Generative Modeling of NAND Flash Voltage Levels

Ziwei Liu, Yi Liu, and Paul H. Siegel   University of California, San Diego

UC San Diego

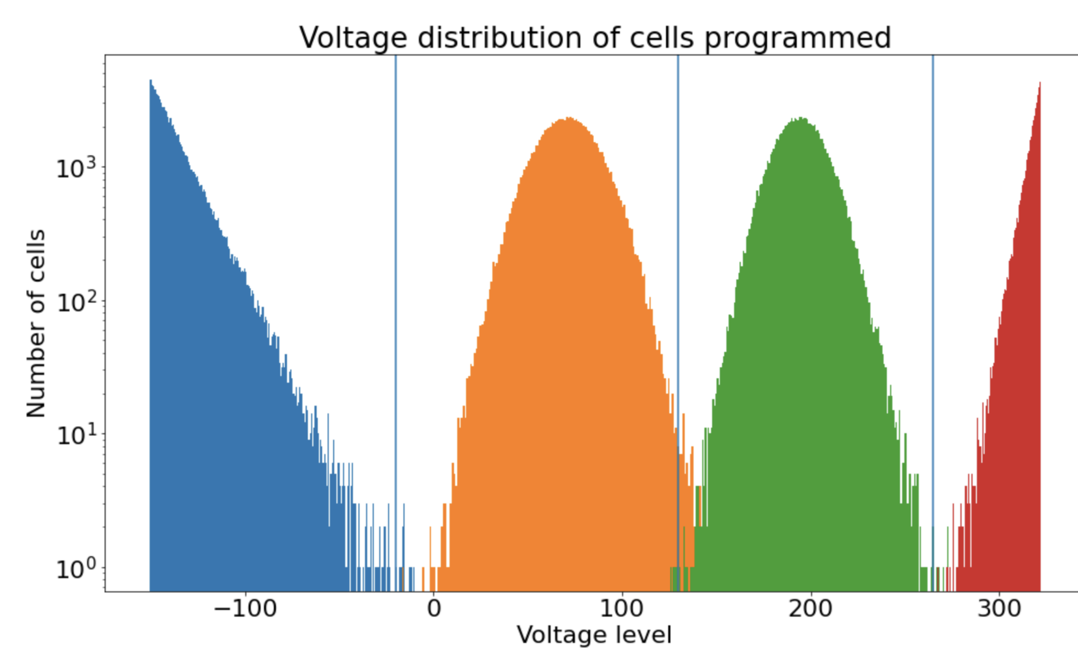
CMRR Center for Memory and Recording Research

Introduction

- Both flash memory hardware characterization and error-correcting code design rely heavily on flash memory cell voltage level data, which is time-consuming to acquire.
- We propose a **voltage level generator** to generate a large amount of realistic NAND flash memory cell voltage level samples using a relatively small number of measured voltage levels.
- This learned generator can generate distributions of authentic voltage level
  - over a range of possible Program/Erase (P/E) cycles
  - for each specified program level.

The Experimental Dataset

- The experimental dataset used in this demonstration in obtained from a **Normal-Laplace model** of MLC flash memory voltage level distributions.
- This Normal-Laplace model can generate realistic MLC NAND flash memory voltage level distributions
  - for P/E cycles from 10 to 1000
  - for each program level



Distribution of voltage levels generated from the Normal-Laplace model

- When the training is enforced, we draw the same amount of samples from the Normal-Laplace generator and our deep learning-based generator, and then compute the loss function.

GMMN

- **Generative Moments Matching Network (GMMN)** measures the similarity of the generated voltage level distribution and the target voltage level distribution using Generative **Maximum Mean Discrepancy (MMD)**.
- MMD computes the mean squared difference of the statistics of the set of samples from the generated distribution and the target distribution.

$$\mathcal{L}_{MMD^2} = \|\frac{1}{N}\sum_{i=1}^{N}\phi(x_i) - \frac{1}{M}\sum_{j=1}^{M}\phi(y_j)\|^2 \quad (1)$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{i'=1}^{N}\phi(x_i)^T\phi(x_{i'}) - \frac{2}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\phi(x_i)^T\phi(y_j) + \frac{1}{M^2}\sum_{j=1}^{M}\sum_{j'=1}^{M}\phi(y_j)^T\phi(y_{j'}) \quad (2)$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{i'=1}^{N}k(x_i, x_{i'}) - \frac{2}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}k(x_i, y_j) + \frac{1}{M^2}\sum_{j=1}^{M}\sum_{j'=1}^{M}k(y_j, y_{j'})$$

- When (1) is expanded in (2), each term only involves inner products between the ϕ vectors, and therefore the **Gaussian kernel trick** can be applied, with

$$k(x, x') = \exp(-\frac{1}{2\sigma}|x - x'|^2) \quad (3)$$

- We can use a Taylor expansion to get an explicit feature map ϕ that contains an infinite number of terms and covers all orders of statistics.
- Minimizing MMD under this feature expansion is then equivalent to minimizing a distance between all moments of the two distributions.

