

멀티 레벨 낸드 플래시 메모리용 연판정 복호를 수행하는 이진 ECC 설계를 위한 EM 알고리즘

김성래*, 신동준^o

EM Algorithm for Designing Soft-Decision Binary Error Correction Codes of MLC NAND Flash Memory

Sung-Rae Kim*, Dong-Joon Shin^o

요 약

멀티 레벨 낸드 플래시 메모리는 한 셀에 2 비트 이상의 정보를 저장하는 구조이고, 비트 위치별 채널 LLR의 밀도 함수 l -밀도가 비대칭 특성을 가지고 있다. 이런 특성은 이진 무기억 대칭 채널 조건에서 설계된 오류 정정 부호의 성능이 제대로 발휘되지 못하게 할 뿐만 아니라, 멀티 레벨 낸드 플래시 메모리용 연판정 복호를 수행하는 이진 오류 정정 부호의 설계도 어렵게 한다. 본 논문에서 밀도 미러링과 EM 알고리즘을 이용하여 오류 정정 부호 설계를 위한 차선책을 소개한다. 밀도 미러링은 EM 알고리즘을 적용하기 전에 0 부호어를 전송한 경우로 가정할 수 있도록 하기 위해서 채널 LLR을 처리하는 과정이고, 이후 채널 LLR l -밀도를 EM 알고리즘을 적용하여 K 개의 성분으로 이루어진 대칭 가우시안 혼합 밀도로 근사화하는 방법을 소개한다.

Key Words : Asymmetric channel, density mirroring, EM algorithm, flash memory, l -density

ABSTRACT

In this paper, we present two signal processing techniques for designing binary error correction codes for Multi-Level Cell(MLC) NAND flash memory. MLC NAND flash memory saves the non-binary symbol at each cell and shows asymmetric channel LLR l -density which makes it difficult to design soft-decision binary error correction codes such as LDPC codes and Polar codes. Therefore, we apply density mirroring and EM algorithm for approximating the MLC NAND flash memory channel to the binary-input memoryless channel. The density mirroring processes channel LLRs to satisfy roughly all-zero codeword assumption, and then EM algorithm is applied to l -density after density mirroring for approximating it to mixture of symmetric Gaussian densities. These two signal processing techniques make it possible to use conventional code design algorithms, such as density evolution and EXIT chart, for MLC NAND flash memory channel.

* 본 연구는 삼성전자의 지원으로 수행되었습니다.

^o 본 연구는 미래창조과학부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음
(NIPA-2013- H0301-13-1001)

• First Author : 한양대학교 전자컴퓨터통신공학과 부호 및 통신 연구실, ksr0812@ccl.hanyang.ac.kr, 학생회원

◦ Corresponding Author : 한양대학교 융합전자공학부, djshin@hanyang.ac.kr, 종신회원

논문번호 : KICS2014-02-052, Received February 10, 2014; Reviewed February 24, 2014; Accepted March 3, 2014

I. 서 론

저장 용량을 증가시키기 위해서 낸드 플래시 메모리는 한 개의 셀(cell)을 여러 개의 문턱 전압으로 구분할 수 있도록 하여 한 개의 셀에 2 비트 이상의 정보를 저장할 수 있도록 발전하고 있다. 이런 구조를 가지는 낸드 플래시 메모리를 멀티 레벨 낸드 플래시 메모리라고 한다. 멀티 레벨 낸드 플래시 메모리는 한 셀에 $q \geq 2$ 개의 비트를 비이진 심벌로 매핑하여 $M=2^q > 2$ 레벨의 전압을 구분할 수 있도록 한다. 저장하고자 하는 심벌에 해당하는 일정량의 전하를 플로팅 게이트에 주입함으로써 셀 안의 전압을 M 레벨로 구분한다. 이로 인해 한 개의 셀에 1 비트의 정보만을 저장하는 단일 레벨 낸드 플래시 메모리에 비해 저장 용량이 증가하지만, 셀에 전압을 쓰고 읽는 과정에서 발생하는 여러 가지 노이즈의 영향 때문에 단일 레벨 낸드 플래시 메모리 보다 저장된 데이터 오류율이 증가하는 단점을 갖는다. 그래서 최근까지 멀티 레벨 낸드 플래시 메모리 채널을 모델링하고 이 채널에 적합한 오류 정정 부호를 설계하고 성능을 평가하는 연구가 많이 진행되었다. 특히, 멀티 레벨 플래시 메모리의 쓰기/읽기 과정에서 발생하는 노이즈를 전압 분포 밀도 함수를 이용하여 모델링하고, 모델링된 메모리 채널을 이용하여 오류 정정 부호와 신호 처리 기법을 멀티 레벨 낸드 플래시 메모리에 적용하는 연구 결과가 있다^{1,2}.

또한 멀티 레벨 낸드 플래시 메모리의 데이터 오류율이 증가하는 문제를 해결하기 위해서 오류 정정 부호를 적용하는 연구가 진행되어왔다^{3,4}. 본 논문에서는 연관정 부호를 수행하는 이진 오류 정정 부호를 고려한다. 예를 들어, 많은 연관정 부호를 수행하는 이진 오류 정정 부호 중에서 LDPC 부호와 Polar 부호는 최근까지 활발히 연구되고 있는 대표적인 부호라 할 수 있다^{5,6}. 특히, 두 부호는 이진 무기억 대칭 채널에서 0 부호어를 전송했다는 가정 하에서 부호를 설계할 수 있는 부호 설계 알고리즘(밀도 진화, EXIT 차트 등)을 가지고 있어서 점근적 성능까지 예측할 수 있는 장점을 가지고 있다^{7,8}. 그러나 멀티 레벨 낸드 플래시 메모리 채널은 비이진 심벌을 사용하고 채널 LLR l -밀도가 비대칭 특성을 갖기 때문에 0 부호어만을 전송하여 부호의 성능을 평가할 수 없을 뿐만 아니라, 기존의 부호 설계 알고리즘을 적용하기도 어렵다⁹. 또, 이진 무기억 대칭 채널에서 설계된 부호를 멀티 레벨 낸드 플래시 메모리에 적용하여도 제대로 된 성능을 발휘하기 어렵게 된다. 물론, 멀티 레벨 낸드

플래시 메모리 채널의 특성(비이진 심벌, 비대칭 특성)을 고려하여 부호를 설계하는 알고리즘을 개발하는 것이 가장 좋은 연구이지만, 이는 어렵기 때문에, 본 논문에서는 부호 설계 알고리즘을 적용하기 전 단계에서 밀도 미러링과 EM 알고리즘을 채널 LLR에 적용하여, 멀티 레벨 낸드 플래시 메모리 채널 환경을 이진 무기억 대칭 채널과 유사한 환경으로 변형시켜 주는 신호 처리 기법을 소개한다. 밀도 미러링은 EM 알고리즘을 적용하기 전에 0 부호어를 전송한 경우로 가정할 수 있도록 채널 LLR을 처리하는 과정이고, 이후 채널 LLR l -밀도를 EM 알고리즘에 적용하여 K 개의 성분으로 이루어진 대칭 가우시안 혼합 밀도로 근사화한다.

EM 알고리즘은 주어진 관측 데이터를 사용하여 최대 우도 추정(maximum likelihood estimation: MLE) 설계 기준을 만족하는 확률 밀도 함수의 파라미터를 E-step과 M-step을 반복적으로 수행하여 파라미터를 추정하는 방법이다. 통계학, 신호 처리, 기계 학습 분야 등에서 많이 사용하는 알고리즘으로써, 특히 관측 데이터를 K 개의 성분으로 구성된 가우시안 혼합 밀도 함수로 추정하는데 매우 효과적인 방법이다¹⁰. 무선 통신 분야에서는 EM 알고리즘을 MIMO-OFDM 디지털 통신 시스템에 적용하여 LDPC 부호를 설계하는 방법이 소개되었다¹¹. [11]에서는 MIMO-OFDM 수신단의 연관정 수신기에서 얻어진 채널 LLR l -밀도를 EM 알고리즘을 사용하여 대칭 가우시안 혼합 밀도 함수로 근사화한 후, 평균 진화 알고리즘⁹을 사용하여 LDPC 부호를 설계하였다. 여기서, 대칭 가우시안 밀도 함수는 분산 σ^2 과 평균 m 이 $\sigma^2 = 2m$ 를 만족하는 가우시안 밀도 함수를 의미하고 $N(m, \sigma^2 = 2m)$ 로 표시한다.

본 논문에서도 이와 유사하게 멀티 레벨 낸드 플래시 메모리 채널 LLR에 대해 EM 알고리즘을 적용하여 K 개의 성분으로 구성된 대칭 가우시안 혼합 밀도 함수로 근사화하는 방법을 소개한다. MIMO-OFDM 시스템에 적용하는 방식과 달리, 본 논문에서는 EM 알고리즘을 적용하기 전에 채널 LLR에 밀도 미러링을 수행한다. 밀도 미러링과 EM 알고리즘을 사용하여 얻어진 대칭 가우시안 혼합 밀도 함수는 0 부호어를 전송하였을 경우 얻어지는 채널 LLR l -밀도로 이해될 수 있고, 구해진 채널 LLR l -밀도를 부호 설계 알고리즘에 적용하여 부호를 효과적으로 설계할 수 있다. 위의 두 가지 과정은 Monte-Carlo 모의 실험을 통해서 수행한다.

