Flash Channel Estimation for Dynamic Voltage Allocation

Haobo Wang, Tsung-Yi Chen, Richard Wesel
UCLA Electrical Engineering
We model the NAND flash memory cell data storage process as [1]

\[ y = x + n_p + n_w + n_r \]

- \( x \): intended programmed state threshold voltage
- \( y \): sensed programmed state threshold voltage
- \( n_p \): programming noise
- \( n_w \): wear-out noise
- \( n_r \): retention noise

Programming Noise \((n_p)\)

- The uncertainty of the programmed threshold voltage immediately after program operation can be modeled by a Gaussian random variable.
- The variance of the programmed threshold voltage is larger when left in the erased state than when actively programmed [2, 3].

\[
f(n_p) = \begin{cases} 
N(0, \sigma_e^2) & \text{if } x = 0 \\
N(0, \sigma_p^2) & \text{if } x > 0 
\end{cases}
\]

where \(\sigma_e > \sigma_p\)


Wear-out Noise \( (n_w) \)

- Wear-out induces P/E-cycle-dependent threshold voltage shift as a result of bulk oxide and oxide interface traps generation and electron trapping/de-trapping during P/E cycling [4, 5, 6, 7].
- Trap behavior is modeled as random telegraph noise (RTN). This causes the distribution of measured thresholds features exponential tails [8].
- In some devices, the positive-shift tail is more significant than the negative-shift one, so we use an exponential distribution to model wear-out noise.

\[
f(n_w) = \begin{cases} 
  \frac{1}{\lambda} e^{-\frac{n_w}{\lambda}} & n_w \geq 0 \\
  0 & n_w < 0 
\end{cases}
\]

Retention Noise \( (n_r) \)

- Retention loss is the reduction of programmed threshold voltage over time caused primarily by electron de-trapping \([9, 10]\).

- Retention noise \( n_r \) is modeled as a Gaussian distribution \( N(\mu_r, \sigma_r^2) \) where the mean and variance are depending on time and number of traps.

- De-trapping scales approximately with the logarithm of post-cycling retention time and the number of traps before the post-cycling retention process \([10]\).

Sample PDF

Programming Noise Only

f (x + n_p)

Probability Density

Voltage
Sample PDF

Programming Noise Only

Programming and Wear-out Noise

\[ f(x + n_p) \]

\[ f(x + n_p + n_w) \]
Sample PDF

Programming Noise Only

Programming and Wear-out Noise

Programming, Wear-out and Retention Noise
Channel degradation is usually modeled as a function of the number of program/erase (P/E) cycles.

The charge passing through dielectrics actually causes the degradation [11].

We use accumulated voltage $V_{acc}$ as a more precise metric of wear-out instead of P/E cycles. $V_{acc}$ indicates the amount of charge passing through the dielectrics since the first write.

Normalized Accumulated Voltage

\[ V_{\text{acc}} = \sum_{j=1}^{N} (V_{p}^{(j)} - V_e) \]

- \( V_{\text{acc}} \): accumulated voltage over N P/E cycles,
- \( V_{p}^{(j)} \): programmed threshold voltage of the jth P/E cycle
- \( V_e \): threshold voltage of the erased state
- \( V_{\text{max}} \): the maximum of \( V_{p}^{(j)} - V_e \), \( \forall j \)

The normalized accumulated voltage is \( V_{\text{acc}} / V_{\text{max}} \).

When using fixed voltage levels, \( V_{\text{acc}} / V_{\text{max}} \approx \# \text{of PE Cycles} \).
Parameter Degradation Model

- Degradation Model
  - Wear-out noise:
    \[ \lambda = C_w + A_w \cdot \left( \frac{V_{acc}}{V_{max}} \right)^{0.62} \]
  - Retention noise:
    \[ \mu_r = -x \cdot \ln \left(1 + \frac{t}{t_0}\right) \cdot \left[ A_r \cdot \left( \frac{V_{acc}}{V_{max}} \right)^{0.62} + B_r \cdot \left( \frac{V_{acc}}{V_{max}} \right)^{0.3} \right] \]
    \[ \sigma_r^2 = 0.1x \cdot \ln \left(1 + \frac{t}{t_0}\right) \cdot \left[ A_r \cdot \left( \frac{V_{acc}}{V_{max}} \right)^{0.62} + B_r \cdot \left( \frac{V_{acc}}{V_{max}} \right)^{0.3} \right]^2 \]
Fixed threshold voltage allocation induces same amount of wear-out to Flash device in each P/E cycle.