References

[1] Y. Li et al., "Generative moment matching networks," in *Proc. 32nd ICML*, 2015.
[2] D. Holden et al., "Phase-functioned neural networks for character control," *ACM Trans. Graph.*, 2017.

Time-dependent Neural Net

- Voltage level distributions at different P/E cycles are different because of the program disturb effects and the wear-out of flash memory cells.
- We designed a **time-dependent generative moments matching network (TD-GMMN)**.
- The weights and biases are estimated using degree-3 polynomial functions

$$f(t) = at^3 + bt^2 + ct + d$$

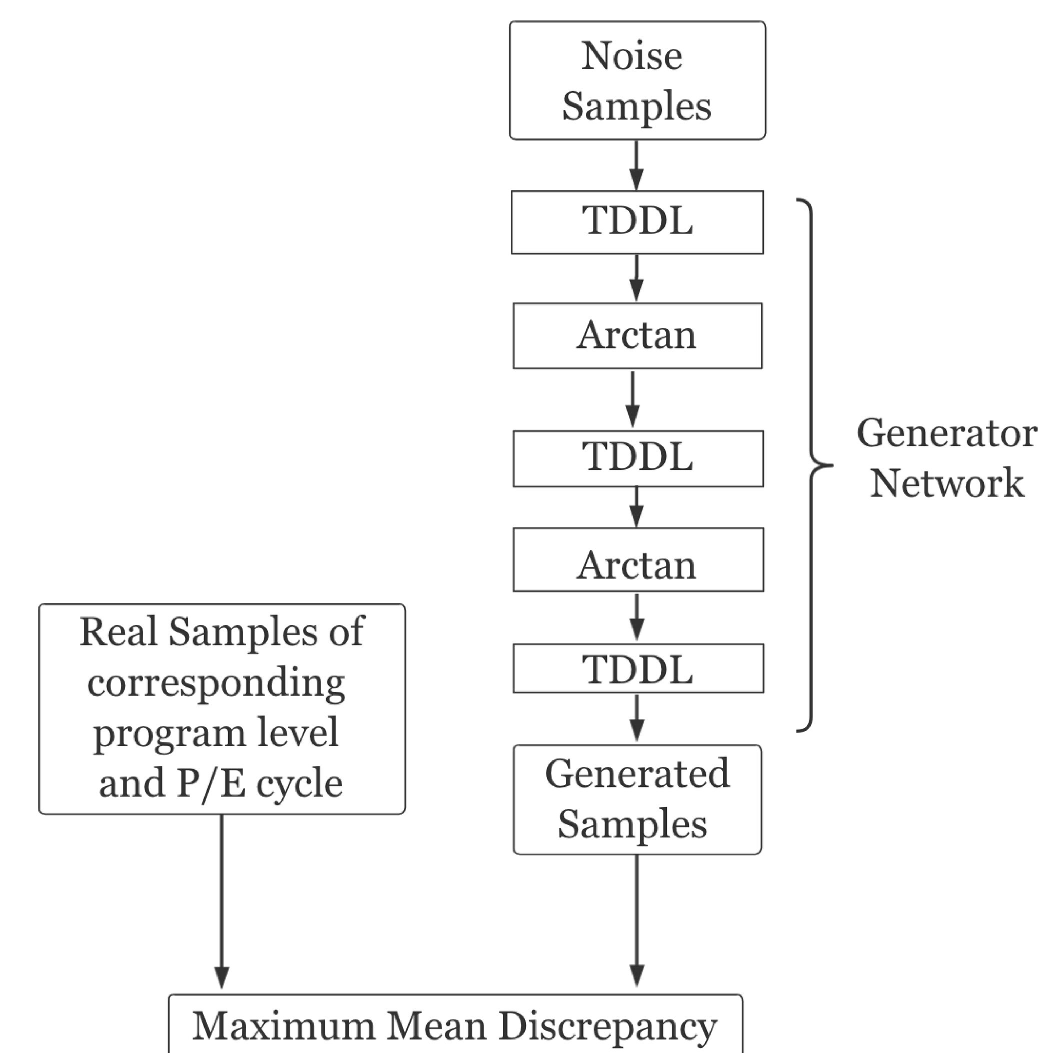
- t is the normalized P/E cycle count.

$$t = \frac{\text{present P/E cycle} - \text{minimum P/E cycle}}{\text{maximum P/E cycle} - \text{minimum P/E cycle}}$$

- a, b, c, d are parameters to be learned in the training

Time-dependent GMMN

- The workflow for training the TD-GMMN voltage level distribution generator.



References (cont.)

[3] Y. Liu et al., "Bad page detector for NAND flash memory," in *11th Annual NVMW*, 2020.

Experimental Results

- The histogram of samples drawn from the Normal-Laplace model (top), generated voltage levels (middle), and the estimated probability density function for both (bottom) for MLC program level 1 at 10, 400, and 800 P/E cycles. Number of samples: 5,000.