본 논문의 구성은 다음과 같다. II에서는 멀티 레벨

낸드 플래시 메모리 채널 모델을 소개한다. 셀에 저장하고자 하는 심벌에 해당하는 전압을 쓰고 읽는 과정에서 발생하는 여러 가지 노이즈를 전압 분포 밀도 함수로 모델링한다. 그리고 III에서는 비이진 심벌, 비대칭 특성을 지닌 멀티 레벨 낸드 플래시 메모리 채널을 밀도 미러링과 EM 알고리즘을 적용하여 대칭 가우시안 혼합 밀도 함수로 근사화하는 방법을 소개한다. IV에서는 두 가지 과정을 사용하여 채널 LLR l -밀도를 대칭 가우시안 혼합 밀도 함수로 근사화하는 방법에 대한 모의 실험을 수행하고 마지막으로 결론을 맺는다.

II. 멀티 레벨 낸드 플래시 메모리 채널

멀티 레벨 낸드 플래시 메모리는 다양한 노이즈의 영향을 받는다. [1, 2]에서 제안된 멀티 레벨 낸드 플래시 메모리 채널 모델은 그림 1과 같이 블록도로 요약하여 나타낼 수 있다. 그림 1에서 나타내어진 멀티 레벨 낸드 플래시 메모리 채널에서는 셀간 간섭 (Cell-to-Cell Interference: CCI), Random Telegraph Noise(RTN), 전하 누수, P/E(프로그래밍/제거) 횟수 등, 여러 가지 요인에 의해서 노이즈가 발생하고 그 효과들은 다음과 같이 요약할 수 있다

2.1 셀 프로그래밍

멀티 레벨 낸드 플래시 메모리는 한 개의 셀에 $q \geq 2$ 개의 비트를 비이진 심벌로 매핑하여 $M=2^q$ 레벨의 전압을 구분할 수 있도록 저장하고자 하는 심벌에 해당하는 일정량의 전하를 플로우팅 게이트에 주입함으로써 셀 안의 전압을 M 레벨로 구분할 수 있도록 한다. 이를 셀 프로그래밍이라 하고, 셀 프로그래밍을 위해서는 반드시 셀 안의 전하를 제거하여 전압을 가장 낮은 전압 레벨로 낮추어 주어야 한다. 이 동

작을 제거라 한다. 예를 들어, 4-레벨 낸드 플래시 메모리는 한 셀에 저장되는 비이진 심벌을 셀 프로그래밍 상태인 E, P_1, P_2, P_3 로 표기하고 구분한다. 셀 프로그래밍 상태에 해당하는 심벌은 2 개의 비트를 그레이 매핑하여 $\{11 \rightarrow E\}, \{01 \rightarrow P_1\}, \{00 \rightarrow P_2\}, \{10 \rightarrow P_3\}$ 로 상태를 결정한다^[1,2].

M 레벨을 구분하는 전압은 여러 가지 변화 요인에 의해서 전압 분포 밀도 함수로 나타낼 수 있다. E 상태의 전압 분포 밀도 함수 $p_E(x)$ 는 가우시안 분포 함수를 따르고 다음과 같이 나타낼 수 있다.

$$p_E(x) = \frac{1}{\sqrt{2\pi}\sigma_E} e^{-\frac{(x-\mu_E)^2}{2\sigma_E^2}} \quad (1)$$

여기서, x 는 전압을 나타내고 μ_E 와 σ_E^2 는 E 상태의 평균과 분산을 나타낸다. E 상태에서 P_m 으로 $m=1, \dots, 2^q-1$, 프로그래밍하기 위해서는 단위 증가 전압 ΔV_{pp} 만큼 셀 전압을 단계적으로 증가시키면서 미리 정해진 프로그래밍 상태 P_m 의 검증 전압 V_{P_m} 에 $m=1, \dots, 2^q-1$, 도달했는지를 검증하는 방법 (program-and-verify)을 사용한다. 이를 단위 증가 펄스 프로그래밍(incremental-step pulse programming: ISPP)이라 하며, 이 방법을 사용하였을 때 $P_m, m=1, \dots, 2^q-1$ 전압 분포 밀도 함수는 각각 $[V_{P_m}, V_{P_m} + \Delta V_{pp}]$ 구간에서 일정한 분포를 가지게 된다. $V_l^{(m)} = V_{P_m}, V_r^{(m)} = V_{P_m} + \Delta V_{pp}$ 으로 표현하면, P_m 의 전압 분포 밀도 함수는 일정 분포이고 다음과 같다.

$$p_{P_m}(x) = \begin{cases} \frac{1}{\Delta V_{pp}}, & \text{if } V_l^{(m)} \leq x \leq V_r^{(m)} \\ 0, & \text{else} \end{cases} \quad (2)$$

$, m=1, \dots, 2^q-1$

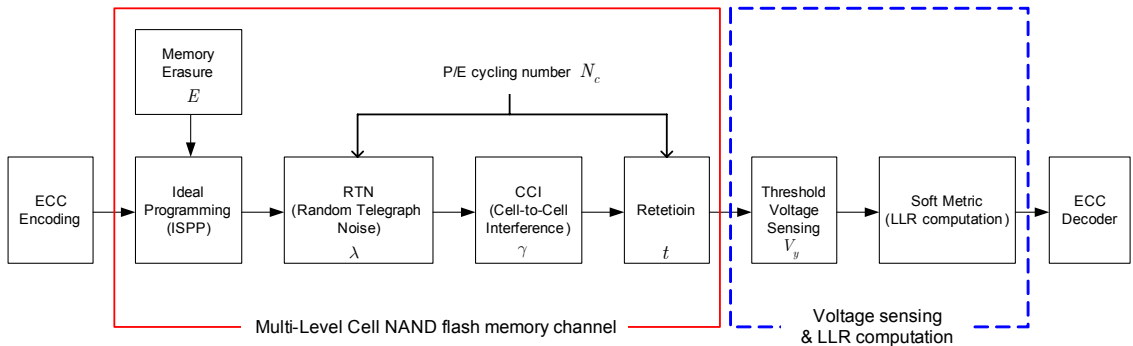


그림 1. 멀티 레벨 낸드 플래시 메모리 채널 모델
Fig. 1. Model of multi-level cell flash memory channel

2.2 P/E 횡수에 의한 RTN(Random Telegraph Noise)와 유지 오류

P/E (프로그래밍/제거) 횡수 N_c 의 증가에 따른 플로우팅 게이트의 빈번한 전하 주입과 제거는 프로그래밍 상태의 전압 분포 밀도 함수에 노이즈로 영향을 미친다. P/E 횡수 N_c 에 따른 노이즈는 크게 두 가지로 구분할 수 있다. 첫째는 셀에 저장된 전압이 변동하는 현상인 RTN(random telegraph noise)이고, 둘째는 시간 t 의 흐름과 P/E 횡수 N_c 의 증가에 의해 발생하는 전하 누수 영향으로 인하여 프로그래밍 상태의 검증 전압 V_{P_m} 가 $m = 1, \dots, 2^q - 1$, 변화하는 현상으로, 이것을 유지 오류이라고 부른다.

두 가지 노이즈는 다음과 같이 모델링 된다. RTN를 나타내는 전압 분포 밀도 함수 $p_{RTN}(x)$ 는 아래와 같은 양방향 지수 분포로 가정한다.

$$p_{RTN}(x) = \frac{1}{2\lambda_{RTN}} e^{-\frac{|x|}{\lambda_{RTN}}} \quad (3)$$

N_c 을 P/E 횡수라고 할 때, $\lambda_{RTN} = K_{RTN} N_c^{0.5}$, $K_{RTN} = 0.00025$ 이고 λ_{RTN} 는 양방향 지수 분포 함수의 파라미터이다. 유지 오류에 의한 전압 분포 밀도 함수 $p_R(x)$ 은 시간 t 와 P/E 횡수 N_c 을 고려하여 계산한 평균 μ_R 과 분산 σ_R^2 를 가지는 가우시안 분포 함수 $N(\mu_R, \sigma_R^2)$ 으로 가정한다. 즉,

$$p_R(x) = \frac{1}{\sqrt{2\pi}\sigma_R} e^{-\frac{(x-\mu_R)^2}{2\sigma_R^2}} \quad (4)$$

이고, 여기서 μ_R, σ_R^2 는 각각 P/E 횡수에 대해서 $N_c^{0.5}$, $N_c^{0.6}$ 만큼의 영향을 받고, 유지 오류 시간에 대해서 $\ln(1+t/t_0)$ 만큼의 영향을 받는다. 또한 셀 프로그래밍 E 상태의 초기 문턱 전압 μ_E 의 영향을 받기 때문에 이를 고려하기 위해서 $K_R(x - \mu_E)$ 를 사용하여 각각 μ_R, σ_R^2 를 아래와 같이 계산한다^[1,2].

$$\begin{cases} \mu_R = K_R(x - \mu_E) K_m N_c^{0.5} \ln(1+t/t_0) \\ \sigma_R^2 = K_R(x - \mu_E) K_v N_c^{0.6} \ln(1+t/t_0) \end{cases} \quad (5)$$

여기서 $K_R = 0.38$, $K_m = 4 \times 10^{-4}$, $K_v = 4 \times 10^{-6}$, $\mu_E = 1.4$ 로 설정하고, t_0 는 초기 시간을 의미하며 단위 시간 $t_0 = 1$ 으로 설정한다.

