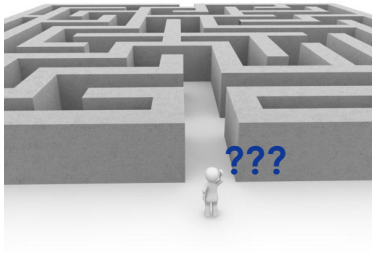
NAND Flash Status Prediction by Machine Learning

Cloud Zeng

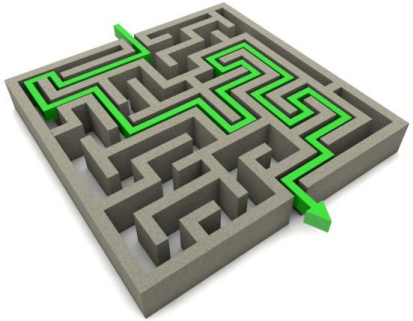
LiteOn/Storage/NVM Lab

The Evolution Of Error Handle

TRIAL & ERROR



ACROSS THE MAZE



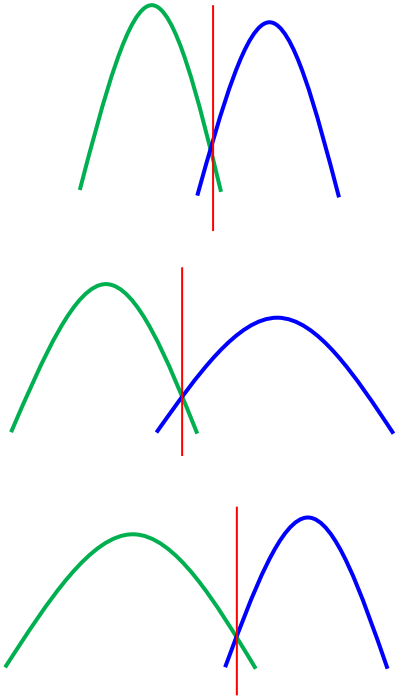
BREAK DOWN THE MAZE

EVOLUTION I

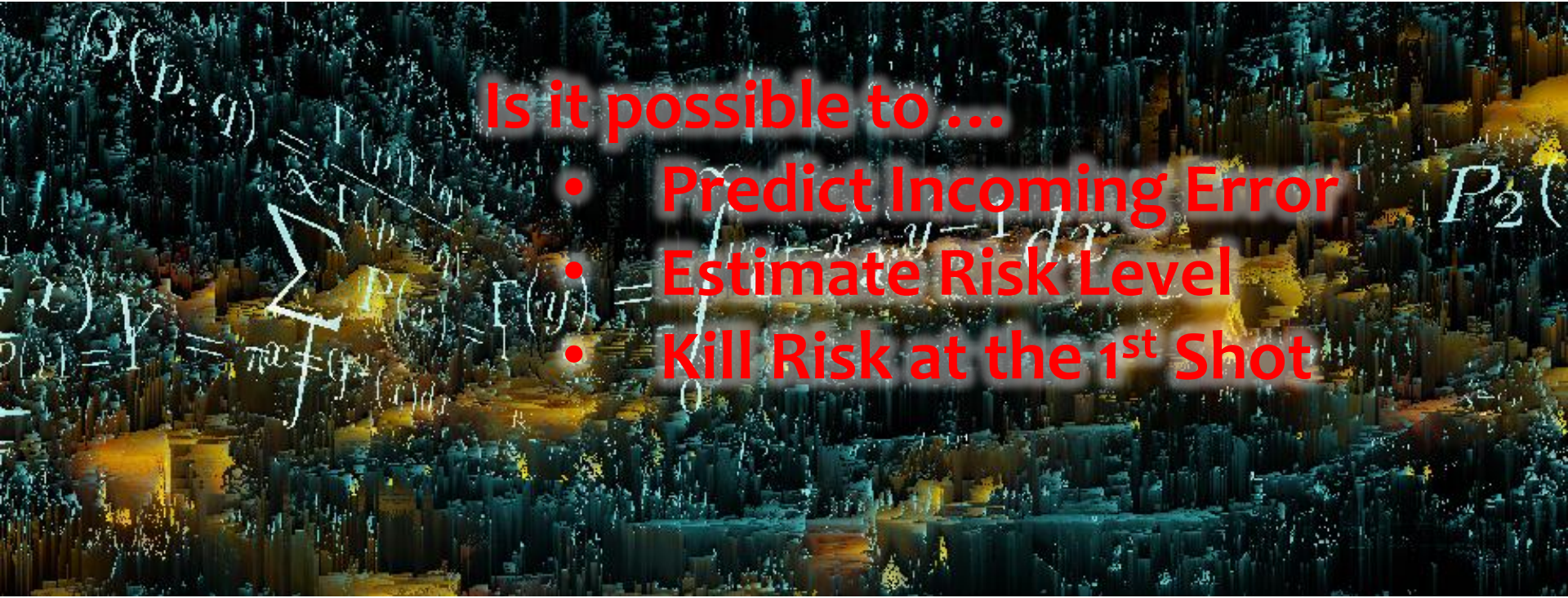
