambda_r},C)\right)\Bigg] \\ & + \frac{1}{4\lambda_r K}\cdot e^{\frac{\sigma_d^2}{2\lambda_r^2}}\Bigg[e^{\frac{1}{\lambda_r}(x+\mu_d)}\left(1+\zeta(x,-\mu_d-\frac{1}{\lambda_r}\sigma_d^2,\sigma_d)\right) \\ & + e^{-\frac{1}{\lambda_r}(x+\mu_d)}\left(1-\zeta(x,-\mu_d+\frac{1}{\lambda_r}\sigma_d^2,\sigma_d)\right)\Bigg]\end{aligned} \tag{27}$$

Since channel parameters (specifically B, C, σ_d, and μ_d) vary for different levels in the flash channel, the overall noise density functions in respect to all K levels (denoted as $L_0 \rightarrow L_{K-1}$) are not exactly the same. Denote the probability density functions of noise-induced threshold voltage shifts for $L_0 \rightarrow L_{K-1}$ as $\widetilde{p}^{(k)}(x)$, $(0 \leq k \leq K-1)$. Based on 2 bits/cell NAND flash memory and the parameters presented in the literature [8], we simulated the distributions of noise-induced threshold voltage shifts and compared the results with the theoretical counterparts found in (27).

Simulations: The normalized σ_e and μ_e of the erased state are set as 0.35 and 1.4, respectively. For the three programmed states, we set the normalized program step voltage ΔV_{pp} as 0.2, and the normalized verify voltages V_p as 2.6, 3.2, and 3.93, respectively. For the RTN distribution function $p_r(x)$, we set the parameter $\lambda_r = K_\lambda \cdot N^{0.5}$, where K_λ equals to 0.00025. Regarding CCI, we ignore the CCI from the two diagonal directions and also fix γ_y to a constant 0.08. For the function $\mathcal{N}(\mu_d, \sigma_d^2)$ to capture trap recovery and electron detrapping during retention, we assume that μ_d scales with $N^{0.5}$ and σ_d^2 scales with $N^{0.6}$, and both scale with $\ln(1+t/t_0)$, where t denotes the memory retention time and t_0 is an initial time set as 1 hour. In addition, both μ_d and σ_d^2 also depend on the initial threshold voltage. Hence we set that both approximately scale $K_s(x - x_0)$, where x is the initial threshold voltage, and x_0 and K_s are constants. Therefore, we have

$$\begin{cases} \mu_d = K_s(x - x_0) K_d N^{0.5} \ln(1 + \frac{t}{t_0}) \\ \sigma_d^2 = K_s(x - x_0) K_m N^{0.6} \ln(1 + \frac{t}{t_0}) \end{cases} \tag{28}$$

where we set $K_s = 0.38$, $x_0 = 1.4$, $K_d = 4 \times 10^{-4}$ and $K_m = 4 \times 10^{-6}$ [8].

Setting the P/E cycling number N as 1000, and the retention time as 1 year, we found the probability density based on (27) for channel noises occurring in all four levels, respectively, as illustrated in Fig. 2.

Accordingly, we carry out Monte Carlo simulations to obtain the cell threshold voltage distribution at different levels under 1000 P/E cycling and 1 year retention limit, as shown in Fig. 3, which gives nearly the same results as in

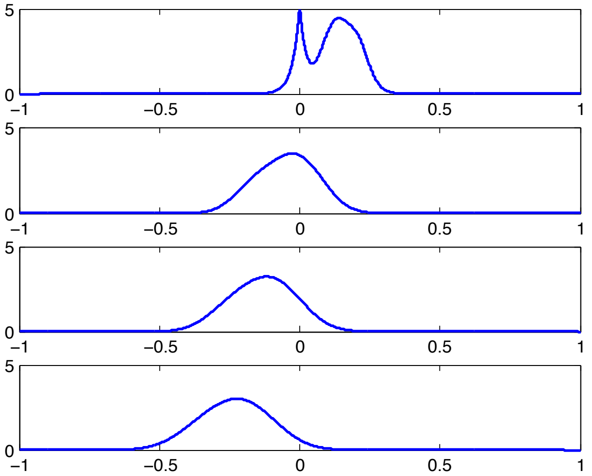

Figure 2: Theoretical curves to show the effects of RTN, CCI, and retention noises on memory cell threshold voltage distribution after 1K P/E cycling and 1 year retention.