2.3 셀간 간섭(Cell-to-Cell Interference : CCI)

낸드 플래시 메모리는 셀에 프로그래밍된 전압이 인접 셀의 프로그래밍된 전압의 영향을 받아서 변하게 되는 간섭 현상이 발생한다. 이를 셀간 간섭(CCI)이라 한다. 셀간 간섭은 셀 사이에 발생하는 패러시틱 정전 용량(parasitic capacitance)으로 인한 커플링(coupling)이 중요한 요인이고, 낸드 플래시 메모리 채널에서 발생하는 노이즈 중에서 영향이 커서 커플링 효과를 제거하기 위한 연구도 진행되었다^[12,13].

셀간 간섭은 낸드 플래시 메모리의 비트라인 구조에 의해서 영향을 받게 되고, 셀간 간섭을 받는 셀의 전압 변화 $F(x)$ 는 다음과 같이 나타낼 수 있다.

$$F(x) = \sum_n (\Delta V_t^{(n)} \times \gamma^{(n)}) \quad (6)$$

여기서 $\Delta V_t^{(n)}$ 은 셀간 간섭을 유발시키는 n -번째 인접 셀이 프로그래밍된 뒤에 간섭을 받는 셀의 변화된 전압을 의미하고, $\gamma^{(n)}$ 은 셀간 간섭을 유발시키는 n -번째 인접 셀과의 커플링 비율이며 패러시틱 정전 용량을 사용하여 다음과 같이 정의된다.

$$\gamma^{(n)} = \frac{C^{(n)}}{C_{total}} \quad (7)$$

여기서 $C^{(n)}$ 은 셀간 간섭을 유발시키는 n -번째 인접 셀과의 패러시틱 정전 용량을 나타내고, C_{total} 은 모든 패러시틱 정전 용량의 합이다.

비트라인 구조는 even/odd 구조와 all-bit-line 구조를 많이 사용하며, 최근에는 all-bit-line 구조를 더 많이 사용하는 것으로 알려져 있다. 이는 all-bit-line 구

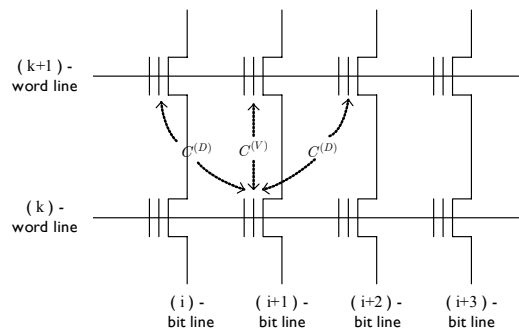


그림 2. All-bit-line 구조를 가진 낸드 플래시 메모리의 셀간 간섭 모델
Fig. 2. Cell-to-Cell interference model of the all-bit-line structure

조에서 발생하는 셀간 간섭이 even/odd 구조를 사용하는 경우보다 작기 때문이다. 셀간 간섭 영향을 각각 패러시틱 정전 용량 $C^{(V)}$, $C^{(D)}$ 로 나타냈으며 all-bit-line 구조에서는 셀간 간섭을 유발시키는 셀은 대각 방향 2 개 셀과 수직 방향 1 개 셀로 이루어진 총 3 개의 셀로부터 간섭을 받게 된다^[12]. $C^{(D)}$, $C^{(V)}$ 에 의해서 결정되는 커플링 계수 $\gamma^{(D)}$, $\gamma^{(V)}$ 는 각각 독립적으로 가우시안 밀도 함수의 양 끝을 절단한 식 (8)과 같은 확률 변수 $p_r(\gamma)$ 를 따른다고 가정한다.

$$p_r(\gamma) = \begin{cases} \frac{c_\gamma}{\sqrt{2\pi}\sigma_\gamma} e^{-\frac{(\gamma-\mu_\gamma)^2}{2\sigma_\gamma^2}}, & \text{if } |\gamma-\mu_\gamma| \leq \omega_\gamma \\ 0, & \text{else} \end{cases} \quad (8)$$

여기서 μ_γ 와 σ_γ^2 는 평균과 분산을 나타내고 ω_γ 는 $p_r(\gamma) > 0$ 인 영역을 나타낸다. c_γ 는 $\int_{\mu_\gamma-\omega_\gamma}^{\mu_\gamma+\omega_\gamma} p_r(\gamma) d\gamma = 1$ 을 만족하기 위한 상수이다. [1]에서와 같이 $\omega_\gamma = 0.1\mu_\gamma$, $\sigma_\gamma^2 = 0.4\mu_\gamma$ 로 사용한다. 수직 방향의 커플링 계수 $\gamma^{(V)}$ 의 평균 $\mathbb{E}[\gamma^{(V)}]$ 과 대각 방향의 커플링 계수 $\gamma^{(D)}$ 의 평균 $\mathbb{E}[\gamma^{(D)}]$ 의 비율을 0.08 : 0.006로 유지하고 커플링 강도 인수(coupling strength factor) s 를 이용하여 식 (6)과 같이 커플링 효과를 조절한다. 즉, $s \geq 0$ 이 주어졌을 때, $\mathbb{E}[\gamma^{(V)}] = 0.08s$, $\mathbb{E}[\gamma^{(D)}] = 0.006s$ 으로 설정하고 커플링 계수 $\gamma^{(V)}$, $\gamma^{(D)}$ 는 각각 서로 독립적으로 $p_r(\gamma)$ 를 따르는 확률 변수이다.

2.4 4-레벨 낸드 플래시 메모리 채널

4-레벨 낸드 플래시 메모리 채널에 그레이 매핑, $\{11 \rightarrow E\}$, $\{01 \rightarrow P_1\}$, $\{00 \rightarrow P_2\}$, $\{10 \rightarrow P_3\}$ 을 사용하여 프로그래밍 상태를 결정한다. 그리고 여러 가지 노이즈는 서로 독립적인 전압 분포를 따르고, 노이즈의 영향을 받은 후의 셀 프로그래밍 상태 전압은 노이즈의 영향을 받기 전 프로그래밍 상태 E, P_m 의 $m=1,2,3$, 전압에 여러 가지 노이즈를 합한 값이 된다. 즉, 노이즈의 영향을 받은 후의 셀 프로그래밍 상태 $\psi(E)$, $\psi(P_m)$ 의 $m=1,2,3$, 전압 분포 밀도 함수 $p_{\psi(E)}(x)$, $p_{\psi(P_m)}(x)$ 는 $m=1,2,3$, 노이즈의 영향을 받기 전의 프로그래밍 상태 전압 분포 밀도 함수 $p_E(x)$, $p_{P_m}(x)$ 에 $m=1,2,3$, 여러 가지 노이즈의 전압 분포 밀도 함수 $p_{RTN}(x)$, $p_R(x)$, $dF(x)/dx$ 를 콘볼루션하여 얻을 수 있다. 또는 Monte-Carlo 모의 실험을 통해서 셀 프로그래밍 상태의 전압 분포 밀도 함수 $p_{\psi(E)}(x)$, $p_{\psi(P_m)}(x)$ 를 $m=1,2,3$, 얻을 수 있다.

그림 3은 4-레벨 낸드 플래시 메모리 채널의 예를 보여주고 있다. E 상태의 $\mu_E = 1.4$, $\sigma_E^2 = 0.35$ 으로 설정하고 P_m 를 $m=1,2,3$, 프로그래밍하기 위하여 검정 전압을 각각 $V_{P_1} = 2.6$, $V_{P_2} = 3.2$, $V_{P_3} = 3.39$ 으로 설정하였다. 또, 단위 증가 전압 $\Delta V_{pp} = 0.2$ 로 설정하였다. RTN 노이즈 효과를 위해서 P/E 횟수 N_c 와 관련된 파라미터 $\lambda_{RTN} = K_{RTN} N_c^{0.5}$, $K_{RTN} = 0.00025$ 으로 설정하고, 유지 오류 노이즈 효과를 위해서 $N(\mu_R, \sigma_R^2)$ 의 평균과 분산을 식 (5)와 같이 설정하였다. 셀간 간섭 효과를 위해서 커플링 계수의 평균을 $\mathbb{E}[\gamma^{(V)}] = 0.08s$, $\mathbb{E}[\gamma^{(D)}] = 0.006s$ 으로 설정하였다. Monte-Carlo 모의 실험 샘플의 수를 각각의 프로그래밍 상태 E, P_m 에 $m=1,2,3$, $N=250000$ 개씩 사용하여 그림 3과 같은 멀티 레벨 낸드 플래시 메모리 채널을 얻었다. P/E 횟수 $N_c = 10000$, 유지 오류 시간 $t = 10 \times 24 \times 365$, 커플링 강도 인수 $s = 0.7$ 로 각각 설정하였다.