EVOLUTION II

Challenges in Error Handle

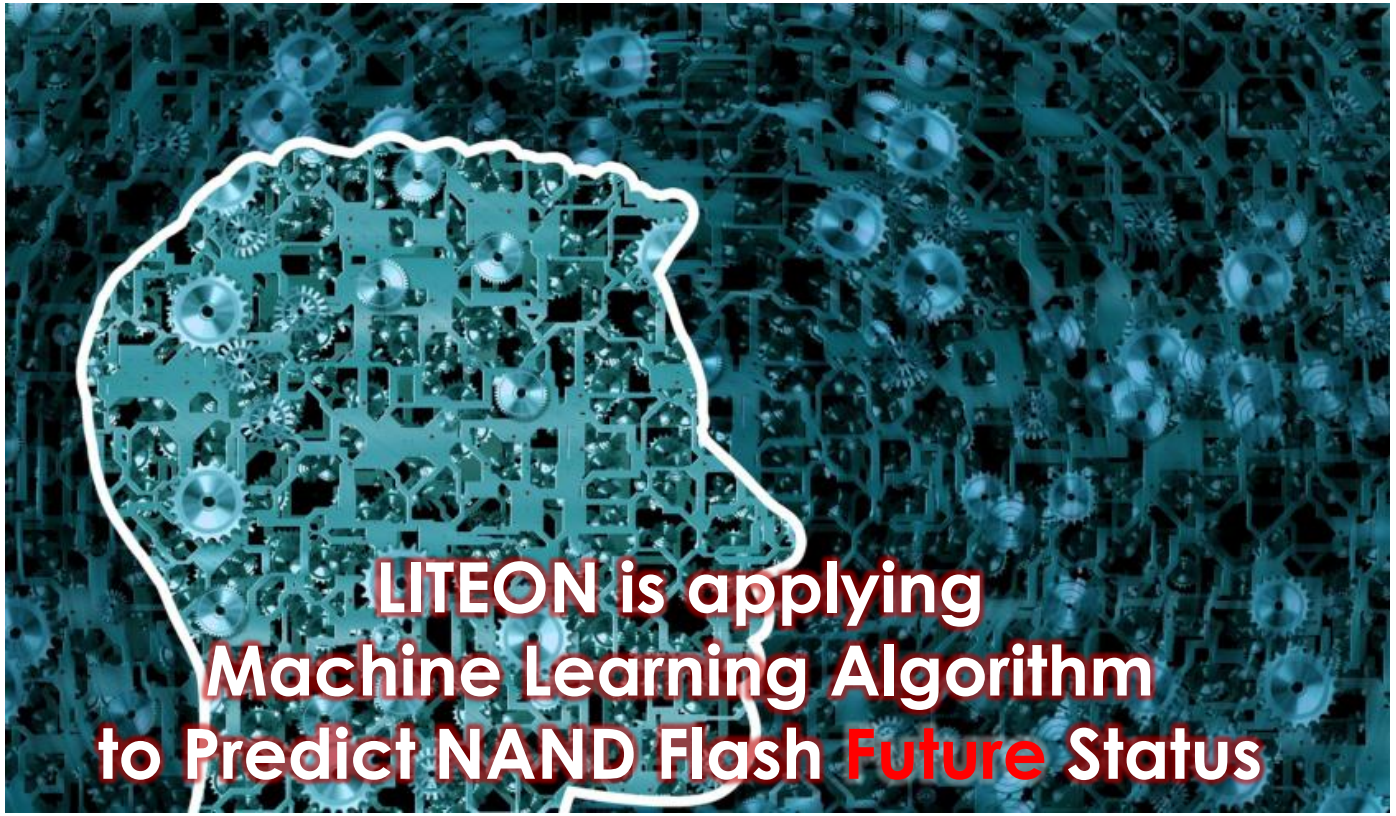
- **During SSD Service Time:**
  - P/E Cycle, Data Retention, Read Disturb
  - Critical R/W Condition(Temperature)
- **Decoding Strategy:**
  - Read Retry
  - Soft Decoding
  - …
- **Challenges:**
  - Keep High Throughput & High Reliability under Variant Operation Condition