Fig. 2. A close resemblance supports the correctness of the above theoretical derivations on the noises-induced threshold voltage shifts.

3.2. Cell Threshold Voltage Distribution

In this section, the threshold voltage distribution of each level is determined, which can be used to calculate the log-likelihood ratio (LLR) of each bit stored in the memory, and further used for flash coding [11].

As shown in Fig. 2, the curves for $L_1 \rightarrow L_3$ are bell-shaped, which is due to the fact that retention noise becomes dominant as the P/E cycling number and the retention time increase. Therefore, we intuitively expect a Gaussian distribution will be a good fit for these probability density functions. The formulas below give the curve-fitting parameters of the proposed normal distribution $(1 \leq k \leq K-1)$.

$$\mu_{(k)} = \int_{-\infty}^{+\infty} x \cdot \widetilde{p}^{(k)}(x) dx. \tag{29}$$

$$\sigma_{(k)}^2 = \int_{-\infty}^{+\infty} (x - \mu_{(k)})^2 \cdot \widetilde{p}^{(k)}(x) dx. \tag{30}$$

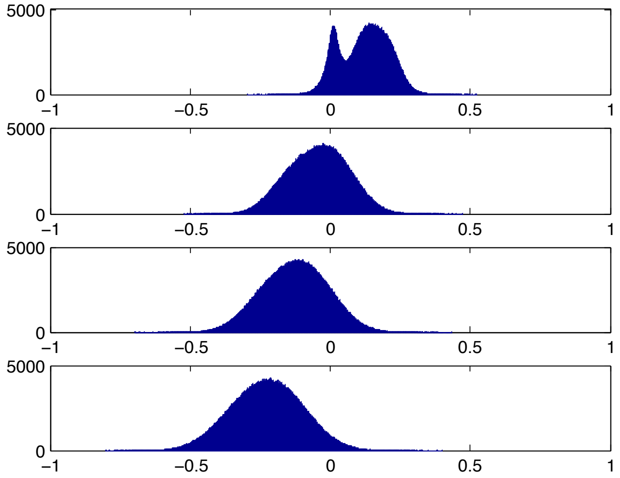

Figure 3: Simulated curves to show the effects of RTN, CCI, and retention noises on memory cell threshold voltage distribution after 1K P/E cycling and 1 year retention.

$$\begin{aligned} p^{(k)}(x) &= g^{(k)}(x) \otimes p_i^{(k)}(x) = g^{(k)}(x) \otimes p_p^{(k)}(x) = \int_{-\infty}^{+\infty} p_p^{(k)}(\tau) g^{(k)}(x-\tau) d\tau \\ &= \int_{V_l^{(k)}}^{V_r^{(k)}} \frac{1}{\Delta V_{pp}} \cdot \frac{1}{\sqrt{2\pi}\sigma_{(k)}} e^{-\frac{\left(x-\tau-\mu_{(k)}\right)^2}{2\sigma_{(k)}^2}} d\tau \\ &= \frac{1}{2\Delta V_{pp}} \left[\zeta(x, V_r^{(k)} + \mu_{(k)}, \sigma_{(k)}) - \zeta(x, V_l^{(k)} + \mu_{(k)}, \sigma_{(k)}) \right], 1 \leq k \leq K-1 \end{aligned} \tag{32}$$

Therefore, the probability density function of the curve-fitted normal distributions for $L_1 \rightarrow L_{K-1}$ are given by

$$g^{(k)}(x) = \frac{1}{\sqrt{2\pi}\sigma_{(k)}} e^{\frac{(x-\mu_{(k)})^2}{2\sigma_{(k)}^2}}, 1 \leq k \leq K-1. \tag{31}$$

Recall the initial distributions of $L_1 \rightarrow L_{K-1}$ after ideal programming in Eq. (3), we can obtain the final probability density functions, denoted as $p^{(k)}(x)$, by convoluting $g^{(k)}(x)$ and $p_i^{(k)}(x)$. The results are given in (32).