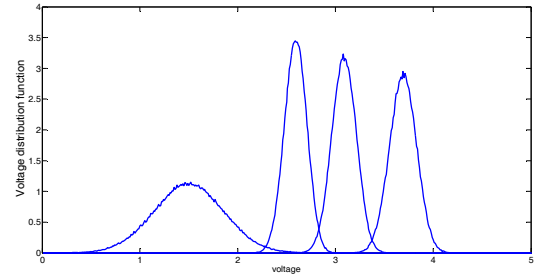


그림 3. 4-레벨 낸드 플래시 메모리 채널의 예
Fig. 3. 4-level cell NAND flash memory channel

III. 밀도 미러링과 EM 알고리즘을 적용한 채널 l-밀도 추정

3.1 밀도 미러링

멀티 레벨 낸드 플래시 메모리에 연관정 복호를 수행하는 이진 오류 정정 부호를 사용하기 위해서는 셀에서 읽어 들인 셀 전압이 V_y 일 때, 심볼을 구성하는 비트 위치 b_i 에 $i=1, \dots, q$, 다른 채널 LLR $LLR(b_i|V_y)$ 을 $i=1, \dots, q$, 계산해야 한다. $LLR(b_i|V_y)$ 을 계산하기 위해서 $\psi(E)$, $\psi(P_m)$ 의 $m=1, \dots, 2^q - 1$, 전압 분포 밀도 함수 $p_{\psi(E)}(x)$, $p_{\psi(P_m)}(x)$ 을 $m=1, \dots, 2^q - 1$, Monte-Carlo 모의 실험을 통해서 얻는다. i -번째 비트 위치 b_i 가 0인 상태 집합을 S_i , 1인 상태 집합을 S_i^c 이라 정의하고 셀에서 읽어 들인 셀 전압이 V_y 일 때, $LLR(b_i|V_y)$ 는 다음과 같이 구할 수 있다.

$$LLR(b_i|V_y) = \ln \frac{\sum_{S_i^c} p_{S_i}(V_y)}{\sum_{S_i^e} p_{S_i}(V_y)}, \quad i = 1, \dots, q \quad (9)$$

여기서 $p_{S_i}(x)$ 는 셀 프로그래밍 상태 $\psi(E)$, $\psi(P_m)$ 중에서 $m=1, \dots, 2^q - 1$, S_i 에 속한 프로그래밍 상태의 전압 분포 밀도 함수를 나타내고 $p_{S_i^c}(x)$ 는 S_i^c 에 속한 프로그래밍 상태의 전압 분포 밀도 함수를 나타낸다. 실제로 셀 전압 V_y 을 읽어 들일 때에는 셀 전압 x 를 J 개의 구간 $(R_i^{(1)}, R_i^{(1)}]$, $(R_i^{(2)}, R_i^{(2)}]$, ..., $(R_i^{(J)}, R_i^{(J)}]$ 으로 나누어 전압을 읽어 들이게 된다. 따라서 셀 전압 V_y 을 정확히 읽어 들이지 못하고 구간 $R_i^{(j)} \leq V_y \leq R_i^{(j)}$ 만을 구별하는 방법을 사용한다^[14]. 이 방법을 고려하면 식 (9)는 수정이 불가피 하지만, 본 논문에서는 셀 전압 V_y 을 부동 소수점으로 정확히 읽어 들일 수 있다고 가정한다.

각각의 프로그래밍 상태 E , P_m 을 $m=1, \dots, 2^q - 1$, N 개 샘플로 모의 실험 하여 멀티 레벨 낸드 플래시 메모리 채널의 전압 분포 밀도 함수 $p_{\psi(E)}(x)$, $p_{\psi(P_m)}(x)$ 을 $m=1, \dots, 2^q - 1$, 얻은 후, i -번째 비트 위치 $b_i = 0$ 조건을 만족하는 $2^{(q-1)N}$ 개의 심벌에 대해서 Monte-Carlo 모의 실험을 통해 채널 LLR $\{l_n^{(0)}\}_{n=1}^{2^{(q-1)N}}$ 을 식 (9)을 이용하여 얻고, 동일하게 $b_i = 1$ 조건을 만족하는 $2^{(q-1)N}$ 개의 심벌에 대한 채널 LLR $\{l_n^{(1)}\}_{n=1}^{2^{(q-1)N}}$ 을 얻는다. i -번째 비트 위치 $b_i \in \{0, 1\}$ 에 따라 구해진 $\{l_n^{(0)}\}_{n=1}^{2^{(q-1)N}}$, $\{l_n^{(1)}\}_{n=1}^{2^{(q-1)N}}$ 의 l -밀도를 각각 $f^i(l|b_i = 0)$, $f^i(l|b_i = 1)$ 이라 하자. 이 과정은 모든 $i = 1, \dots, q$ 에 동일하게 적용시킬 수 있고, $f^i(l|b_i = 0)$ 와 $f^i(l|b_i = 1)$ 는 히스토그램을 이용하여 얻을 있다. 그러나 부호 설계 알고리즘을 이용하기 위해서는 채널 LLR l -밀도에 대한 신호 처리 과정인 밀도 미러링이 필요하다. 밀도 미러링을 소개하기 위해서 대칭 채널의 정의를 소개한다.

정의 1. 이진 무기억(Binary-input Memoryless) 채널 중에서 0 부호어를 전송했다는 가정 하에서 얻어진 채널 LLR l -밀도 $f(l|0)$ 가

$$f(l|0) = e^{lf}(-l|0)$$

성질을 만족할 때, $f(l|0)$ 을 대칭을 만족한다고 말하고 이 채널을 이진 무기억 대칭(Binary Memoryless

Symmetric: BMS)이라 한다^[15-17].

이진 무기억 대칭 채널의 대표적인 예는 이진 대칭 채널(binary symmetric channel: BSC), 이진 소거 채널(binary erasure channel: BEC), 이진 AWGN 채널 등이 있다. 또 대칭 채널은 $f^i(l|0) = f^i(-l|1)$ 를 만족한다. 그러나 멀티 레벨 낸드 플래시 메모리 채널에서 얻은 i -번째 비트 위치에 해당하는 채널 LLR l -밀도 $f^i(l|b_i = 0)$, $f^i(l|b_i = 1)$ 는 둘 다 대칭 조건을 만족하지 못할 뿐만 아니라, $b_i = 0$ 과 $b_i = 1$ 인 두 가지 상황을 함께 고려하여야만 부호 설계 알고리즘으로 좋은 부호를 설계할 수 있다.

밀도 미러링이란 $b_i = 0$ 과 $b_i = 1$ 인 두 가지 상황에서 서로 다른 $f^i(l|b_i = 0)$, $f^i(l|b_i = 1)$ 를 둘 다 고려하여 좋은 부호를 설계하기 위해서, $b_i = 1$ 인 조건으로 얻어진 LLR $\{l_n^{(1)}\}_{n=1}^{2^{(q-1)N}}$ 의 부호(양수, 음수)를 바꾸어 $b_i = 0$ 인 조건에서 얻어진 채널 LLR $\{l_n^{(0)}\}_{n=1}^{2^{(q-1)N}}$ 과 함께 EM 알고리즘에 적용시키기 위한 채널 LLR에 대한 사전 신호 처리이다. 즉, 두 가지 상황을 함께 고려하기 위해서 $b_i = 1$ 인 조건에서 얻어진 채널 LLR $\{l_n^{(1)}\}_{n=1}^{2^{(q-1)N}}$ 의 부호(양수, 음수)를 바꾸어주고 $b_i = 0$ 인 조건에서 얻어진 채널 LLR $\{l_n^{(0)}\}_{n=1}^{2^{(q-1)N}}$ 과 함께 즉, $\{l_n^{(1)}\}_{n=1}^{2^{(q-1)N}} = \left\{ \{l_n^{(0)}\}_{n=1}^{2^{(q-1)N}}, -\{l_n^{(1)}\}_{n=1}^{2^{(q-1)N}} \right\}$ 를 EM 알고리즘에 사용한다.

그림 4는 그림 3과 같은 조건의 4-레벨 낸드 플래시 메모리 채널 상황에서 첫 번째 비트 위치 b_1 에 대한 $f^1(l|b_1 = 0)$ (그림 4-1), $f^1(l|b_1 = 1)$ (그림 4-2)와 밀도 미러링의 개념을 보여주고 있다. 각각의 프로그래밍 상태 E , P_m 에 $m=1, 2, 3$, $N=250000$ 개의 샘플을 사용하여 모의 실험을 수행하여 채널 LLR $\{l_n^{(0)}\}_{n=1}^{2^N}$, $\{l_n^{(1)}\}_{n=1}^{2^N}$ 를 얻었다. 그리고 밀도 미러링을 수행하여 EM 알고리즘에 적용하기 위한 채널 LLR $\{l_n^{(1)}\}_{n=1}^{2^N} = \left\{ \{l_n^{(0)}\}_{n=1}^{2^N}, -\{l_n^{(1)}\}_{n=1}^{2^N} \right\}$ 를 사용하여 $f^1(l) = f^1(l|b_1 = 0)/2 + f^1(-l|b_1 = 1)/2$ (그림 4-4)를 얻었다. 또한 최대/최소 LLR 값을 각각 40/-40으로 정하였다. 밀도 미러링은 [18]에서 설명된 i.i.d. 채널 어댑터와 동일한 기능을 한다.