Challenges in Error Handle

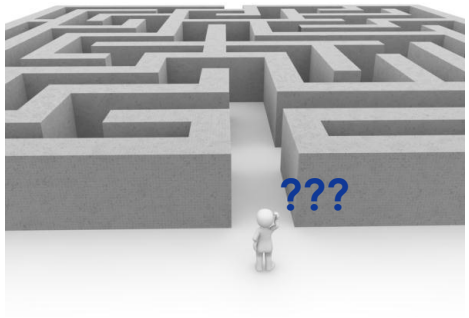


NAND Flash Status Prediction



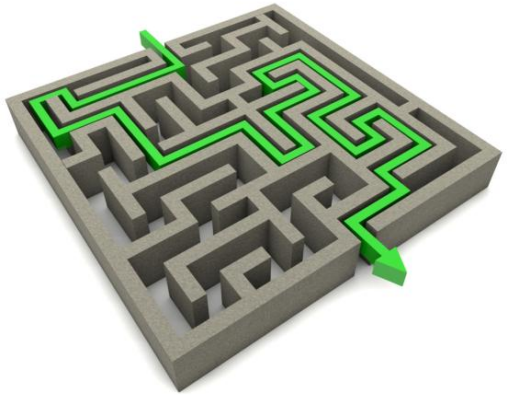
The Evolution I

TRIAL & ERROR



EVOLUTION I

ACROSS THE MAZE

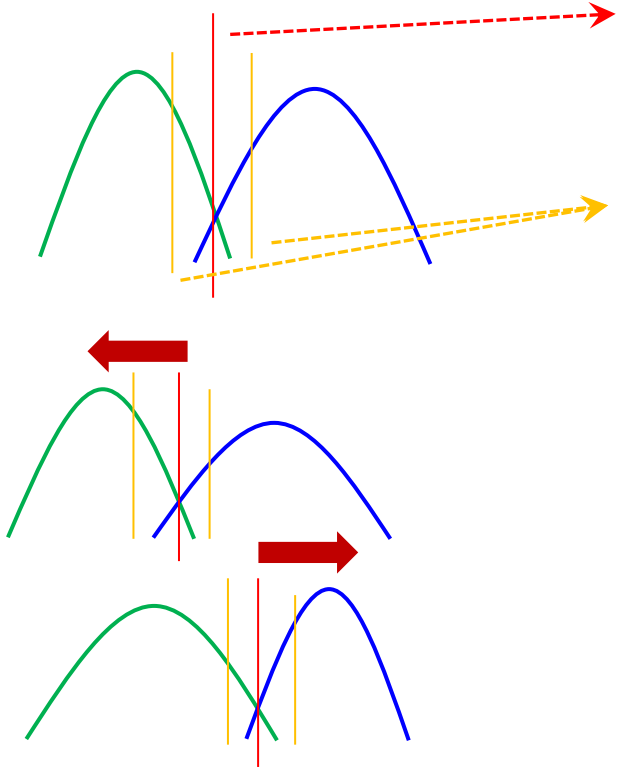


Efficient and Powerful

Optimal Decoding Parameters Prediction

Optimal Decoding Parameters

Example : Binary LDPC - MSA



1. Read Level for **Hard Bit**
→**Minimize Error Bits**

2. Read Level for **Soft Bit** & **LLR Value**
→ **Maximize the Decoding Capability**

- **Optimal Read Level & LLR Prediction**
  - Maximize Decoding Capability. Extend the Endurance
  - Vary with Operation Condition (P/E Cycle, Retention Time, Read Count, Temperature…)

Status Prediction

- **Input Parameters:**
  - Some factors will affect NAND Flash Status. (P/E Cycle, Retention Time, Read Count, Temperature…)
