DATA QUALITY METRIC (DQM) – HOW ACCURATE DOES IT NEED TO BE?

ITC 2023

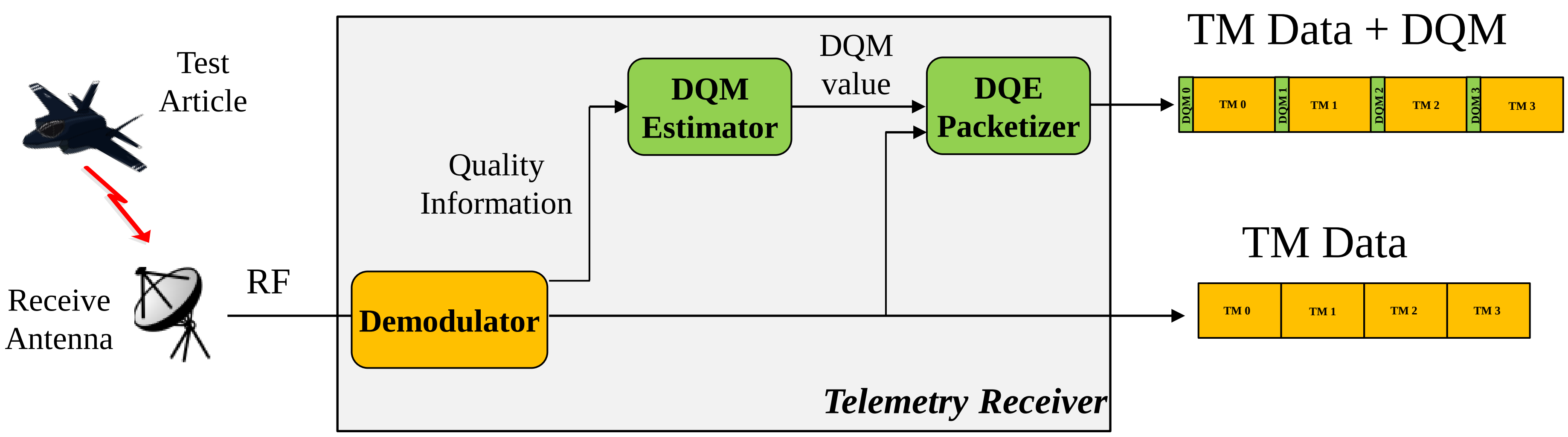
Mark Geoghegan
Quasonix, West Chester, OH
markg@quasonix.com

DQE/DQM Background

Milestones	Contributions
2015 Hill ITC Paper - "Metrics and Test Procedures for Data Quality Estimation in the Aeronautical Telemetry Channel"	DQE/DQM transport protocol to support efficient multiple source TM combining
2015 Rice/Perrins ITC Best Paper -"Maximum Likelihood Detection from Multiple Bit Sources"	Optimal combining recipe and performance analysis - Maximum Likelihood Bit Detection (MLBD)
IRIG 106-22 Chapter 2, Appendix 2-G - "Standards for Data Quality Metrics and Data Quality Encapsulation"	Standardized in IRIG 106-22
2022 Temple ITC Paper - "Some Thoughts on Testing the Data Quality Metric"	Proposed DQE/DQM verification test methods
IRIG 118-22 Release 2 Volume 2., Chapter 11, 2022 "Test procedures for assessing telemetry receiver data quality metrics"	DQE/DQM verification test methods standardized in IRIG 118-22
Industry Day on DQE/DQM assessment 10/28/22	Initial multi-vendor DQE/DQM performance/compatibility testing

Data Quality Metric (DQM)

- 16 bit value that indicates the BEP for a ‘single’ block of TM data bits
- Periodically inserted in TM data output stream

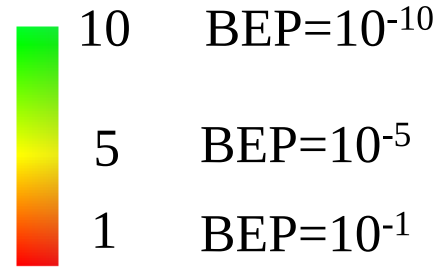


Quasonix

"Gambling Quality Metric (GQM)"

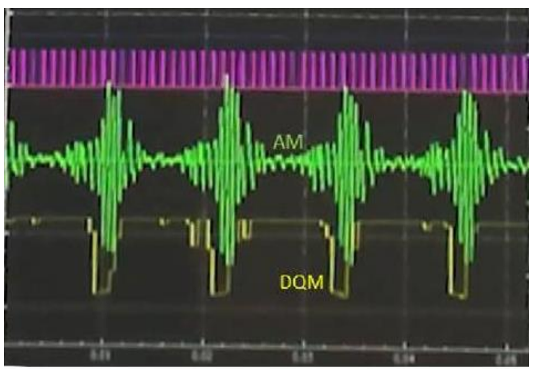
- Probability of winning for the next N plays
- Play multiple games simultaneously to further improve odds!
- Don't get excited, it doesn't exist…

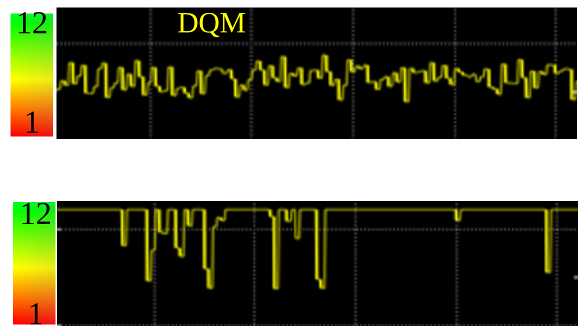
DQM Telemetry Applications



- Provides estimate of TM quality without a priori knowledge of data.
- Provides real-time indication of the TM channel.

Prop Chop



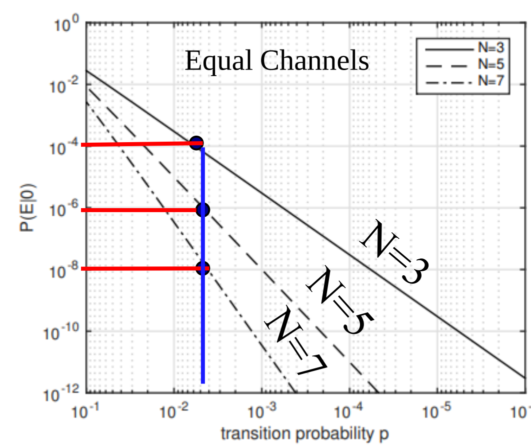


Noise BEP=10^{-5}

Interference BEP=10^{-5}

Which one would you choose?

- Provides sufficient information for optimal multi-channel telemetry reception.
  - *Potential for virtually 'error-free' telemetry*.

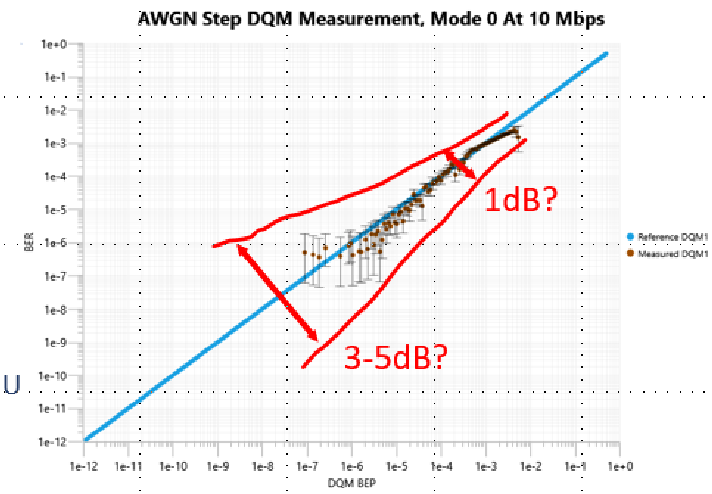


QUASONIX

Industry Day (10/22) - Preliminary DQE/DQM integration
(RFVWG meets RCC TG RF Systems - Yuma March 2023 - Final.pdf)

DQM assessment results at a glance

- Test scenarios
  - 50 test scenarios tested
    - 22 for 11.1 – Sweep AWGN - static
    - 24 FOR 11.3 – ACI - static
    - 4 for 11.2 – Step AWGN - dynamic
  - All results logged into one consolidated matrix
    - BER & BEP (along with Min & Max - spread)
    - Raw matrix not meaningful to be shared with U
- Conclusions
  - All TM receiver vendor present are COMPLIANT with the DQE encapsulation format (version 0)
    - Note Version 1 was not tested