Dynamic Voltage Allocation can reduce unnecessary wear-out, and thus increase lifetime by using lower threshold voltage for early writes [1].

The threshold voltages can be gradually increased as needed to combat channel degradation [1].

The optimize target is to maintain a minimum channel capacity as long as possible.

Dynamic Voltage Allocation (DVA) on MLC based on the perfect prediction of future channel condition can extend the block’s lifetime by 93%.
Practical DVA using Histogram-based Channel Estimation

- Histogram Measurement
- Parameter Estimation
- Dynamic Voltage Allocation
- Channel Model
Channel parameters can be estimated from measured histograms [12].

Channel parameter estimation workflow:

- Measured Histogram
- Channel Model
- Least Square Algorithm
- Estimated Parameters

Parameter Vector

- **Parameter Vector**

  - \( \alpha = [\lambda, \sigma_{\text{programming}}, \sigma_{\text{erase}}, \sigma_{\text{retention}}, \mu_{\text{retention}}] \)

  - We actually estimate \( [\lambda, \sigma_p, \sigma_e, m_r, n_r] \), where

    \[
    \mu_{\text{retention}} = -x \cdot n_r \\
    \sigma_{\text{retention}}^2 = x \cdot m_r^2
    \]
Estimation Objective Function

- Estimation Objective Function is the squared Euclidean distance between predicted histogram and measured histogram.

\[
C_M = \sum_{i=0}^{M-1} \left( \frac{\hat{N}_{\text{bin},i} - N_{\text{bin},i}}{N} \right)^2
\]

- \( N \): total number of cells in a page
- \( N_{\text{bin},i} \): total number of cells in ith bin of measured histogram
- \( \hat{N}_{\text{bin},i} \): total number of cells in ith bin by estimation
- \( M \): total number of bins
Parameter Estimation

\[ \alpha = [\lambda, \sigma_p, \sigma_e, m_r, n_r] \Rightarrow [0.0500, 0.1997, 0.6000, 0.1002, -0.1000] \]
Voltage Levels Adapted to Degraded Channel

Histogram Measurement

Parameter Estimation

Dynamic Voltage Allocation

Probability Density

Voltage

Probability Density

Voltage

Voltage Levels Adapted to Degraded Channel
More about Least Square Algorithms

- **Objective**
  - Minimize the cost function.

- **Algorithm 1 – Gradient Descent**
  - Follow the descending gradient with a fixed step size.

- **Algorithm 2 – Gauss–Newton Algorithm**
  - Take each step based on quadratic approximation at current point.

- **Algorithm 3 – Levenberg–Marquardt Algorithm**
  - Rotate Gauss-Newton increment vector toward the direction of descending gradient.
Least Square Algorithm Comparison

Convergence Speed

Convergence Success Rate

Gradient Descent
Gauss-Newton
Levenberg-Marquart
Histogram Types

- **Equal Interval Histogram** [12]
  - Not actually equal. Bins covering erased state distribution can be slightly wider.

- **Maximum Mutual Information (MMI) Histogram** [13]

- **Equal Probability Histogram**
  - Each bin has the same number of cells.


Equal Interval Histogram Does not Adapt Well to Retention Loss.
MMI Histogram Has Poor Resolution.
Equal Interval Histogram [12]
- Equal interval histogram does not adapt well to retention loss.

Maximum Mutual Information Histogram [13]
- This histogram optimizes decoder performance, but may not be the best for channel parameter estimation.

Equal Probability Histogram
- Every bin has an equal number of cells, intuitively good for parameter estimation.

Prediction requires knowledge of parameter degradation models.

In reality, the parameters derived from degradation models may not always represent the performance of every block.

Dynamic Voltage Allocation can utilize current channel condition with a fixed margin.

The fixed margin should be set to allow channel capacity remain above the optimization target until the next adaptation.
There is a 500 PE loss.

Dynamic Voltage Allocation using current channel condition can still extend the block’s lifetime by 77%.
SIMULATION RESULT
Comparison of fixed margin DVA with perfect channel knowledge and estimation of channel using idealized equal probability histogram of 9 bins.
Comparison of fixed margin DVA with channel estimation using idealized equal probability histogram of 9 bins and realistic Monte Carlo generated histogram.
CoDESS is founded in 2013.

Webpage: http://www.uclacodess.org/

Mission:
- Push the frontiers in emerging data storage systems through integrated research program.
- Create highly-trained workforce of graduate students and post-doctoral researchers.

For more information, please email to
- Prof. Lara Dolecek (dolecek@ee.ucla.edu)
- Prof. Richard Wesel (wesel@ee.ucla.edu)