  - Some information from NAND Flash are also collected as Input Parameters (Program/Erase Time, …)
- **Status Prediction:**
  - Our target is to predict NAND Flash Status (Ex: Optimal Read Level, Error Recovery Flow …..) by Input Parameters.

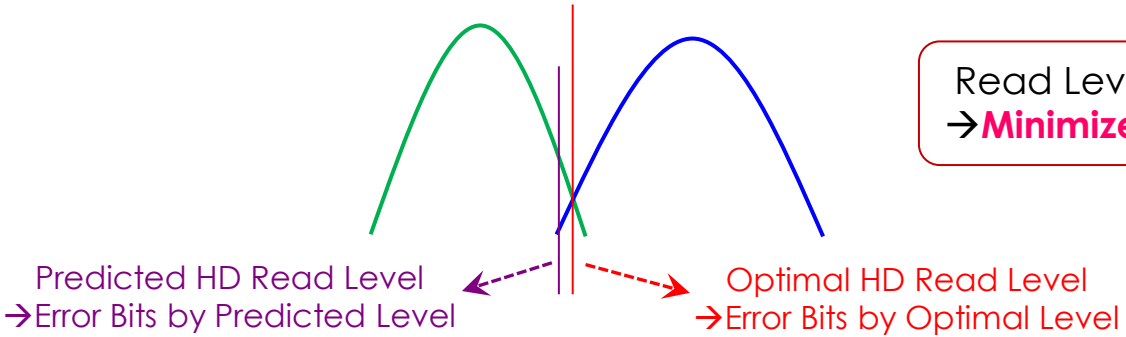
Optimal HD Read Level - Prediction

Example: Data Format

	Input Para 1	Input Para 2	Input Para 3	Input Para 4	Input Para 5	Input Para 6	Optimal HD Read Level
Data 1	1100	589	1794	6322	1000	1000	6
Data 2	932	908	1503	7849	500	500	-5
…	…						
Data N	990	842	1894	5692	300	400	3

- **What's the Optimal HD Read Level after n Days/Weeks?**
- **Regression Problem:**
  - Ordinary Least Square(OLS) Regression
  - Ridge Regression (Hoerl and Kennard, 1970)
  - Other Regression Analysis can be used to solve this problem