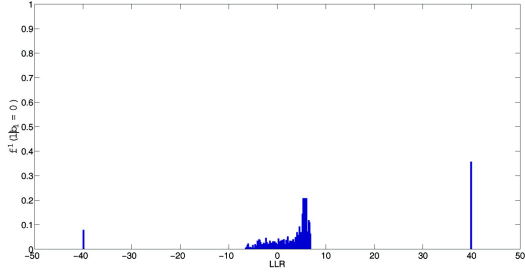


그림 4-1. 그림 3에서 주어진 4-레벨 낸드 플래시 메모리 채널의 $f^l(lb_1=0)$
 Fig. 4-1. $f^l(lb_1=0)$ from 4-level cell NAND flash memory channel on Fig. 3

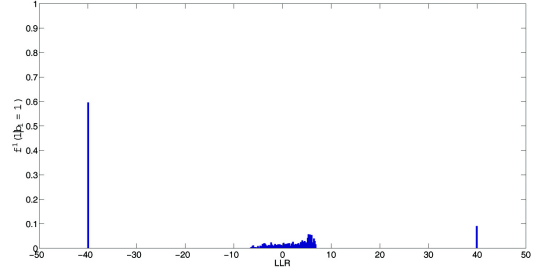


그림 4-2. 그림 3에서 주어진 4-레벨 낸드 플래시 메모리 채널의 $f^l(lb_1=1)$
 Fig. 4-2. $f^l(lb_1=1)$ from 4-level cell NAND flash memory channel on Fig. 3

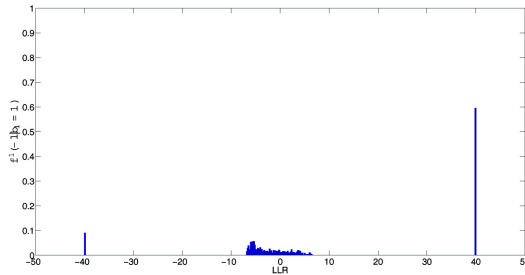


그림 4-3. $f^l(lb_1=1)$ 를 밀도 미러링한 후 $f^l(-lb_1=1)$
 Fig. 4-3. $f^l(-lb_1=1)$ after applying density mirroring to $f^l(lb_1=1)$

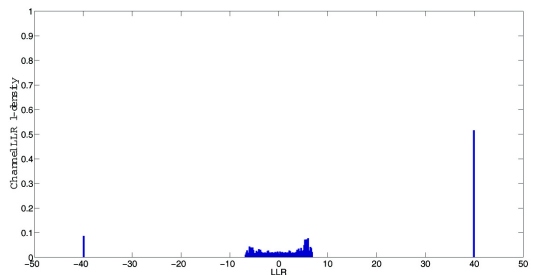


그림 4-4. 채널 LLR l -밀도 히스토그램
 $f^l(l) = f^l(lb_1=0)/2 + f^l(-lb_1=1)/2$
 Fig. 4-4. Histogram of channel LLR l -density
 $f^l(l) = f^l(lb_1=0)/2 + f^l(-lb_1=1)/2$

그림 4. 4-레벨 낸드 플래시 메모리 채널에서 밀도 미러링을 수행 한 후 채널 LLR l -밀도
 Fig. 4. Channel LLR l -density from 4-level cell NAND flash memory on Fig. 3

3.2 EM 알고리즘을 이용한 채널 l -밀도 추정

본 장에서는 밀도 미러링으로 얻어진 $\{l_n^i\}_{n=1}^{2^N}$ 을 이용하여 채널 LLR l -밀도를 K 개의 성분을 가진 대칭 가우시안 혼합 밀도 함수로 근사화하는 EM 알고리즘을 소개한다. 그림 3과 같이 주어진 멀티 레벨 낸드 플래시 메모리 채널에서 얻은 채널 LLR $\{l_n^i\}_{n=1}^{2^N}$ 에서 $l=40$ 과 $l=-40$ 인 경우는 EM 알고리즘 적용에서 제외한다. 왜냐하면 $l=40$ 과 $l=-40$ 은 경우는 델타 함수 $\delta(l-40)$, $\delta(l+40)$ 를 이용하여 l -밀도를 표현할 수 있기 때문이다. 즉, $\sum_{k=1}^K \pi_k N(m_k, 2m_k) + \alpha \delta(l+40) + \beta \delta(l-40)$ 형태로 표현할 수 있다. 또한 0 부호어를 전송했을 때의 m_k 는 양의 값을 가진다. 수식 표현의 편의와 이해를 돕기 위하여, 비트 위치 i 를 생략하여 $f(l)$, $\{l_n^i\}_{n=1}^{2^N}$ 으로 나타내고, $N(m, 2m)$ 는 EM 알고리즘의 적용에 편리하도록 대칭 가우시안 밀도 함수의 파라미터인 평균 m 을 이용하여 $N(l; m, 2m)$ 으로도 나타내어 혼용

하여 사용한다. l -밀도 $f(l)$ 를 EM 알고리즘을 적용하여 식 (10)과 같이 근사화한다.

$$f(l) \cong p_L(l; \theta) = p_L(l; \pi, \mathbf{m}) = \sum_{k=1}^K \pi_k N(l; m_k, 2m_k) \quad (10)$$

여기서 π_k 는 k -번째 성분 $N(m_k, 2m_k)$ 이 발생하는 확률을 의미하는 혼합 계수이고 $\pi = \left\{ \pi_1, \dots, \pi_K \mid 0 \leq \pi_k \leq 1, \sum_{k=1}^K \pi_k = 1 \right\}$ 는 K 개 성분이 발생할 사전 확률(a priori)을 의미하는 혼합 계수 파라미터이다.

EM 알고리즘은 통계학이나 신호 처리 분야에서 미지의 파라미터 θ 의 최대 우도 추정 $\hat{\theta}^*$ 를 E-step과 M-step을 반복적으로 수행하여 얻는 방법이다. 본 장에서는 관측 데이터 $\{l_n^i\}_{n=1}^{2^N}$ 에서 $l=40$ 과 $l=-40$ 인 경우를 제외한 나머지 채널 LLR $\mathbf{X} = \{l_n\}_{n=1}^N$ 가 주어

졌을 때, 식 (10)과 같이 $\mathbf{X} = \{l_n\}_{n=1}^{N_c}$ 가 따르는 확률 밀도 함수 $p_L(l; \theta)$ 의 미지의 파라미터 $\theta = \{\pi = \{\pi_k\}_{k=1}^K, \mathbf{m} = \{m_k\}_{k=1}^K\}$ 를 추정한다. 즉, 주어진 관측 데이터 $\mathbf{X} = \{l_n\}_{n=1}^{N_c}$ 를 이용하여 아래 식 (11)과 같이 우도 함수 $p_L(\mathbf{X}; \theta) = \prod_{n=1}^{N_c} p_L(l_n; \theta)$ 혹은 로그 우도 함수 $\ln p_L(\mathbf{X}; \theta)$ 를 최대로 만드는 파라미터 $\theta^* = \{\pi^*, \mathbf{m}^*\}$ 를 EM 알고리즘을 적용하여 추정한다^[10].

$$\begin{aligned} \theta^* &= \operatorname{argmax}_{\theta} \ln p_L(\mathbf{X}; \pi, \mathbf{m}) \\ &= \operatorname{argmax}_{\theta} \sum_{n=1}^{N_c} \ln \left\{ \sum_{k=1}^K \pi_k N(l_n; m_k, 2m_k) \right\} \end{aligned} \quad (11)$$

EM 알고리즘에서는 파라미터 $\theta = \{\pi, \mathbf{m}\}$ 를 추정하기 위해서, 주어진 $\mathbf{X} = \{l_n\}_{n=1}^{N_c}$ 에 은닉 변수를 정의한다. 은닉 변수 $\mathbf{z}_n = (z_{n1}, \dots, z_{nK})^T$ 는 n -번째 LLR l_n 이 k -번째 성분 $N(m_k, 2m_k)$ 에서 발생한 경우에 \mathbf{z}_n 의 k -번째 성분은 $z_{nk} = 1$ 이고, 나머지 성분은 $\{z_{nj} = 0, j = 1, \dots, K, j \neq k\}$ 인 K 차원 이진 표시 벡터로 정의된다. 즉, \mathbf{z}_n 은 l_n 이 K 개 성분 $N(m_k, 2m_k)$ 중에서 $k = 1, \dots, K$, 어떤 성분에서 발생된 것인지를 명확히 설명하는 변수이다. 그러나 일반적으로 \mathbf{z}_n 은 주어지지 않기 때문에 \mathbf{z}_n 을 은닉 변수라 부른다.

$\mathbf{X} = \{l_n\}_{n=1}^{N_c}$ 뿐만 아니라 은닉 변수 $\mathbf{Z} = \{\mathbf{z}_n\}_{n=1}^{N_c}$ 도 함께 주어지는 경우에는 \mathbf{z}_n 을 이용하여 l_n 이 발생한 성분을 알 수 있기 때문에 θ^* 를 쉽게 구할 수 있다. 반면, $\mathbf{X} = \{l_n\}_{n=1}^{N_c}$ 만 주어졌을 경우에는 은닉 변수 $\mathbf{Z} = \{\mathbf{z}_n\}_{n=1}^{N_c}$ 를 알지 못하기 때문에 \mathbf{z}_n 의 기대치 $\gamma(\mathbf{z}_n) = (\gamma(z_{n1}), \gamma(z_{n2}), \dots, \gamma(z_{nK}))^T$ 를 먼저 구해야한다. $\gamma(z_{nk})$ 는 l_n 이 주어졌을 때, l_n 이 k -번째 성분 $N(m_k, 2m_k)$ 에서 발생했을 확률 $P(z_{nk} = 1 | l_n; \pi, \mathbf{m})$ 을 의미하고, 사후 확률(a posteriori)로 해석할 수 있다. 그리고 $\gamma(\mathbf{z}_n)$ 를 구하기 위해 필요한 k -번째 성분 $N(m_k, 2m_k)$ 이 발생할 사전 확률 $P(z_{nk} = 1; \pi, \mathbf{m})$ 는 파라미터 $\mathbf{m} = \{m_k\}_{k=1}^K$ 에 독립이고, k -번째 성분 $N(m_k, 2m_k)$ 이 발생하는 확률은 π_k 으로 가정하였으므로 $P(z_{nk} = 1; \pi, \mathbf{m}) = P(z_{nk} = 1; \pi) = \pi_k$ 이다. 이와 같은 관계를 고려하여 EM 알고리즘을 적용하면 최적 파라미터 $\theta^* = \{\pi^*, \mathbf{m}^*\}$ 는 아래와 같이 얻을 수 있다^{[10][11]}.