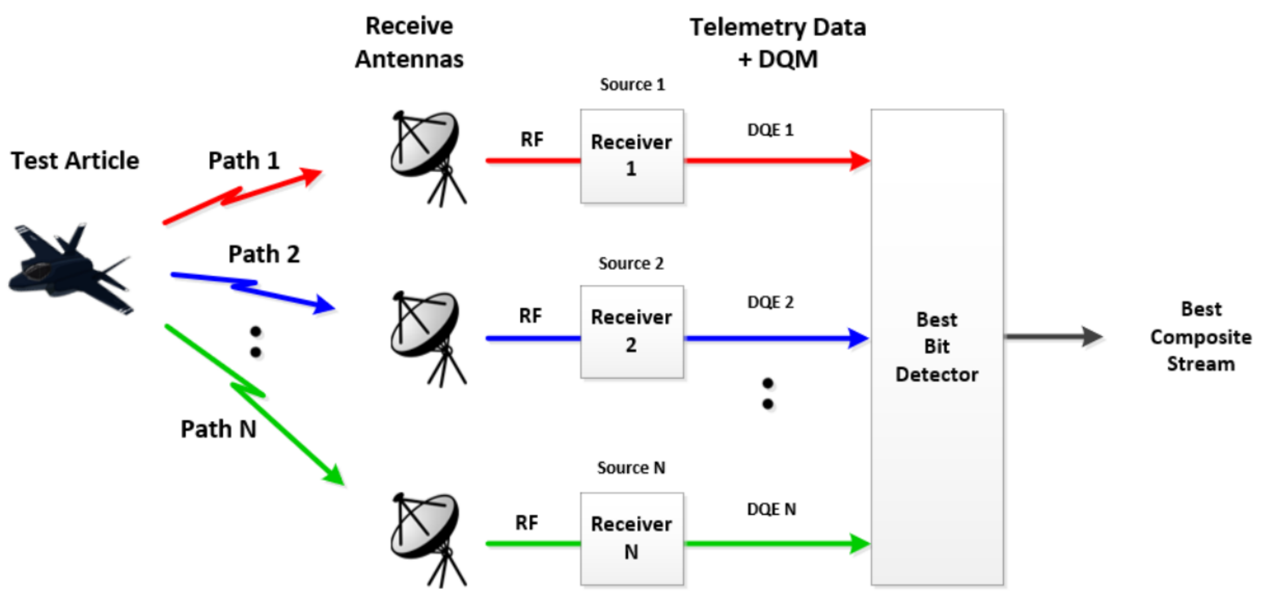


- As for validating the assessment of the DQM, the IRIG 118 ch11 misses a Criteria or threshold for PASS/FAIL.
  - Concept of Variance (distance) between BER and BEP
    - 1?dB for noisy to very noise channel (ie BER lower than 10-4)
    - 3-5?dB for noisy to very noise channel (ie BER better than 10-6)
    - Other thresholds to define a “trumpet mask” as shown in red above

Quasonix

DQM Accuracy:
How good is good enough?

- How accurate does the DQM estimate need to be to not "***significantly***" degrade the MLBD output performance?
- *Need to understand relationship between DQM and MLBD*



DQM Estimation Error ??? MLBD BEP Loss

Quasonix

Aeronautical Telemetry channels

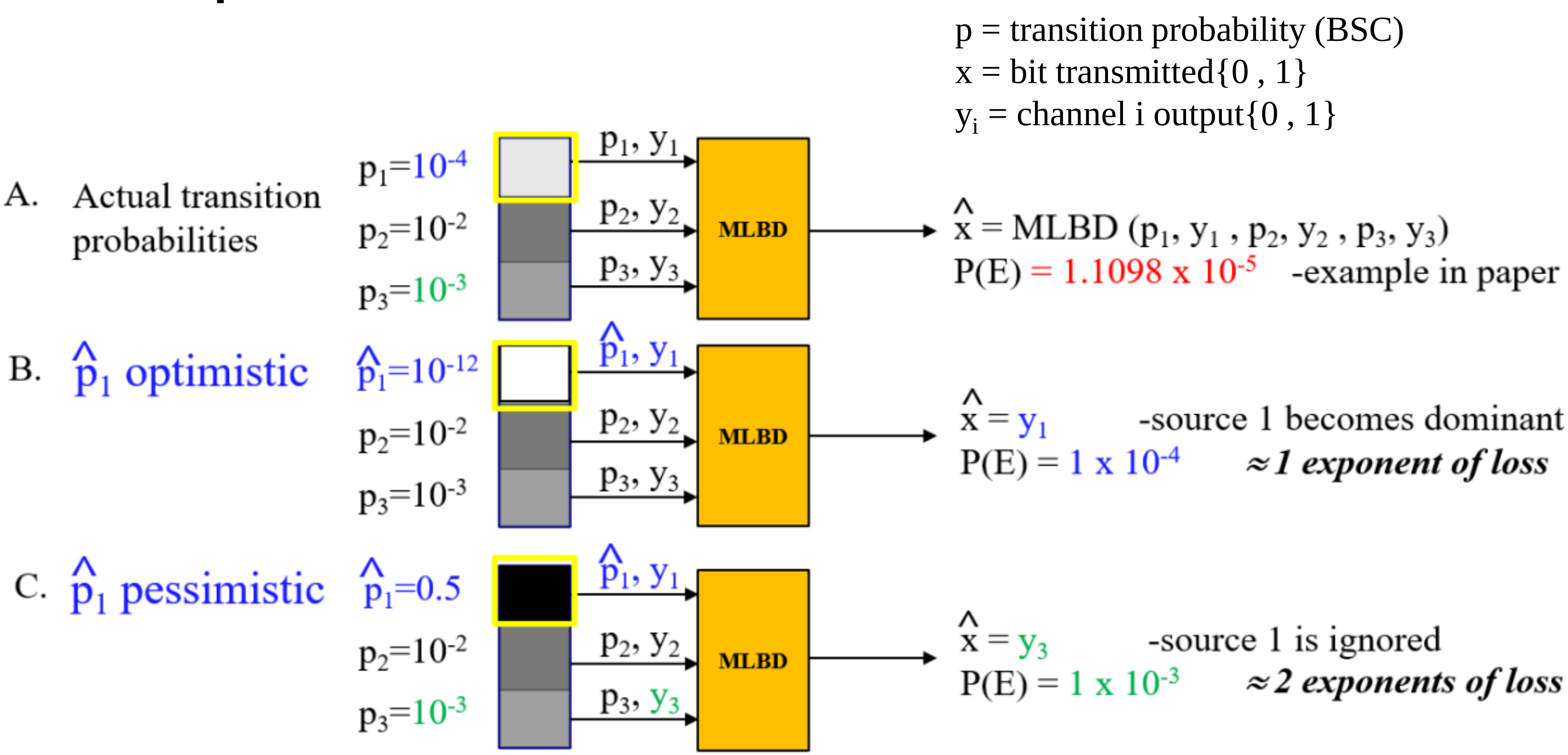
- Variety of channel conditions
  - AWGN, ACI, Multipath, Flat Fading, Interference
- 2022 ITC Kip Temple paper (DQM testing)
  - ***"The DQM must be relevant for all channel conditions or sources of signal corruption that may exist during telemetry operations."***
  - ***"The key to assigning a DQM value is an <u>accurate</u> and <u>consistent</u> estimate of the bit error probability that applies to the <u>data that follows the estimate</u>."***
- Can't control the hand that you're dealt!

⋮

Must be able to play all hands well

How does DQM affect MLBD?

- Example: Extreme over and under estimation



Misestimation of p can degrade MLBD performance

MLBD Calculations (Ideal and estimated DQM)

- Ideal DQM
  - Bit decision is based on comparison of LLR sums of group that votes 0 versus 1

$$\sum_{n\in\mathcal{N}_0} \log\left(\frac{1-p_n}{p_n}\right) \gtrless \sum_{n\in\mathcal{N}_1} \log\left(\frac{1-p_n}{p_n}\right) \xrightarrow{\text{x = 0}} P_e = \sum_{i\epsilon\mathcal{I}_+} P[i]$$

- Estimated DQM

$$\sum_{n\epsilon \mathrm{N}_0} \log\left(\frac{1-\hat{p}_n}{\hat{p}_n}\right) \gtrless \sum_{n\epsilon \mathrm{N}_1} \log\left(\frac{1-\hat{p}_n}{\hat{p}_n}\right) \xrightarrow{\text{x = 0}} \hat{P}_e = \sum_{i\epsilon\hat{\mathcal{I}}_+} P[i]$$

MLBD probability of error for ideal and estimated DQM

- Bias

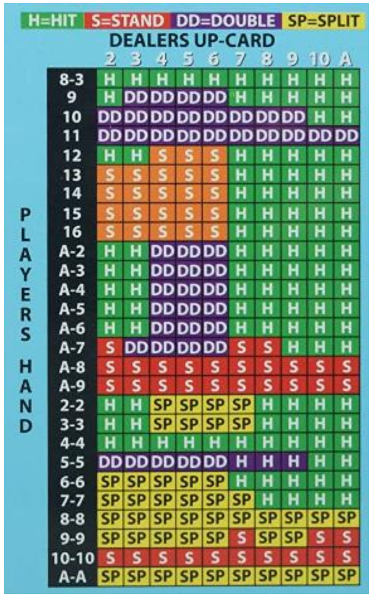
$$\log\left(\frac{1-\hat{p}_n p_{bias}}{\hat{p}_n p_{bias}}\right) \xrightarrow{\hat{p}_n p_{bias} \ll 0} -\log\left(\hat{p}_n\right) - \log\left(p_{bias}\right)$$

Systematic estimation bias offsets all terms

MLBD calculations are just like "Perfect Play" tables (fixed)

- Procedure to play each hand for maximum possible return
- Note: These are fixed tables (best way to play all possible hands)