Optimal HD Read Level - Target



Read Level for **Hard Bit**
→**Minimize Error Bits**

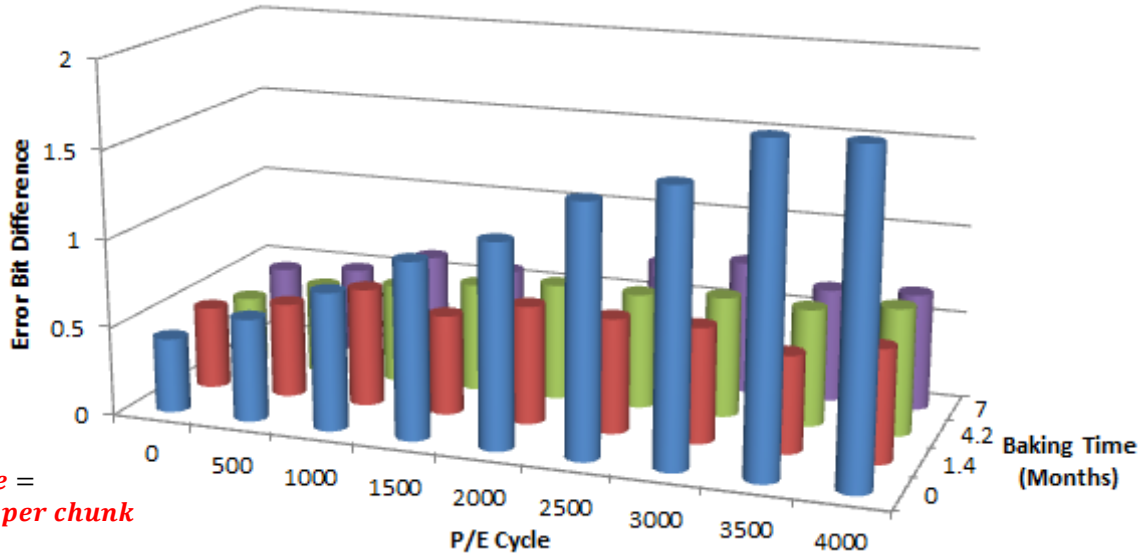
- **Target – Algorithm Level**
  - $minimize\ E[(Predicted\ HD\ Read\ Level\ -\ Optimal\ HD\ Read\ Level)^2]$
    $make\ Prediction\ Error\ as\ small\ as\ possible$
- **Target – Real NAND Application**
  - $minimize\ Error\ Bits\ Difference\ between\ Predicted\ Read\ Level\ and\ Optimal\ Read\ Level$

Optimal HD Read Level - Evaluation

Category	Item	Description	Remark
P/E	Cycle	0, 500, …~	
	Temperature	(Random)	
	Dwell	(Random)	
Test Item	Data Retention	0, 24, … ~ (HRs)	Room Temperature
	Date Retention	0, 6, … ~ (HRs)	High Temperature
	Read Disturb	0, 1000, … ~	

- **One** prediction model is applied to the NAND Flash under different operation condition {PE, DR, RD, …}
- Advantage of our prediction model is **we don't have to know {PE, DR, RD}**
- We can predict the optimal read level for **any combination of {PE, DR, RD}**

Optimal HD Read Level – HT DR

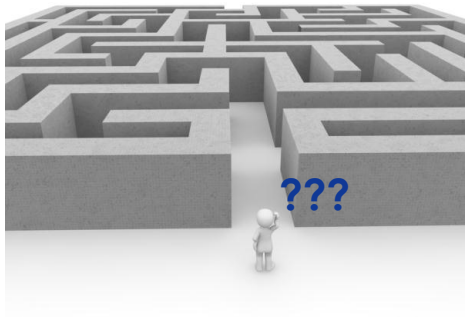


$$Error\ Bit\ Difference = (Predict - Opt)\ Error\ Bits\ , per\ chunk$$

- Our prediction model is suitable for variant operation condition {PE, DR}
- Maximum Error Bit Difference is only 2 bits/chunk