Since the retention noise has minor influence on the threshold voltage of L_0, the probability density function of L_0 shown in Fig. 2 is not bell-shaped, thus it would not be correct to use a normal distribution for curve fitting in this case. However, the final distribution of L_0 can be easily determined with the results in (27) as the distribution of L_0 after ideal programming is Gaussian. To this end, the two parameters: μ_d and σ_d in (27), should be revised as follows.

$$\begin{cases} \mu'_d = \mu_d - \mu_e \\ \sigma'_d = \sqrt{\sigma_e^2 + \sigma_d^2} \end{cases} \tag{33}$$

Accordingly, B and C, should be revised as well as follows.

$$\begin{cases} B' = -\mu'_d - \mu_e\gamma_y = -\mu_d - \mu_e\gamma_y + \mu_e \\ C' = \sqrt{\sigma_e^2\gamma_y^2 + \sigma_d'^2} = \sqrt{\sigma_e^2\gamma_y^2 + \sigma_d^2 + \sigma_e^2} \end{cases} \tag{34}$$

Consequently, the final threshold voltage distribution of L_0, i.e. $p^{(0)}(x)$, can be calculated in (35).

$$\begin{aligned} p^{(0)}(x) =& \widetilde{p}^{(0)}(x) \otimes p_e(x) \\ =& \frac{1}{2\gamma_y \Delta V_{pp} K} \cdot \sum_{n=1}^{K-1} \Big[\zeta(x, B' + \gamma_y V_r^{(n)}, C') - \zeta(x, B' + \gamma_y V_l^{(n)}, C') \Big] + \\ & \frac{e^{\frac{C'^2}{2\lambda_r^2}}}{4\gamma_y \Delta V_{pp} K} \cdot \sum_{n=1}^{K-1} \Big[e^{\frac{1}{\lambda_r}\left(-x+B'+\gamma_y V_r^{(n)}\right)} \Big(1 - \zeta(x, B' + \gamma_y V_r^{(n)} + \frac{C'^2}{\lambda_r}, C')\Big) \\ & + e^{-\frac{1}{\lambda_r}\left(-x+B'+\gamma_y V_l^{(n)}\right)} \Big(1 + \zeta(x, B' + \gamma_y V_l^{(n)} - \frac{C'^2}{\lambda_r}, C')\Big) \\ & - e^{-\frac{1}{\lambda_r}\left(-x+B'+\gamma_y V_r^{(n)}\right)} \Big(1 + \zeta(x, B' + \gamma_y V_r^{(n)} - \frac{C'^2}{\lambda_r}, C')\Big) \\ & - e^{\frac{1}{\lambda_r}\left(-x+B'+\gamma_y V_l^{(n)}\right)} \Big(1 - \zeta(x, B' + \gamma_y V_l^{(n)} + \frac{C'^2}{\lambda_r}, C')\Big) \Big] \\ & + \frac{1}{4\lambda_r K} \cdot e^{\frac{\sigma_d'^2}{2\lambda_r^2}} \Big[e^{\frac{1}{\lambda_r}(x+\mu'_d)} \Big(1 + \zeta(x, -\mu'_d - \frac{1}{\lambda_r}\sigma_d'^2, \sigma'_d)\Big) \\ & + e^{-\frac{1}{\lambda_r}(x+\mu'_d)} \Big(1 - \zeta(x, -\mu'_d + \frac{1}{\lambda_r}\sigma_d'^2, \sigma'_d)\Big) \Big] \end{aligned} \tag{35}$$