EM 알고리즘 반복 횟수 t 에서 얻어진 파라미터 $\theta^{(t)} = \{\pi^{(t)}, \mathbf{m}^{(t)}\}$ 를 이용하여 $t+1$ 에서의 파라미터

$\theta^{(t+1)} = \{\pi^{(t+1)}, \mathbf{m}^{(t+1)}\}$ 를 구하기 위해서 E-step에서는 $\{\gamma^{(t+1)}(\mathbf{z}_n)\}_{n=1}^{N_c}$ 를 아래와 같이 구한다.

$$\begin{aligned} \gamma^{(t+1)}(z_{nk}) &\doteq \frac{P(z_{nk} = 1 | l_n; \pi^{(t)}, \mathbf{m}^{(t)})}{\sum_{j=1}^K P(z_{nj} = 1; \pi^{(t)}) p(l_n | z_{nj} = 1; \mathbf{m}^{(t)})} \\ &= \frac{\pi_k^{(t)} N(l_n; m_k^{(t)}, 2m_k^{(t)})}{\sum_{j=1}^K \pi_j^{(t)} N(l_n; m_j^{(t)}, 2m_j^{(t)})} \end{aligned} \quad (12)$$

여기서 $\gamma^{(t+1)}(z_{nk})$ 는 l_n 이 주어졌을 때, l_n 이 k -번째 $N(m_k, 2m_k)$ 성분에서 발생했을 사후 확률이고 $\pi_k = P(z_{nk} = 1; \pi)$ 는 k -번째 $N(m_k, 2m_k)$ 성분이 발생할 사전 확률이다. 그리고 $p(l_n | z_{nk} = 1; \pi, \mathbf{m})$ 는 $z_{nk} = 1$ 이 주어졌을 때, l_n 이 가지는 조건부 확률 밀도 함수를 의미한다. $p(l_n | z_{nk} = 1; \pi, \mathbf{m})$ 는 파라미터 $\pi = \{\pi_k\}_{k=1}^K$ 에 독립이고, $z_{nk} = 1$ 는 k -번째 $N(m_k, 2m_k)$ 성분이 발생했다는 조건을 의미하므로 $p(l_n | z_{nk} = 1; \pi, \mathbf{m}) = p(l_n | z_{nk} = 1; \mathbf{m}) = N(l_n; m_k, 2m_k)$ 이 성립한다. 즉, k -번째 $N(m_k, 2m_k)$ 성분이 주어졌을 때 l_n 이 발생할 확률 밀도 함수를 의미한다.

E-step에서 얻어진 $\{\gamma^{(t+1)}(\mathbf{z}_n)\}_{n=1}^{N_c}$ 를 이용하여 M-step에서는 반복 횟수 $t+1$ 에서의 파라미터 $\theta^{(t+1)} = \{\pi^{(t+1)}, \mathbf{m}^{(t+1)}\}$ 를 구한다. EM 알고리즘의 성질에 의하여 $\theta^{(t+1)}$ 은 $\theta^{(t)}$ 보다 더욱 θ^* 에 가까운 추정이 되고, E-step과 M-step을 반복적으로 수행하여 θ^* 를 구할 수 있다. M-step에서는 식 (13)과 같이 정의된 함수 $Q(\theta^{(t+1)})$ 를 최대로 만드는 파라미터 $\theta^{(t+1)} = \{\pi^{(t+1)}, \mathbf{m}^{(t+1)}\}$ 를 편미분을 이용하여 구한다^[11].

$$\begin{aligned} Q(\theta^{(t+1)}) &= \sum_{n=1}^{N_c} \sum_{k=1}^K \gamma^{(t+1)}(z_{nk}) \\ &\quad \times \left\{ \ln \pi_k^{(t+1)} + \ln N(l_n; m_k^{(t+1)}, 2m_k^{(t+1)}) \right\} + C \\ &= \sum_{n=1}^{N_c} \sum_{k=1}^K \gamma^{(t+1)}(z_{nk}) \\ &\quad \times \left\{ \ln \pi_k^{(t+1)} - \frac{1}{2} \ln 2m_k^{(t+1)} - \frac{(l_n - m_k^{(t+1)})^2}{4m_k^{(t+1)}} \right\} + C \end{aligned} \quad (13)$$

먼저, $\pi^{(t+1)} = \{\pi_k^{(t+1)}\}_{k=1}^K$ 을 구하기 위해서 $Q(\theta^{(t+1)})$ 를 편미분한다.

알고리즘 1. 비대칭 채널 l -밀도를 대칭 가우시안 혼합 밀도 함수로 근사화하는 EM 알고리즘

Algorithm 1. EM algorithm for approximating asymmetric channel l -density to mixture of symmetric Gaussian densities

$\mathbf{X} = \{l_n\}_{n=1}^{N_e}$ 이 주어졌을 때, 식 (10)과 같이 $p_L(l; \boldsymbol{\pi}, \mathbf{m})$ 로 근사화하기 위해서는 $\mathbf{X} = \{l_n\}_{n=1}^{N_e}$ 의 은닉 변수 $\mathbf{Z} = \{z_n\}_{n=1}^{N_e}$ 의 기대치 $\{\gamma(z_n)\}_{n=1}^{N_e}$ 와 파라미터 $\boldsymbol{\theta} = \{\boldsymbol{\pi} = \{\pi_k\}_{k=1}^K, \mathbf{m} = \{m_k\}_{k=1}^K\}$ 를 반복 횟수 $t = 0, \dots, T$ 에 따라 아래와 같이 구한다.

① $t = 0$ 으로 설정하고, $t = 0$ 에서의 파라미터 $\boldsymbol{\theta}^{(0)}$ 값을 초기화 한다. 즉, $\{\boldsymbol{\pi}^{(0)}, \mathbf{m}^{(0)}\}$ 초기 값을 정한다. 일반적으로 K -means 알고리즘^[10]을 많이 사용한다.

② **E-step** : t 에서의 파라미터 $\boldsymbol{\theta}^{(t)} = \{\boldsymbol{\pi}^{(t)}, \mathbf{m}^{(t)}\}$ 를 이용하여, 관측 데이터 $\mathbf{X} = \{l_n\}_{n=1}^{N_e}$ 의 은닉 변수 기대치 $\{\gamma(z_n)\}_{n=1}^{N_e}$ 를 아래와 같이 구한다.

$$\gamma^{(t+1)}(z_{nk}) \doteq P(z_{nk} = 1 | l_n; \boldsymbol{\pi}^{(t)}, \mathbf{m}^{(t)}) = \frac{\pi_k^{(t)} N(l_n; m_k^{(t)}, 2m_k^{(t)})}{\sum_{j=1}^K \pi_j^{(t)} N(l_n; m_j^{(t)}, 2m_j^{(t)})}, \quad n = 1, \dots, N_e, k = 1, \dots, K$$

③ **M-step** : Step-2에서 구해진 $\{\gamma(z_n)\}_{n=1}^{N_e}$ 을 사용하여 $t + 1$ 에서의 새로운 파라미터 $\boldsymbol{\theta}^{(t+1)} = \{\boldsymbol{\pi}^{(t+1)}, \mathbf{m}^{(t+1)}\}$ 를 아래와 같은 식을 이용하여 구한다.

$$\pi_k^{(t+1)} = \frac{1}{N_e} \sum_{n=1}^{N_e} \gamma^{(t+1)}(z_{nk}), \quad k = 1, \dots, K$$

$$m_k^{(t+1)} = -1 + \sqrt{1 + \frac{\sum_{n=1}^{N_e} \gamma^{(t+1)}(z_{nk}) l_n^2}{\sum_{n=1}^{N_e} \gamma^{(t+1)}(z_{nk})}}, \quad k = 1, \dots, K$$

④ 새로운 파라미터 $\boldsymbol{\theta}^{(t+1)} = \{\boldsymbol{\pi}^{(t+1)}, \mathbf{m}^{(t+1)}\}$ 를 사용하여 로그 우도

$$\ln p_L(\mathbf{X}; \boldsymbol{\pi}^{(t+1)}, \mathbf{m}^{(t+1)}) = \sum_{n=1}^{N_e} \ln \left\{ \sum_{k=1}^K \pi_k^{(t+1)} N(l_n; m_k^{(t+1)}, 2m_k^{(t+1)}) \right\}$$

를 계산하고 $\ln p_L(\mathbf{X}; \boldsymbol{\pi}^{(t+1)}, \mathbf{m}^{(t+1)})$ 혹은 파라미터 $\boldsymbol{\theta}^{(t+1)}$ 가 수렴하는지 검토한다. 만약 수렴이 되었다면, $\boldsymbol{\theta}^* = \boldsymbol{\theta}^{(t+1)}$ 이고 알고리즘을 종료한다. 이와 달리 수렴되지 않았다면, 반복 횟수를 $t + 1$ 로 증가시키고 step-2로 돌아가서 알고리즘을 수행한다.

$$\frac{\partial Q(\boldsymbol{\theta}^{(t+1)})}{\partial \pi_k^{(t+1)}} = 0 \rightarrow \pi_k^{(t+1)} = \frac{1}{N_e} \sum_{n=1}^{N_e} \gamma^{(t+1)}(z_{nk}) \quad (14)$$