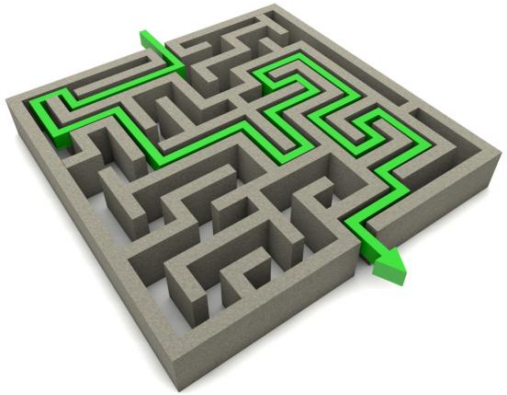
The Evolution I

TRIAL & ERROR



EVOLUTION I

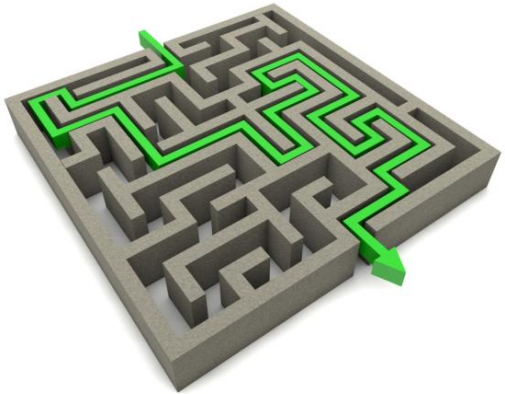
ACROSS THE MAZE



- **Current/Future Optimal Read Level & LLR Prediction**
  - Reduce the Latency → **Efficient**
  - Maximize the Decoding Capability → **Powerful**

The Evolution II

ACROSS THE MAZE



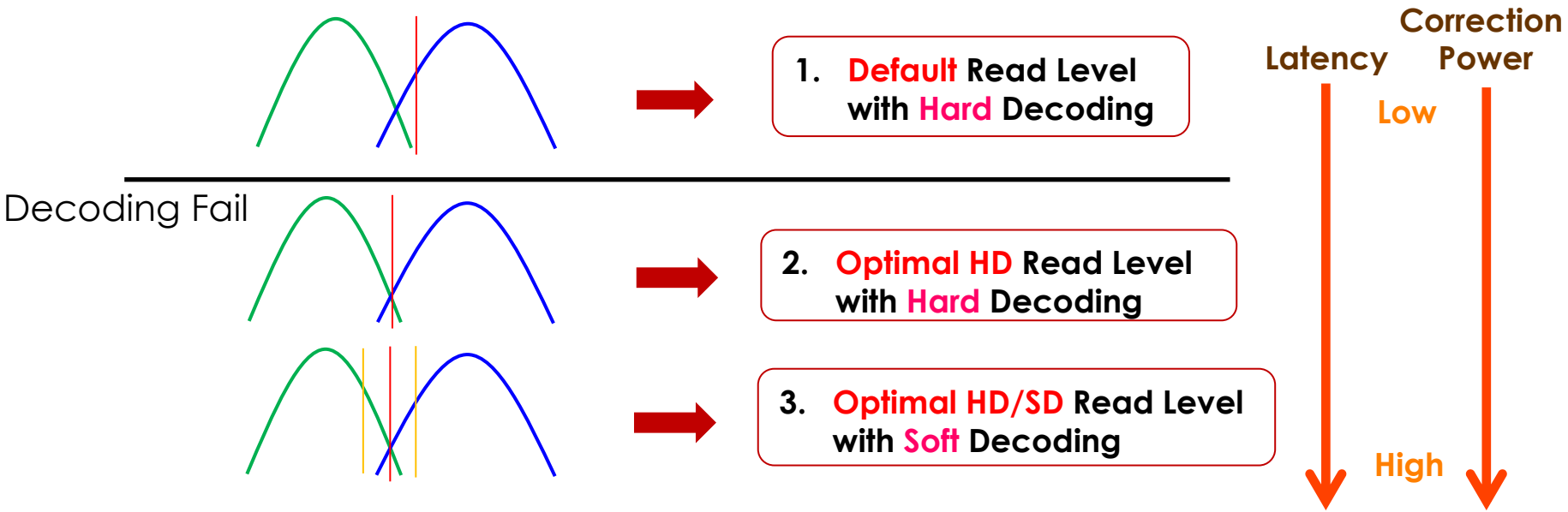
Decoding Parameters
Prediction

EVOLUTION II

BREAK DOWN THE MAZE

Error Recovery Flow
Prediction

Error Recovery Flow



- **Error Recovery Flow Prediction**
  - Throughput Control, End of Life Prediction
  - Vary with Operation Condition , Decoding Capability …

Error Recovery Flow - Prediction

Error Recovery Flow

1. Default Read Level with Hard Decoding
2. Optimal HD Read Level with Hard Decoding
3. Optimal HD/SD Read Level with Soft Decoding
4. Fail