The aforementioned 4-level MLC flash and the related channel parameters have been used as well in this section to show the probability density functions derived above. The P/E cycling number is set as 1000 and retention time as 1 year. Fig. 4 illustrates the final threshold voltage distributions of all four levels according to (32) and (34). As shown, the channel modeled in this work is quite similar to the models used in the literatures [5, 4, 7, 8]. Nonetheless, as only reasonable Gaussian approximation is used and all the channel noises have been considered in the derivations above, the model proposed here is more precise to the real flash channel. Moreover, the exact formula of cell threshold voltage distribution, i.e., the results given in (34), makes the soft decisions calculated based on our work more accurate for the flash decoding.

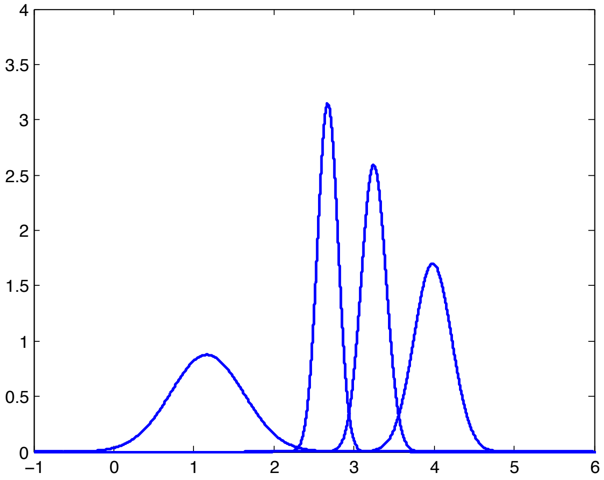

Figure 4: Cell threshold voltage distributions 4-level MLC flash memory

3.3. Calculation of Soft Decisions

Soft decisions, or in other words, the log-likelihood-ratios (LLRs) information are critical in decoding the advanced flash error correction codes. This section proposes the mathematical formulation for calculating LLRs based upon the threshold-voltage distributions presented above, and the results are used to instruct the implementation of decoding schemes in the design practice.

Assume each bit is programmed to a flash cell with equal probability, i.e., all the storage levels in one memory cell have equal *a priori* probability, and let

V_{th} represents the sensed threshold voltage of one memory cell, we can calculate the LLR of the ith bit stored in one cell as

$$L(b_i) = \log\frac{p(b_i = 1|V_{th})}{p(b_i = 0|V_{th})} = \log\frac{p(V_{th}|b_i = 1)}{p(V_{th}|b_i = 0)} \tag{36}$$

For K levels MLC flash memory, there are $N_b = \log_2(K)$ bits stored in each cell, and the probability density function of the threshold voltage of kth storage level, $p^{(k)}(x), 0 \le k \le K-1$, is known from (32) and (35). Let $\mathbb{S}_i$ denote the set of levels whose ith bit is 1. Hence, given the threshold voltage V_{th} of a cell, the LLR of each bit would be calculated as

$$L(b_i) = \log\frac{\sum_{k\in\mathbb{S}_i} p^k(V_{th})}{\sum_k p^k(V_{th}) - \sum_{k\in\mathbb{S}_i} p^k(V_{th})} \tag{37}$$

In respect to practical NAND flash memories, the cell threshold voltages cannot be precisely obtained while being sensed by the comparison with a series of reference voltages. Assume that the threshold voltage V_{th} falls into the range $(R_l, R_r]$ (where R_l and R_r are two adjacent reference voltages), we can estimate the corresponding LLR of the ith bit as

$$L(b_i) = \log\frac{\int\limits_{R_l}^{R_r} \sum_{k\in\mathbb{S}_i} p^{(k)}V_{th}}{\int\limits_{R_l}^{R_r} \sum_k p^{(k)}V_{th} - \int\limits_{R_l}^{R_r} \sum_{k\in\mathbb{S}_i} p^{(k)}V_{th}} \tag{38}$$