Blackjack

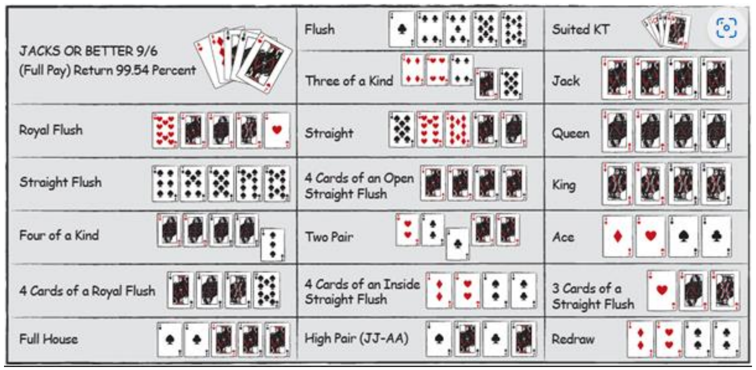


H=HIT S=STAND DD=DOUBLE SP=SPLIT

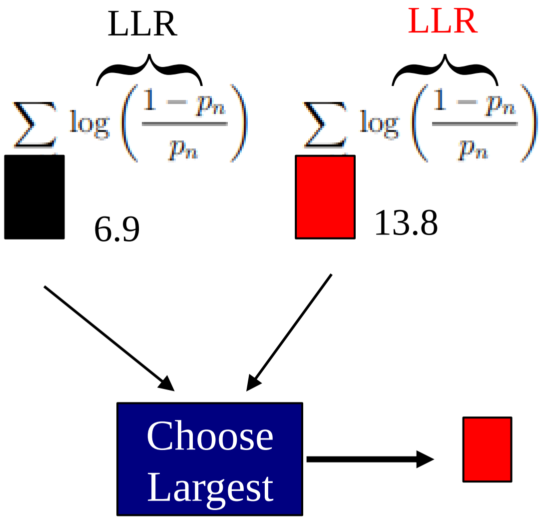
DEALERS UP-CARD

PLAYERS HAND	2	3	4	5	6	7	8	9	10	A
8-3	H	H	H	H	H	H	H	H	H	H
9	H	DD	DD	DD	DD	H	H	H	H	H
10	DD	DD	DD	DD	DD	DD	DD	DD	H	H
11	DD	DD	DD	DD	DD	DD	DD	DD	DD	DD
12	H	H	S	S	S	H	H	H	H	H
13	S	S	S	S	S	H	H	H	H	H
14	S	S	S	S	S	H	H	H	H	H
15	S	S	S	S	S	H	H	H	H	H
16	S	S	S	S	S	H	H	H	H	H
A-2	H	H	DD	DD	DD	H	H	H	H	H
A-3	H	H	DD	DD	DD	H	H	H	H	H
A-4	H	H	DD	DD	DD	H	H	H	H	H
A-5	H	H	DD	DD	DD	H	H	H	H	H
A-6	H	H	DD	DD	DD	H	H	H	H	H
A-7	S	DD	DD	DD	DD	S	S	H	H	H
A-8	S	S	S	S	S	S	S	S	S	S
A-9	S	S	S	S	S	S	S	S	S	S
2-2	H	H	SP	SP	SP	SP	H	H	H	H
3-3	H	H	SP	SP	SP	SP	H	H	H	H
4-4	H	H	H	H	H	H	H	H	H	H
5-5	DD	DD	DD	DD	DD	H	H	H	H	H
6-6	SP	SP	SP	SP	SP	H	H	H	H	H
7-7	SP	SP	SP	SP	SP	SP	H	H	H	H
8-8	SP	SP	SP	SP	SP	SP	SP	SP	SP	SP
9-9	SP	SP	SP	SP	SP	S	SP	SP	S	S
10-10	S	S	S	S	S	S	S	S	S	S
A-A	SP	SP	SP	SP	SP	SP	SP	SP	SP	SP

Video Poker
(Jacks or Better)

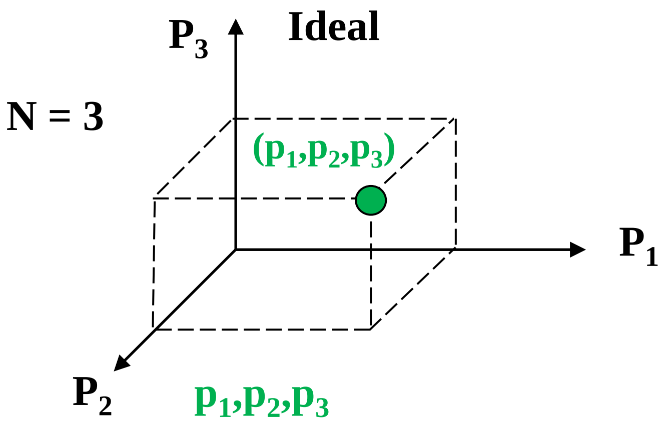


MLBD



Fun Fact - Brute force DQM MLBD card has $2^{((16+1)N)}$ rows!

MLBD Simulation Approach (compute loss over all combinations)

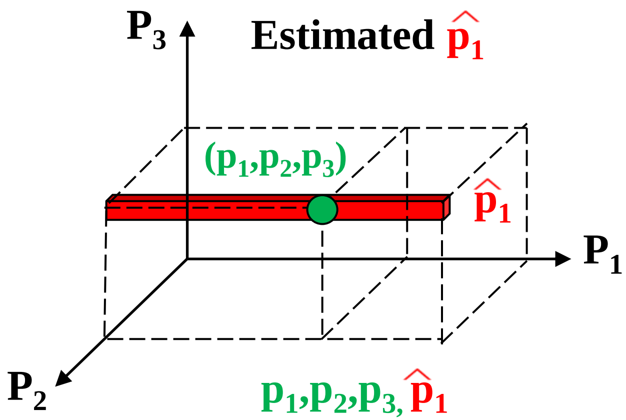


p_1, p_2, p_3

MLBD Probability of Error Calculation

$P(E|p_1, p_2, p_3)$

Ideal MLBD with perfect p



$p_1, p_2, p_3, \hat{p}_1$

MLBD Probability of Error Calculation

$P(E|p_1, p_2, p_3, \hat{p}_1)$

MLBD output as a function of p_1 estimate

Probability Cost $= P(E|p_1, p_2, p_3) - P(E|p_1, p_2, p_3, \hat{p}_1)$

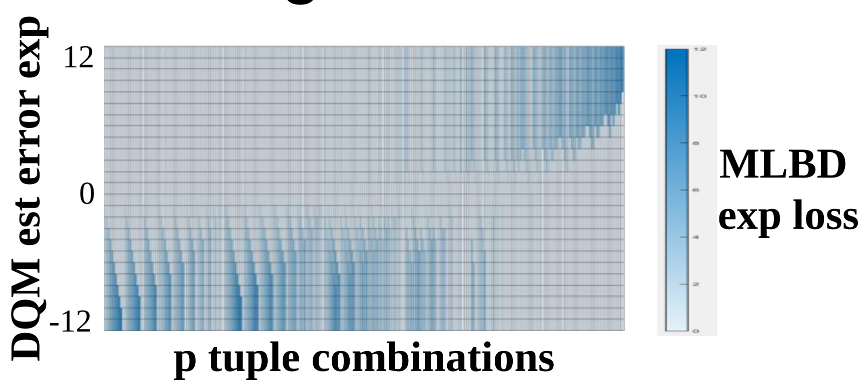
Error Exp Cost $= \log 10(P(E|p_1, p_2, p_3)) - \log 10(P(E|p_1, p_2, p_3, \hat{p}_1))$

Simulation Parameters (start out simple)

- DQM resolution
  - IRIG standard 2^{16} = 65536 values BEP=0.5 to 10^{-12}
  - Start with reduced resolution of 1 exponent (0 to 12)

- N > 2 but unknown
  - Start with reasonable number of sources N=3,5,7
- Math Equation vs Monte Carlo
  - Prove that P(E) formulas agree with brute force
- Gain understanding of basic relationships

N=3 Example:



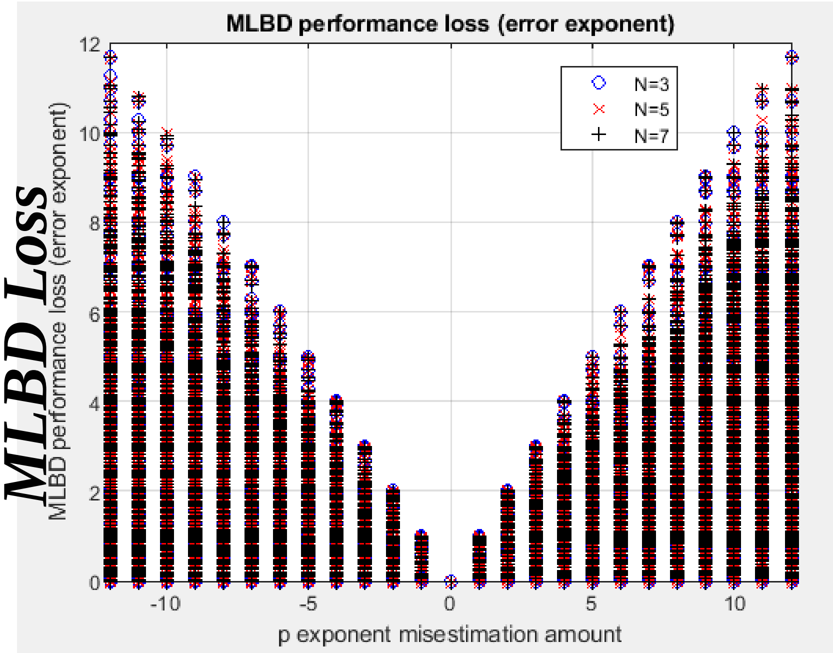
Simulation Results

- Calculate MLBD performance loss vs estimation error

Error Exp Cost = log10(P(E|p1,p2,p3)) –log10(P(E|p1,p2,p3,$\hat{p}_1$))

Coarse Sampling (1 exp)

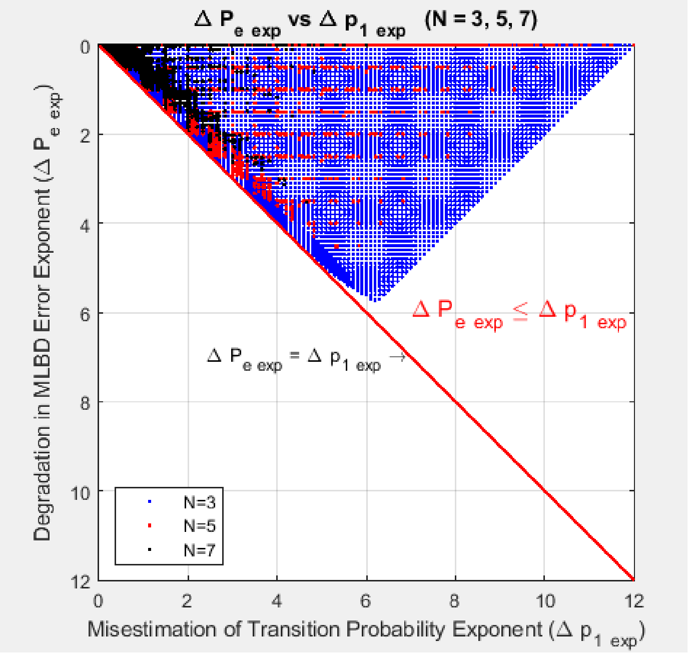
MLBD Loss



DQM Estimation Error

Fine Sampling (0.1 exp) (| | & axis flipped)

MLBD Loss



|DQM Estimation Error|

Simulations suggest that MLBD loss exponent is bounded by DQM estimation error exponent!