Example: Data Format

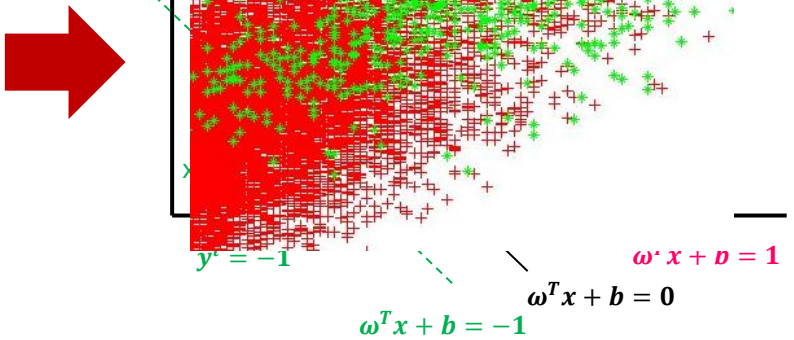
	Input Para 1	Input Para 2	Input Para 3	Error Recovery Flow
Data 1	1100	589	1794	1
Data 2	932	908	1503	2
…	…			
Data N	990	842	1894	1

- **What's the appropriate Error Recovery Flow after n Days/Weeks?**
- **Classification Problem:**
  - SVM (Support Vector Machine) or other Classification Algorithm can be used to solve this problem

Error Recovery Flow - Target

Error Recovery Flow

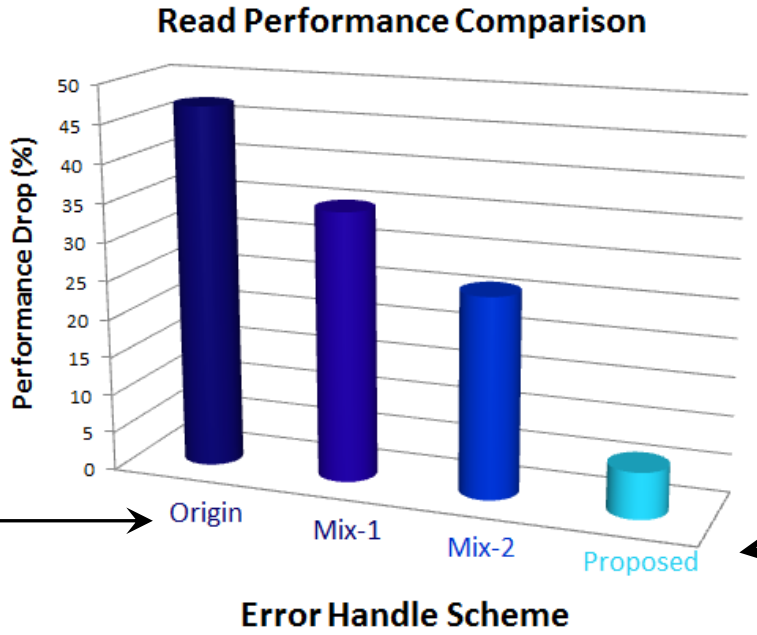
1. Default Read Level with Hard Decoding
2. Optimal HD Read Level with Hard Decoding



Simple Classification Example

- **Target**
  - *minimize* $Misclassification\ Rate$
    $Denote\ P(Predict\ Flow = 2 \mid Real\ Flow = 1)\ as\ False\ Alarm$
    $Denote\ P(Predict\ Flow = 1 \mid Real\ Flow = 2)\ as\ Miss$

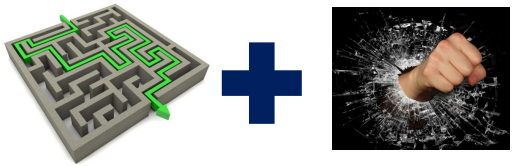
Read Performance Comparison



TRIAL & ERROR

ACROSS THE MAZE

BREAK DOWN THE MAZE



- TRIAL & Error (Origin) : Drop 47 %
- Proposed Scheme: Drop 6 %

Summary

ACROSS THE MAZE

BREAK DOWN THE MAZE

- **Future Error Recovery Flow Prediction**
  - FW handle possible fail blocks in advance
  - Warn our User when the Drive approaching EOL

Thank you

Q & A