We assume the parameters σ_e, μ_e, σ_d, μ_d in (32) and (35) are known and treat the coupling ratio γ_y as a known constant. As the number of reference voltages in memory sensing is fixed, all possible LLRs can be calculated before the decoding starts. Therefore, only a lookup table is required in the run time to obtain the LLRs, which reduces the system complexity and improves the speed and throughput. Let K_s denote the number of reference voltages being used in memory sensing, the LLR lookup table only contains $N_b(K_s + 1)$ entries for N_b-bit/cell NAND flash memory [12].

4. Conclusions

We have derived the probability density function for the channel noises in the MLC NAND flash memories, and formulated the threshold voltage distribution of each symbol. This was possible by using characteristic functions for the various noises. The results presented here are novel in a theoretical way as the first time the flash channel has been characterized by incorporating the distributions of all the interferences and noises. Using the 4-level MLC flash as a case study, the derived threshold voltage distributions were demonstrated to be consistent with simulated channel outputs.

References

[1] R. Gallager, Low-density parity check codes, IRE Trans. Information Theory (1962) 21–28.

[2] X. Hu, B. Kumar, Z. Li, R. Barndt, Error Floor Estimation of Long LDPC Codes on Partial Response Channels, in: IEEE GLOBECOM, 2007, pp. 259–264.

[3] Q. Xu, T. Chen, Y. Hu, P. Gong, Write Pattern Format Algorithm for Reliable NAND-Based SSDs, IEEE Trans. onCircuits and Systems II: Express Briefs 61 (7) (2014) 516–520.

[4] J. Wang, T. Courtade, H. Shankar, R. Wesel, Soft information for LDPC decoding in flash: mutual-information optimized quantization, in: IEEE GLOBECOM, 2011, pp. 1–6.

[5] F. Sun, S. Devarajan, K. Rose, T. Zhang, Multilevel flash memory on-chip error correction based on trellis coded modulation, in: 2006 IEEE Int. Symp. on Circuits and Systems (ISCAS 2006), 2006.

[6] H. Zhou, A. Jiang, J. Bruck, Error-correcting schemes with dynamic thresholds in non-volatile memories, in: International Symposium on Information Theory, 2011, pp. 2143–2147.

[7] D. hwan Lee, W. Sung, Estimation of NAND Flash Memory Threshold Voltage Distribution for Optimum Soft-Decision Error Correction, IEEE Trans. on Signal Processing 61 (2) (2013) 440–449.

[8] G. Dong, Y. Pan, N. Xie, C. Varanasi, T. Zhang, Estimating Information-Theoretical NAND Flash Memory Storage Capacity and its Implication to Memory System Design Space Exploration, IEEE Trans. on Very Large Scale Integration (VLSI) Systems 20 (9) (2012) 1705–1714.

[9] J. Moon, J. No, S. Lee, S. Kim, S. Choi, Y. Song, Statistical Characterization of Noise and Interference in NAND Flash Memory, IEEE Trans. on Circuits and Systems I: Regular Papers 60 (8) (2013) 2153–2164.

[10] J. Moon, J. No, S. Lee, S. Kim, J. Yang, Statistical Analysis of Flash Memory Read Data, in: IEEE GLOBECOM, 2011, pp. 1–6.

[11] G. Dong, S. Li, T. Zhang, Using Data Postcompensation and Predistortion to Tolerate Cell-to-cell Interference in MLC NAND Flash Memory, IEEE Trans. on Circuit and System: Part I (2010) 2718–2728.

[12] G. Dong, N. Xie, T. Zhang, On the use of soft-decision error-correction codes in nand flash memory, IEEE Trans. on Circuits and Systems I: Regular Papers 58 (2) (2011) 429–439.