Quasonix

Can this bound be shown to apply for all N and p tuples?

- When all else fails, dig into the math…
- What does it take to degrade P(E)?
  - Answer: Adjust p_1 to flip the smallest positive M[i] negative
- Math required to degrade MLBD (see appendix for details)

$$\log\left(\frac{1-\hat{p}_1}{\hat{p}_1}\right) = \log\left(\frac{1-p_1}{p_1}\right) + M[i_{\text{min+}}] \implies \hat{p}_1 = \frac{p_1}{p_1 + (1-p_1)e^{M[i_{\text{min+}}]}}$$

$$\hat{P}_e = \sum_{i \epsilon \mathrm{I}_+} P[i] - P[i_{\text{min+}}] + P[i_{\text{max-}}]$$

Smallest change in $\hat{p}_1$ to degrade $\hat{P}_e$

$$\Delta p_{1_{\text{exp}}} = log_{10}\left(\frac{\hat{p}_1}{p_1}\right) \quad , \quad \Delta P_{e_{\text{exp}}} = log_{10}\left(\frac{\hat{P}_e}{P_e}\right)$$

MLBD Loss DQM Error

$$\Delta P_{e_{\text{exp}}} \leq |\Delta p_{1_{\text{exp}}}|$$

Mathematical analysis shows that MLBD loss exponent is bounded by DQM estimation error exponent regardless of N or p!