또, $\mathbf{m}^{(t+1)} = \{m_k^{(t+1)}\}_{k=1}^K$ 를 구하기 위해서 $Q(\boldsymbol{\theta}^{(t+1)})$ 를 편미분한다.

$$\begin{aligned} \frac{\partial Q(\boldsymbol{\theta}^{(t+1)})}{\partial m_k^{(t+1)}} &= 0 \\ \rightarrow \{m_k^{(t+1)}\}^2 + 2\{m_k^{(t+1)}\} - \frac{\sum_{n=1}^{N_e} \gamma^{(t+1)}(z_{nk}) l_n^2}{\sum_{n=1}^{N_e} \gamma^{(t+1)}(z_{nk})} &= 0 \\ \rightarrow \{m_k^{(t+1)}\} &= 1 + \sqrt{1 + \frac{\sum_{n=1}^{N_e} \gamma^{(t+1)}(z_{nk}) l_n^2}{\sum_{n=1}^{N_e} \gamma^{(t+1)}(z_{nk})}} \quad (15) \end{aligned}$$

EM 알고리즘을 사용한 채널 LLR l -밀도를 식 (10) 과 같이 근사화하는 과정을 요약하면 알고리즘 1과 같다. 알고리즘 1을 통해 최대 우도 추정 $\boldsymbol{\theta}^* = \{\boldsymbol{\pi}^*, \mathbf{m}^*\}$ 를 얻은 후 $l = -40$ 과 $l = 40$ 인 경우를 채널 LLR l -밀도에 포함시키기 위해서 $\boldsymbol{\pi}^*$ 를 조절한다. $l = -40$ 일 때의 확률 밀도 함수를 $\alpha \delta(l + 40)$ 이라하고 $l = 40$ 일 때의 확률 밀도 함수를 $\beta \delta(l - 40)$ 라 하면 $\boldsymbol{\pi}^*$ 는 아래와 같은 식 (16)을 이용하여 조절한다.

$$(1 - \alpha - \beta) \pi_k^*, \quad k = 1, \dots, K. \quad (16)$$

IV. 모의 실험

모의 실험을 위해서 본문에 소개된 4-레벨 낸드 플래시 메모리 채널을 사용하였다. Monte-Carlo 모의

실험을 위해서 E 상태의 $\mu_E = 1.4$, $\sigma_E^2 = 0.35$ 로 설정하고, P_m 를 $m = 1, 2, 3$, 프로그래밍하기 위하여 검정 전압을 각각 $V_{P_1} = 2.6$, $V_{P_2} = 3.2$, $V_{P_3} = 3.39$ 로 설정하였다. 또, 단위 증가 전압 $\Delta V_{pp} = 0.2$ 로 설정하였다. RTN 노이즈 효과를 위해서 P/E 횟수 N_c 와 관련된 파라미터 $\lambda_{RTN} = K_{RTN} N_c^{0.5}$, $K_{RTN} = 0.00025$ 로 설정하고, 유지 오류 노이즈 효과를 위해서 $N(\mu_R, \sigma_R^2)$ 의 평균과 분산을 식 (5)와 같이 설정하였다. 셀간 간섭 효과를 위해서 커플링 계수의 평균을 $\mathbb{E}[\gamma^{(V)}] = 0.08s$, $\mathbb{E}[\gamma^{(D)}] = 0.006s$ 로 설정하였다. Monte-Carlo 모의 실험 샘플의 수를 각각의 프로그래밍 상태 E , P_m 에 $m = 1, 2, 3$, $N = 250000$ 개씩 사용하여 그림 3과 같은 멀티 레벨 낸드 플래시 메모리 채널을 얻었다. P/E 횟수 $N_c = 10000$, 유지 오류 시간 $t = 10 \times 24 \times 365$, 커플링 강도 인수 $s = 0.7$ 로 각각 설정하였다. 그리고 첫 번째 비트 위치 b_1 에 대한 조건부 l -밀도 $f^l(l|b_1 = 0)$ 과 $f^l(l|b_1 = 1)$ 를 $\{l_n^{(0)}\}_{n=1}^{2N}$, $\{l_n^{(1)}\}_{n=1}^{2N}$ 를 구하여 그림 4에서처럼 히스토그램을 이용하여 구하였다. 모든 비트 위치 b_i 에 $i = 1, 2$, 위와 같은 과정을 동일하게 적용하여 $f^l(l|b_i = 0)$ 과 $f^l(l|b_i = 1)$ 를 구할 수 있다. 그림 4에서 볼 수 있듯이 채널 LLR이 $l = 40$ 과 $l = -40$ 인 경우는 델타 함수 $\delta(l - 40)$ 과 $\delta(l + 40)$ 으로 각각 표현이 가능하기 때문에 EM 알고리즘을 적용하여 대칭 가우시안 혼합 밀도 함수로 l -밀도를 근사화하는 것은 $l = 40$ 과 $l = -40$ 인 샘플을 제외한 나머지 샘플 $\{l_n^i\}_{n=1}^{N_c}$ 만을 이용한다. 즉, 밀도 미러링 후에 얻어진 $\{l_n^i\}_{n=1}^{4N} = \left\{ \{l_n^{(0)}\}_{n=1}^{2N}, -\{l_n^{(1)}\}_{n=1}^{2N} \right\}$ 에서 $l = 40$ 과

$l = -40$ 인 샘플을 제외한 나머지 샘플 $\{l_n^i\}_{n=1}^{N_c}$ 만을 이용하여 EM 알고리즘을 수행하였다. K 개의 성분으로 이루어진 대칭 가우시안 혼합 밀도 함수를 근사화하기 위해서 $K = 1, 2, 3$ 을 설정하였고, 그림 5는 $K = 1, 2, 3$ 에 따라서 $\{l_n^i\}_{n=1}^{N_c}$ 를 이용하여 EM 알고리즘으로 근사화한 채널 LLR l -밀도를 보여주고 있다. 그림 5-1는 $\{l_n^i\}_{n=1}^{N_c}$ 에 가우시안 혼합 밀도 함수를 가정하여 EM 알고리즘을 적용한 l -밀도를 근사한 결과를 보여주고 있는 반면, 그림 5-2는 대칭 가우시안 밀도 함수로 근사한 결과를 보여주고 있다. 공교롭게도 $s = 0.7$ 인 그림 3과 같은 채널 조건에서는 대칭 특성을 만족해야 하는 제한 조건에 의해서 $K = 1, 2, 3$ 인 경우에 모두 동일한 대칭 가우시안 밀도 함수로 근사화되었다. 표 1은 $s = 0.4, 0.7, 1$ 에 따른 대칭 가우시안 밀도 함수로 근사화한 결과를 나타내었다. 표 1에 나타난 멀티 레벨 낸드 플래시 메모리 채널은 그림 3에서 설정된 채널 파라미터 중에서 s 만을 $s = 0.4, 0.7, 1$ 으로 바꿔가며 모의 실험을 수행한 결과를 나타낸 표이다. 0 부호어를 전송한 경우이므로 평균 $m_k > 0$ 이고, 이에 의해서 그림 5에서처럼 평균 m_k 과 분산 $\sigma^2 = 2m_k$ 이 매우 큰 대칭 가우시안 밀도 함수로 근사화되었다. 그리고 혼합 성분의 개수 $K = 1, 2, 3$ 에 따라서 대칭 가우시안 밀도 함수의 성질을 만족시켜야 하기 때문에 가우시안 혼합 밀도 함수인 경우에 비해서 혼합 성분의 구분이 어려운 특성도 보이고 있다. 즉, 오직 한 개의 대칭 가우시안 혼합 밀도 함수로 근사화되고 있다. 이는 채널의 대칭 조건에 의해서 발생하는 현상으로 판단된다.

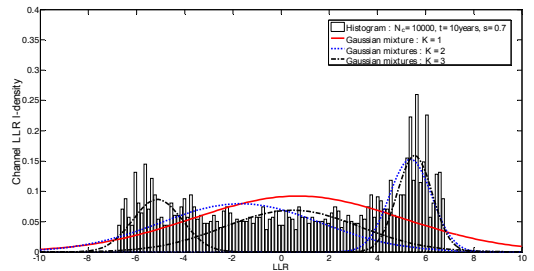


그림 5-1. $N_c = 10000$, $t = 10 \times 24 \times 365$, $s = 0.7$ 인 채널에서 가우시안 혼합 밀도 함수로 근사화한 l -밀도
Fig. 5-1. Approximated l -density by Gaussian mixture model in 4-level cell memory channel with parameters $N_c = 10000$, $t = 10 \times 24 \times 365$, $s = 0.7$

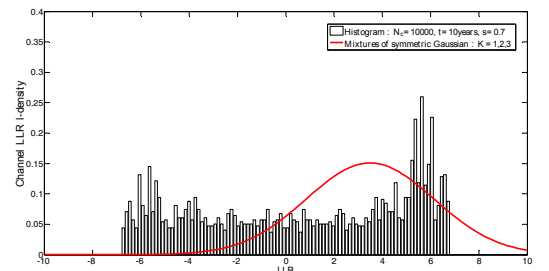


그림 5-2. $N_c = 10000$, $t = 10 \times 24 \times 365$, $s = 0.7$ 인 채널에서 대칭 가우시안 혼합 밀도 함수로 근사화한 l -밀도
Fig. 5-2. Approximated l -density by mixture of symmetric Gaussian densities in 4-level cell memory channel with parameters $N_c = 10000$, $t = 10 \times 24 \times 365$, $s = 0.7$

그림 5. EM 알고리즘을 적용한 l -밀도의 근사화
Fig. 5. Approximation of l -density using EM algorithm

그림 6은 표 1의 모의 실험 환경과 달리 $N_c = 100000$, $t = 30 \times 24 \times 365$, $s = 0.7$ 를 파라미터로 하여 표 1보다 더욱 노이즈의 영향을 많이 받는 환경을 고려한 경우이다. 이와 같은 채널 환경에서 그림 6은 밀도 미러링과 EM 알고리즘을 적용하여 얻은 근사화된 채널 LLR l -밀도를 보여 주고 있다. 또, 표 1에서 보다 노이즈의 영향이 더 많은 환경에서 오히려 $K=1,2,3$ 에 따라서 혼합 성분이 잘 구별되고 있음을 확인할 수 있다.