Quasonix

Implication on DQM accuracy

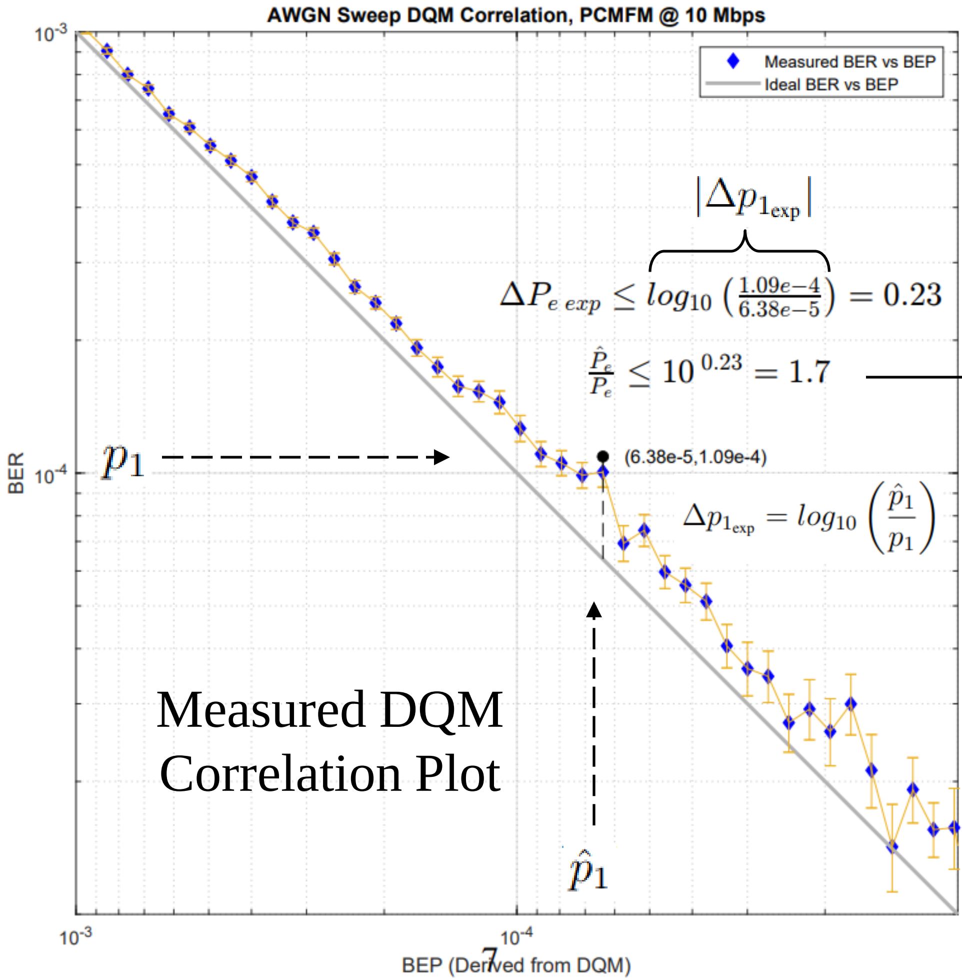


Figure 5: DQM Correlation Plot with Bound Calculation Example

MLBD BEP performance loss could be approximately twice that of ideal system

DQM measurement thresholds along with maximum MLBD system loss

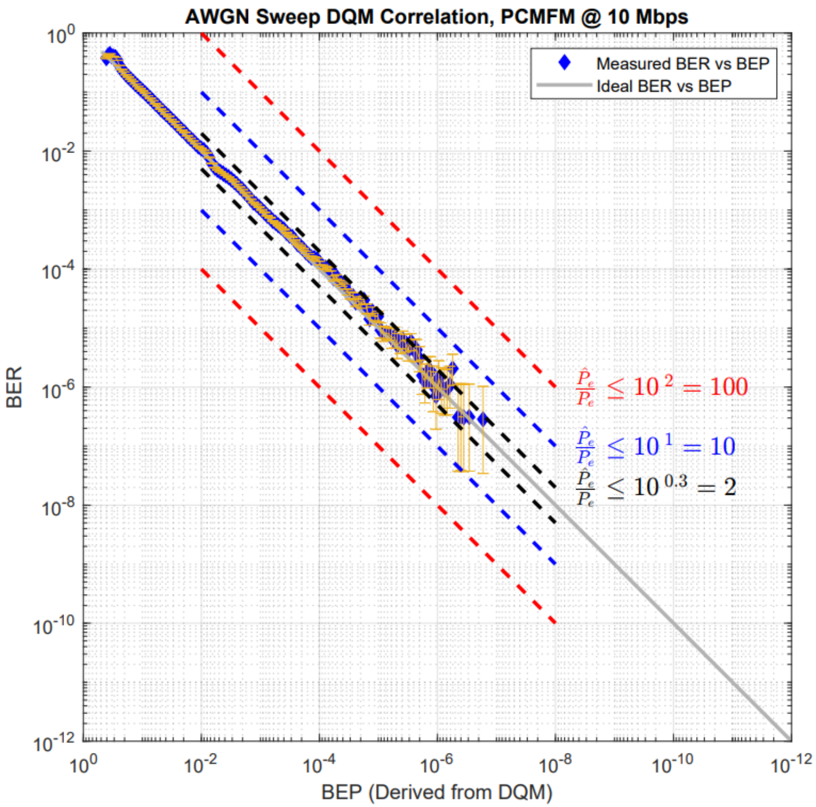


Figure 6: DQM Correlation Plot with various thresholds

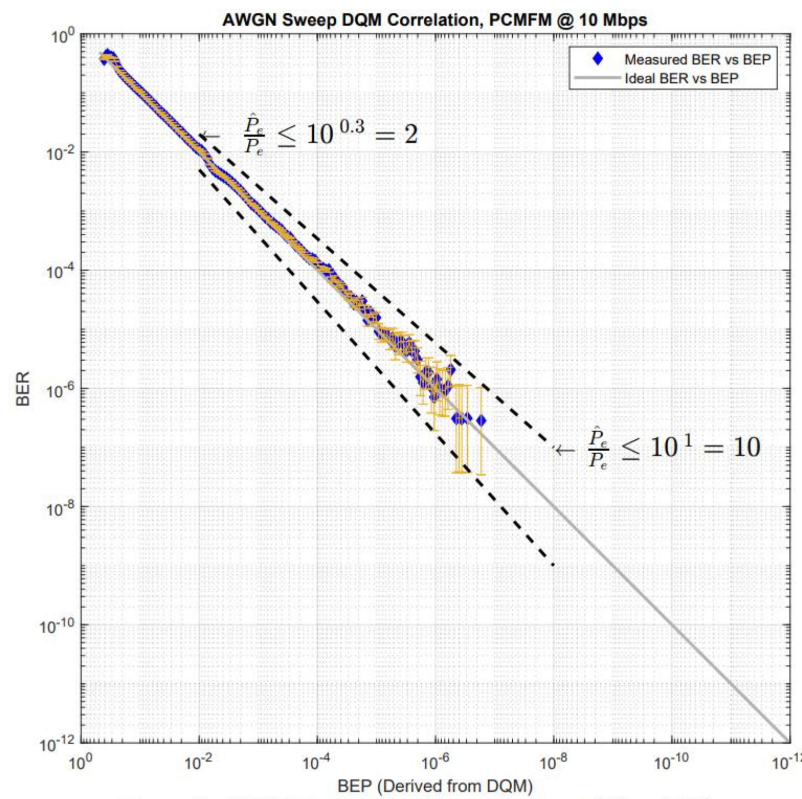


Figure 7: DQM Correlation Plot with tapered threshold

RCC has a new tool to understand implications of DQM accuracy thresholds

Quasonix

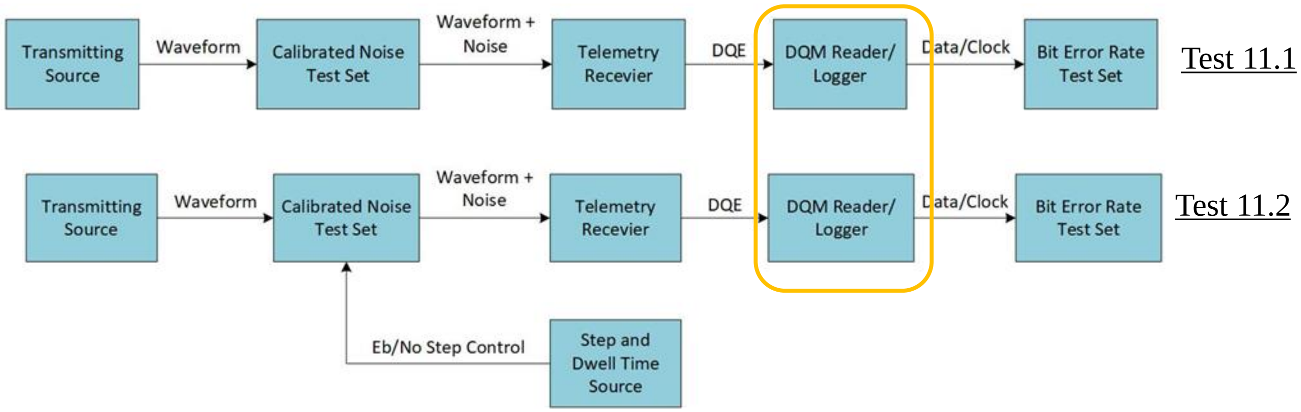
DQM testing standard

- 2022 ITC Temple paper – “Some thoughts on testing the Data Quality Metric”
  - ***“The key to assigning a DQM value is an <u>accurate</u> and <u>consistent</u> estimate of the bit error probability that applies to the <u>data that follows the estimate</u>.”***
  - ***“The DQM must be relevant for all channel conditions or sources of signal corruption that may exist during telemetry operations.”***
- IRIG 118-22 Release 2 Volume 2 Chapter 11

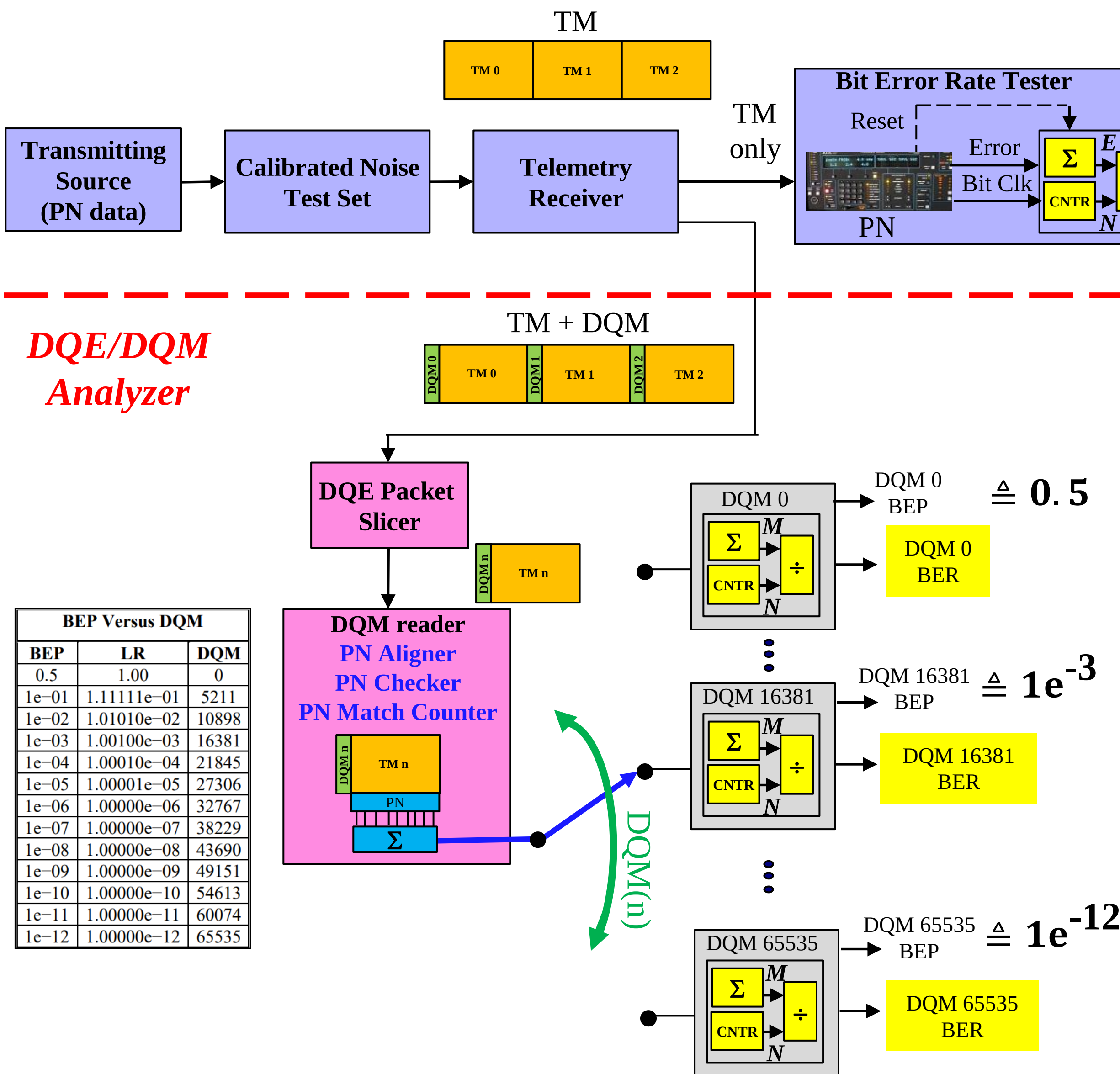
https://www.trmc.osd.mil/wiki/download/attachments/113019772/118-22_Vol%202R2-Test_Methods_TM_Systems_Subsystems.pdf?version=1&modificationDate=1665520822889&api=v2

Table 11-2. Test Matrix for Data Quality Metric Testing

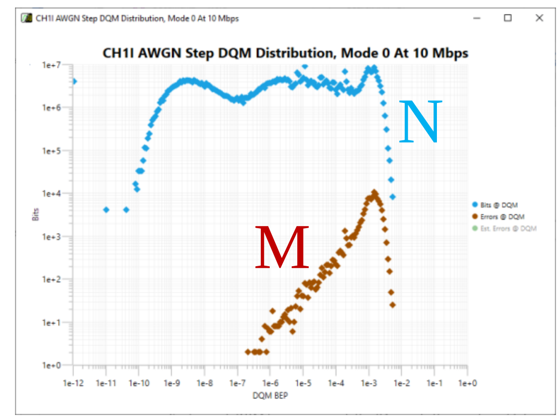
Test Number	Test Description
11.1	BER vs BEP with Additive Noise
11.2	DQM (BEP) Step and Dwell Response
11.3	BER vs BEP with Adjacent Channel Interference
11.4	BER vs BEP for Static 3-Ray Multipath Channel Conditions
11.5	BER vs BEP for Static 2-Ray Multipath Channel Conditions
11.6	DQM (BEP) Resynchronization Response

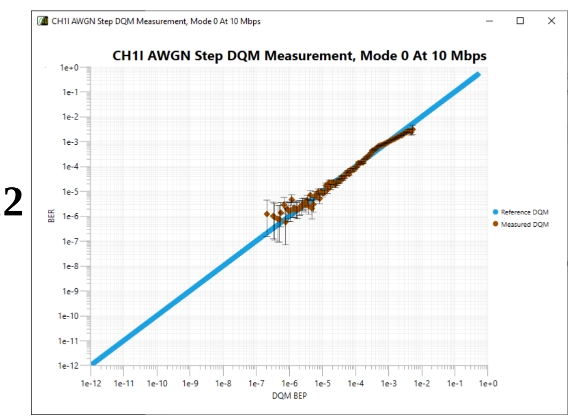


DQM testing – setup and measurements



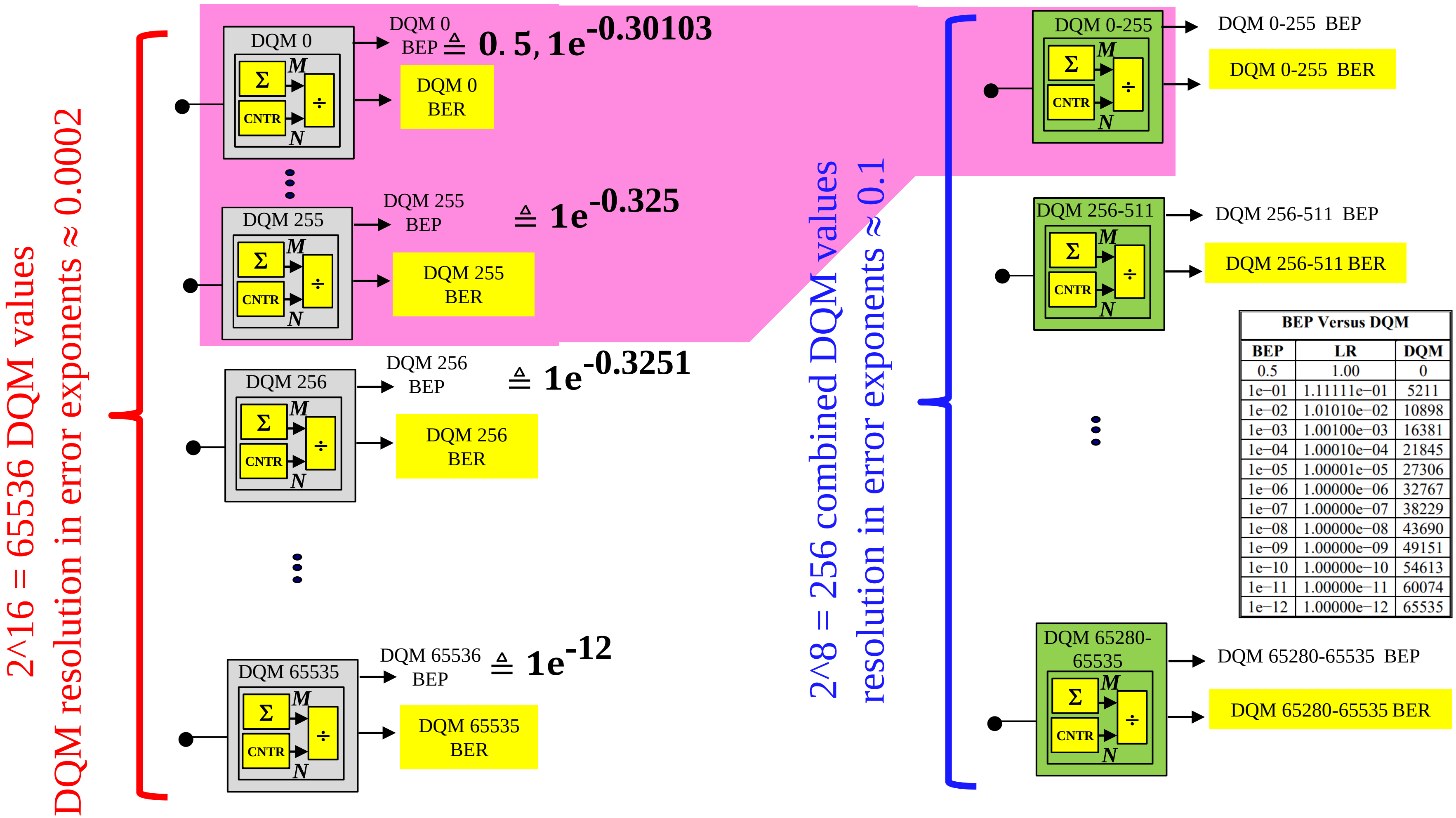
BEP Versus DQM		
BEP	**LR**	**DQM**
0.5	1.00	0
1e−01	1.11111e−01	5211
1e−02	1.01010e−02	10898
1e−03	1.00100e−03	16381
1e−04	1.00010e−04	21845
1e−05	1.00001e−05	27306
1e−06	1.00000e−06	32767
1e−07	1.00000e−07	38229
1e−08	1.00000e−08	43690
1e−09	1.00000e−09	49151
1e−10	1.00000e−10	54613
1e−11	1.00000e−11	60074
1e−12	1.00000e−12	65535





Quasonix

DQM BER resolution vs measurement time ("binning")



BEP Versus DQM

BEP	LR	DQM
0.5	1.00	0
1e−01	1.11111e−01	5211
1e−02	1.01010e−02	10898
1e−03	1.00100e−03	16381
1e−04	1.00010e−04	21845
1e−05	1.00001e−05	27306
1e−06	1.00000e−06	32767
1e−07	1.00000e−07	38229
1e−08	1.00000e−08	43690
1e−09	1.00000e−09	49151
1e−10	1.00000e−10	54613
1e−11	1.00000e−11	60074
1e−12	1.00000e−12	65535

High resolution, long test time

≈ 4 Hrs vs 1 min

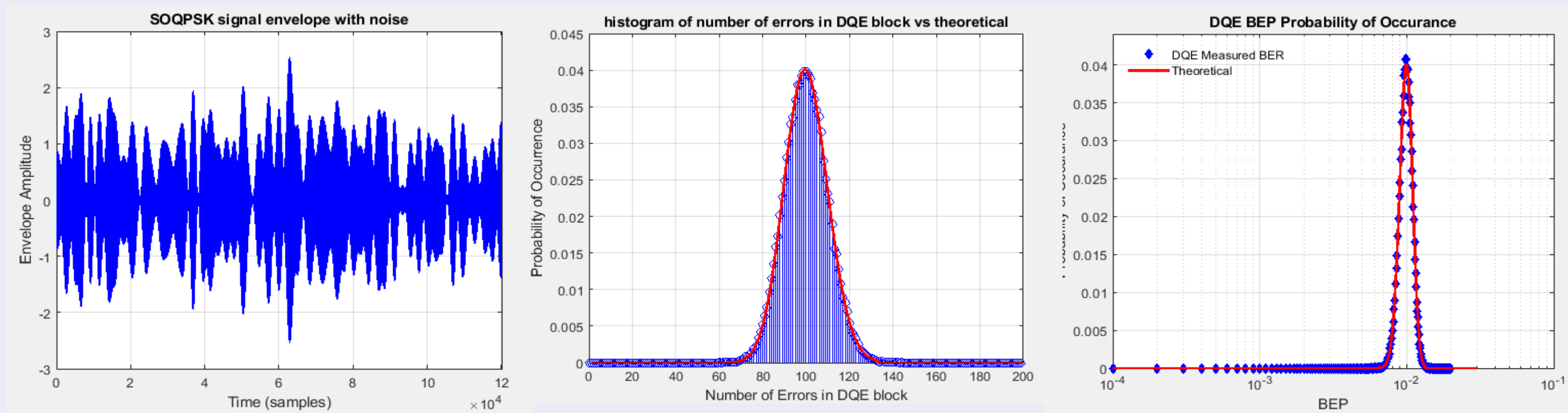
More reasonable resolution reduces measurement time by ≈ 256x

Ideal DQM distribution vs p Example (AWGN) n=10000

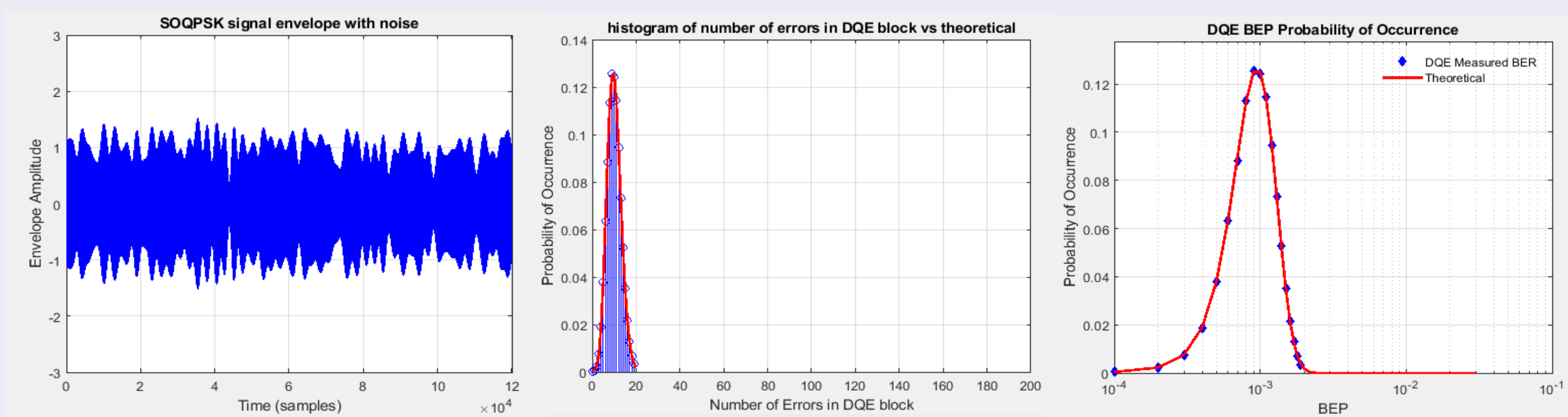
blocks of data are not guaranteed to occur so uniformly. Let E be the number of bit errors in a block of n bits. For a transmitted word of size n with independent bit errors, the probability of having fewer than or equal to t errors, $\Pr(E \le t)$ can be straightforwardly obtained from (6) as

$$\Pr(E \le t) = \sum_{i=0}^{t} \binom{n}{i} P_b{}^i (1 - P_b)^{n-i}. \qquad (7)$$