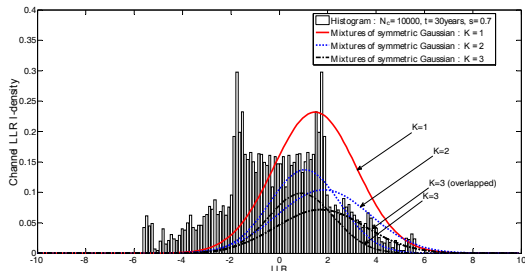


그림 6. $N_c = 100000$, $t = 30 \times 24 \times 365$, $s = 0.7$ 인 채널에서 대칭 가우시안 혼합 밀도 함수로 근사화된 채널 l -밀도
Fig. 6. Approximated l -density by mixture of symmetric Gaussian densities in 4-level cell memory channel with parameters $N_c = 100000$, $t = 30 \times 24 \times 365$, $s = 0.7$

표 1. EM 알고리즘을 적용하여 커플링 강도 인수 $s = 0.4, 0.7, 1$ 인 셀간 간섭의 영향에 따른 채널 LLR l -밀도 근사화

Table 1. Approximated l -density by mixtures of symmetric Gaussian densities using EM algorithm according to various $s = 0.4, 0.7, 1$

s	K	approximated l -density
$s = 0.4$	$K = 1, 2, 3$	$0.3583N(m_1 = 3.5796, 2m_1)$ $+ 0.1253\delta(l + 40)$ $+ 0.5164\delta(l - 40)$
$s = 0.7$	$K = 1, 2, 3$	$0.3953N(m_1 = 3.4926, 2m_1)$ $+ 0.0828\delta(l + 40)$ $+ 0.5219\delta(l - 40)$
$s = 1.0$	$K = 1, 2, 3$	$0.4285N(m_1 = 3.51, 2m_1)$ $+ 0.0994\delta(l + 40)$ $+ 0.4771\delta(l - 40)$

V. 결 론

본 논문에서는 멀티 레벨 낸드 플래시 메모리 채널을 소개하고 밀도 미러링과 EM 알고리즘을 채널 LLR에 적용하는 신호 처리 기법을 소개하였다.

두 과정은 비이진 심벌과 비대칭 특성을 지닌 멀티 레벨 낸드 플래시 메모리 채널에서 좋은 성능을 갖는 연관정 복호를 수행하는 이진 오류 정정 부호를 설계하기 위해서 필요한 신호 처리 기법으로써 멀티 레벨

낸드 플래시 메모리 채널을 이진 무기억 대칭 채널로 바꾸어주는 역할을 수행한다. 즉, 멀티 레벨 낸드 플래시 메모리 채널에서 좋은 성능을 갖는 LDPC 부호나 Polar 부호를 설계하기 위해서 0 부호어가 전송되었다는 가정을 만족하도록 해줄 뿐만 아니라 채널 LLR l -밀도가 대칭 특성을 만족하도록 해준다. 이렇게 얻어진 채널 LLR l -밀도를 연관정 복호를 수행하는 이진 오류 정정 부호인 LDPC 부호 혹은 Polar 부호의 부호 설계 알고리즘에 적용하여 멀티 레벨 낸드 플래시 메모리 채널에서 좋은 성능을 갖는 부호를 설계할 수 있을 것으로 기대된다. 또한 멀티 레벨 낸드 플래시 메모리 채널과 유사한 환경에서도 좋은 성능을 가진 LDPC 부호나 Polar 부호를 설계하는데 도움을 줄 것으로 기대된다. 물론, 멀티 레벨 낸드 플래시 메모리가 physically degraded 채널^[5]이 아님을 유념해야 하며, 밀도 미러링과 EM 알고리즘은 멀티 레벨 낸드 플래시 메모리용 연관정 복호를 수행하는 이진 오류 정정 부호를 설계하기 위한 차선책으로 사용될 수 있다.

References

- [1] G. Dong, N. Xie, and T. Zhang, "On the use of soft-decision error-correction codes in NAND flash memory," *IEEE Trans. Circuits and Syst.*, vol. 58, no. 2, pp. 429-439, Feb. 2011.
- [2] G. Dong, S. Li, and T. Zhang, "Using data post-compensation and pre-distortion to tolerate cell-to-cell interference in MLC NAND flash memory," *IEEE Trans. Circuits and Syst.*, vol. 57, no. 10, pp. 2718-2728, Oct. 2010.
- [3] C. Yang, Y. Emre, and C. Chakrabatri, "Product code schemes for error correction in MLC NAND flash memories," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* vol. 29, no. 12, pp. 2302-2314, Dec. 2012.
- [4] H. Choi, W. Liu, and W. Sung, "VLSI implementation of BCH error correction for multilevel cell NAND flash memory," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 18, no. 5, pp. 843-847, May 2010.
- [5] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes

- under message passing decoding,” *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599-618, Feb. 2001.
- [6] E. Arikan, “Channel polarization : A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels,” *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051-3073, Jul. 2009.
- [7] S. ten Brink, “Convergence behavior of iteratively decoded parallel concatenated codes,” *IEEE Trans. Commun.*, vol. 49, pp. 1727-1737, Oct. 2001.
- [8] R. Mori and T. Tanaka, “Performance and construction of polar codes on symmetric binary-input memoryless channels,” in *Proc. IEEE Symp. Inf. Theory*, pp. 1496-1500, Seoul, South Korea, Jun. 2009.
- [9] C.-C. Wang, S. R. Kulkarni, and H. V. Poor, “Density evolution for asymmetric memoryless channels,” *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4216-4236, Dec. 2005.
- [10] C. M. Bishop, *Pattern Recognition and Machine Learning*, Cambridge, UK: Springer, 2006.
- [11] B. Lu, G. Yue, and X. Wang, “Performance analysis and design optimization of LDPC-coded MIMO OFDM systems,” *IEEE Trans. Signal Processing*, vol. 52, no. 2, pp. 348-361, Feb. 2004.
- [12] H. G. Lee, S.-R. Kim, M.-K Lee, J.-J. Kong and D.-J. Shin, “The threshold voltage control method for mitigating cell-to-cell interference in multi-level cell NAND flash memory,” in *Proc. KICS Int. Conf. Commun. 2012(KICS ICC 2012)*, pp. 79-80, Seoul, South Korea, Nov. 2012.
- [13] D. Park and J. Lee, “Performance of the coupling canceller with the various window size on the multi-level cell NAND flash memory channel,” *J. KICS*, vol. 37A, no. 8, pp. 706-711, Aug. 2013.
- [14] D. Lee and W. Sung, “Adaptive quantization scheme for multi-level cell NAND flash memory,” *J. KICS*, vol. 38C, no. 6, pp. 540-549, Jun. 2013.
- [15] S.-Y. Chung, T. J. Richardson, and R. L. Urbanke, “Analysis of sum product decoding of low-density parity-check codes using a Gaussian approximation,” *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 657-679, Feb. 2001.
- [16] S.-Y. Chung, G. D. Forney, T. J. Richardson, and R. L. Urbanke, “On the design of low-density parity-check codes within 0.0045 dB of the shannon limit,” *IEEE Commun. Lett.*, vol. 5, pp. 58-60, Feb. 2001.
- [17] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, “Design of capacity-approaching irregular low-density parity-check codes,” *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619-637, Feb. 2001.
- [18] J. Hou, P. H. Siegel, L. B. Milstein, and H. D. Pfister, “Capacity-approaching bandwidth-efficient coded modulation schemes based on low-density parity-check codes,” *IEEE Trans. Inf. Theory*, vol. 49, no. 9, pp. 2141-2155, Sep. 2003.

김 성 래 (Sung-Rae Kim)



2005년 2월 : 한양대학교 전자
전컴퓨터 공학부 졸업
2005년 3월~현재 : 한양대학교
전자컴퓨터통신공학과 석박
사 통합 과정
<관심분야> 오류 정정 부호,
디지털 통신, 신호 처리, 최
적화 이론, 기계 학습

신 동 준 (Dong-Joon Shin)



1990년 2월 : 서울대학교 전자
공학과 졸업

1991년 12월 : Northwestern
University 전자공학과 석사

1998년 12월 : University of
Southern California 전기공
학과 박사

1999년 1월~1999년 4월 : University of Southern
California, Research Associate

1999년 4월~2000년 8월 : Hughes Network Systems,
Member of Technical Staff

2002년 9월~현재 : 한양대학교 융합전자공학부 교수
<관심분야> 오류 정정 부호, 시퀀스, 디지털 통신,
정보 보호, 이산 수학