p = 1e-2



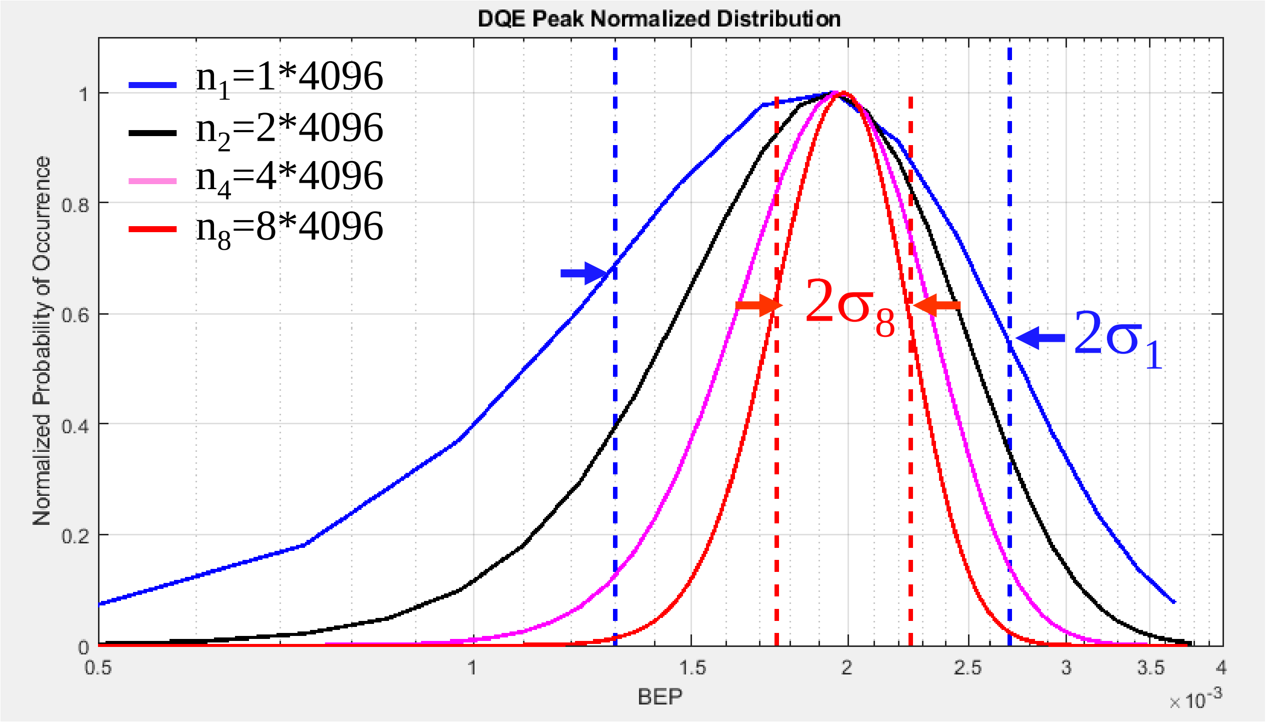
p = 1e-3



Distribution of ideal DQM values changes with p

Ideal DQM distribution vs Block Size n

- Example: n=4096, p = 2e^{-3}
- As span increases, DQM distribution narrows

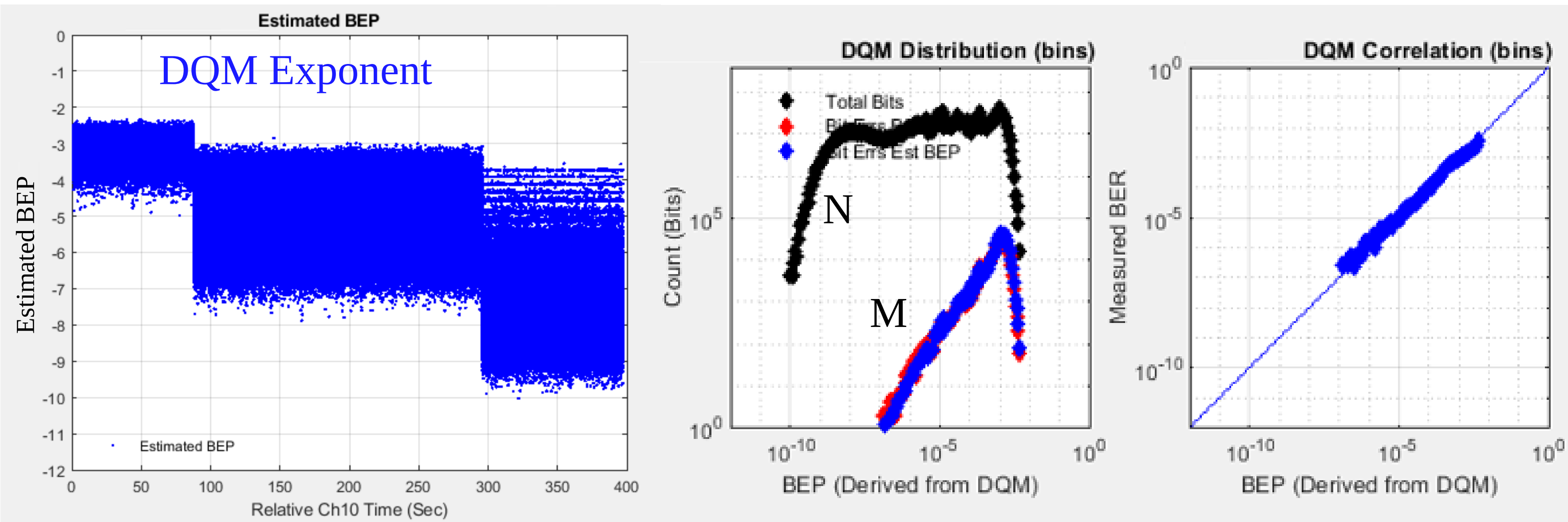


Distribution of ideal DQM values changes with block length n

Quasonix

Measured example from Industry Day Testing 10/28/22

4	ITC_5M_2022-301T16_05_35Z.ch10		B						Eb/N0 = 6dB
5	0	397		11.1 (AWGN sweep)	S-band	2225MHz	Tier0 (MS)	5Mbps (high rate)	Eb/N0 = 8dB
6									Eb/N0 = 10dB



Time Domain

DQM Distribution

DQM Correlation

Key Performance Plot!

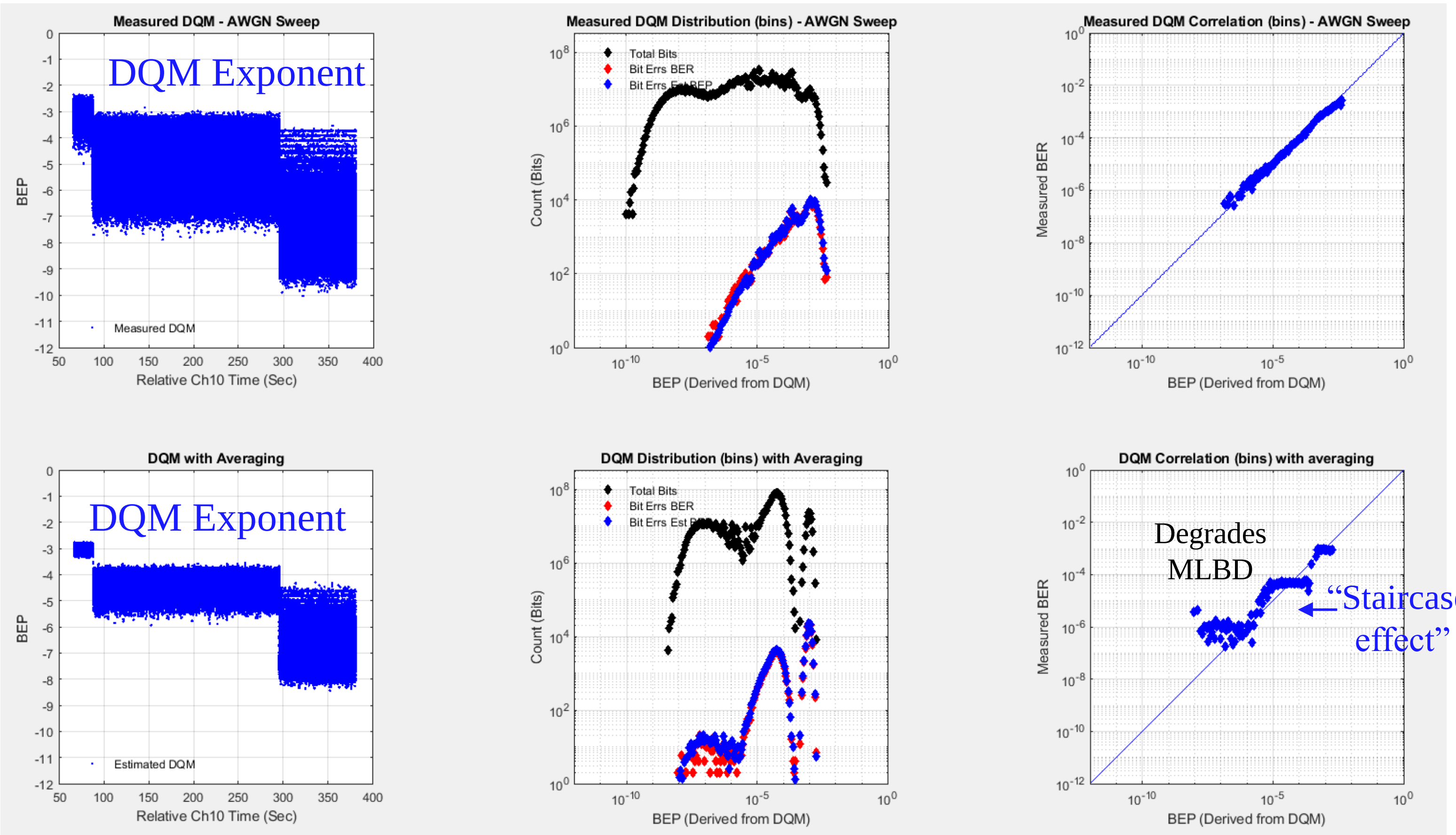
Quasonix

Effect of DQM "Averaging"

(Possibly using a parameter estimated at a rate much slower than DQE frames)

No Averaging

Averaging factor = 8



DQM averaging beyond one DQE packet degrades MLBD even in a static channel.
(Why? Result is based on a different DQM distribution than the MLBD is observing)

Quasonix

DQM Accuracy Conclusions

- Multiple receive channels can significantly improve the system BEP.
- The DQE/DQM RCC IRIG standard supports optimal processing for multi-channel telemetry reception.
- Accurate DQM values are a vital ingredient in achieving good performance.
- Common DQM estimation biases between receivers tend to cancel due to the differential nature of the metric calculations.
- Testing and verification methods are currently being developed to ensure vendor compatibility and consistent performance.

- ***It was shown that the performance degradation of the MLBD in error exponents is bounded by the estimation error exponent in DQM regardless of p's or N!***

MLBD Loss DQM error

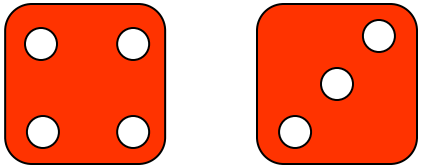
$$\Delta P_{e_{\exp}} \leq |\Delta p_{1_{\exp}}|$$

- ***Results from this paper can assist standards organizations in developing meaningful DQM testing thresholds.***

QUASONIX

Thank you !!! Questions?

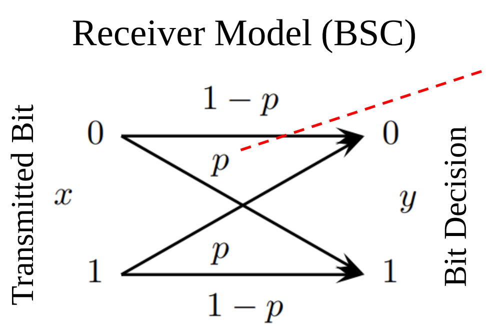
GOOD LUCK!



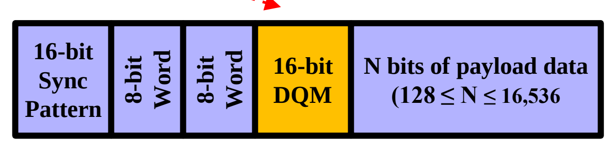
Backup Slides

2015 Hill ITC Paper – DQE/DQM protocol

- Diversity can provide tremendous improvements in link performance
- Problem: How to provide combiner with quality information from multiple sites?
- Proposed a protocol to insert timely quality information directly into the TM stream



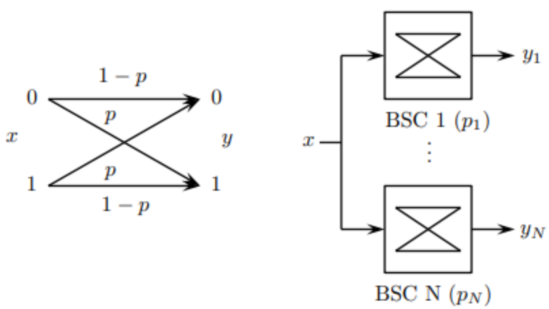
BEP Versus DQM		
BEP	**LR**	**DQM**
0.5	1.00	0
1e−01	1.11111e−01	5211
1e−02	1.01010e−02	10898
1e−03	1.00100e−03	16381
1e−04	1.00010e−04	21845
1e−05	1.00001e−05	27306
1e−06	1.00000e−06	32767
1e−07	1.00000e−07	38229
1e−08	1.00000e−08	43690
1e−09	1.00000e−09	49151
1e−10	1.00000e−10	54613
1e−11	1.00000e−11	60074
1e−12	1.00000e−12	65535



DQE Packet

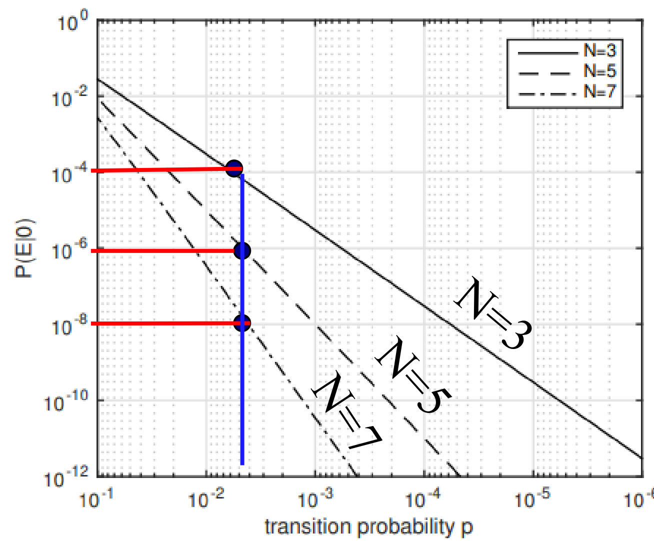
2015 ITC Rice/Perrins Paper – MLBD

- Analyzed the optimal bit detection strategy using multiple sources
- Proved that the transition probability p is a sufficient metric for optimum detection performance.
- Derived MLBD detection rule and probability of error.
- Best ITC paper of 2015



$$\hat{x} = 0 \iff \sum_{n \in \mathcal{N}_0} \log\left(\frac{1-p_n}{p_n}\right) > \sum_{n \in \mathcal{N}_1} \log\left(\frac{1-p_n}{p_n}\right)$$

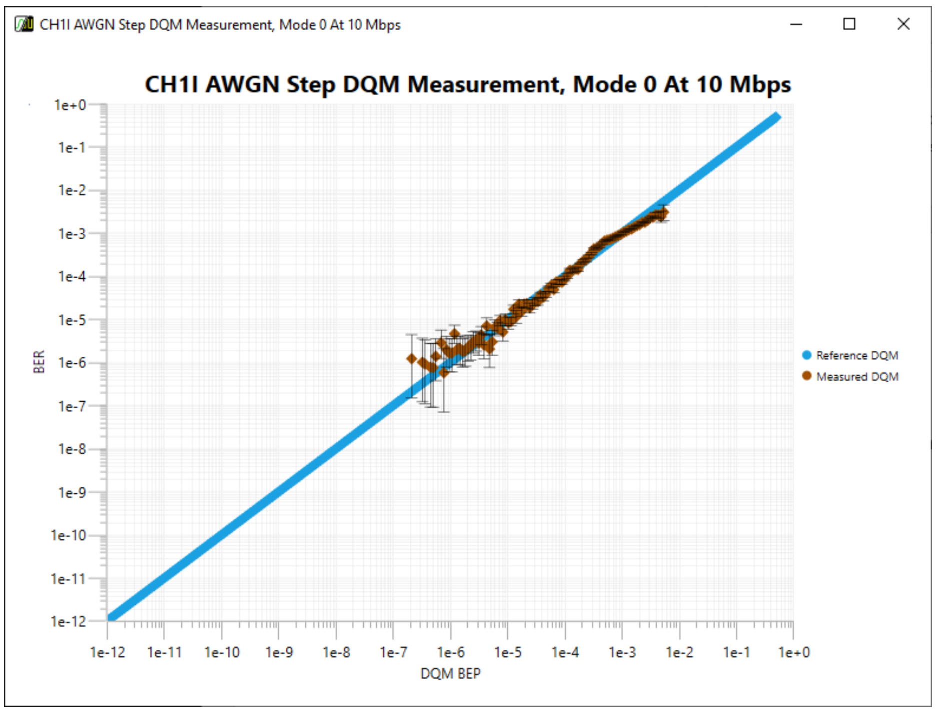
Equal Channels ($p_1=p_2=\ldots=p_N$)



DQM Correlation Plot (measured BER vs. estimated BEP)

- The y-axis is the measured BER for each point (labeled “BER”).
- The x-axis is the estimated BEP associated with each point (labeled “DQM BEP”)
- Ideal performance is when measured BER equals estimated BEP (blue line)

Measured BER

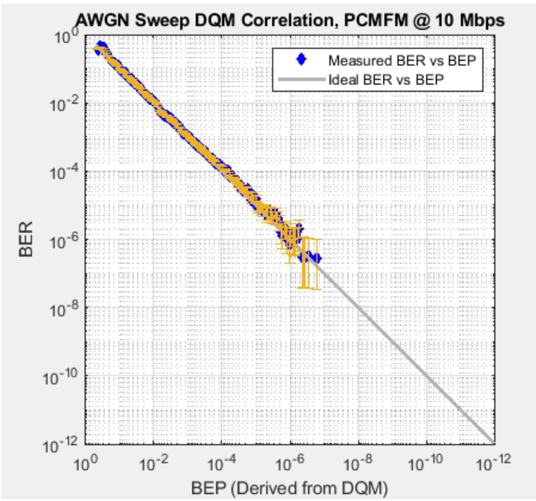


Estimated DQM BEP

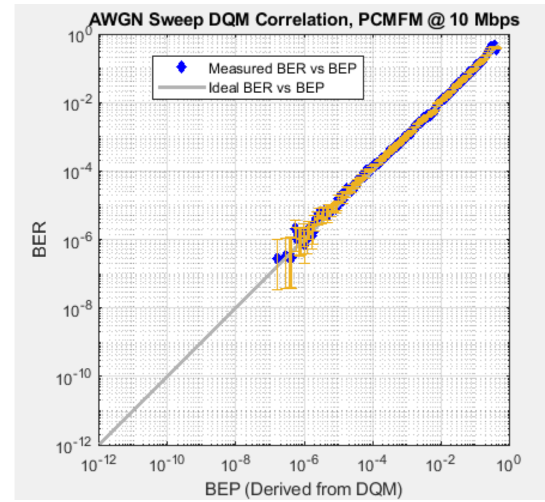
Key Performance Plot!

Quasonix

DQM Correlation Plot Representation



Equivalent Information



Representation used in the paper – looks more like a traditional BER curve

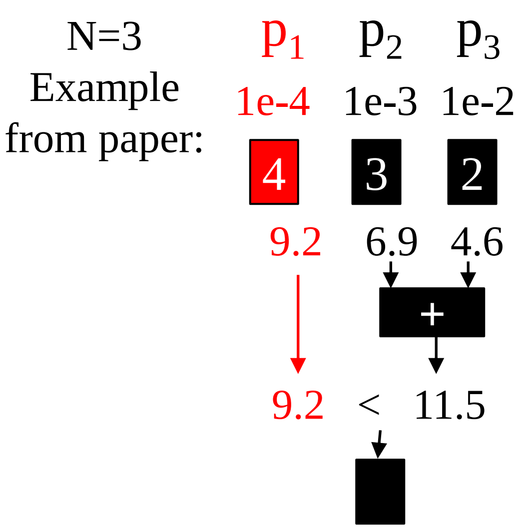
Conventional correlation plot with positive slope that clearly indicates direct or inverse relationship

MLBD Calculations

- Bit decision is based on comparison of LR sums of group that votes 0 versus 1

$$\sum_{n \in \mathcal{N}_0} \log\left(\frac{1-p_n}{p_n}\right) \gtrless \sum_{n \in \mathcal{N}_1} \log\left(\frac{1-p_n}{p_n}\right)$$

Bit Decision



Probability of Error

Assume x=0

y_1	y_2	y_3	probability	$\sum_{\mathcal{N}_0}$	$\sum_{\mathcal{N}_1}$
0	0	0	$(1-p_1)(1-p_2)(1-p_3) = 9.8891 \times 10^{-1}$	20.7	0
0	0	1	$(1-p_1)(1-p_2)p_3 = 9.9890 \times 10^{-3}$	16.1	4.6
0	1	0	$(1-p_1)p_2(1-p_3) = 9.8990 \times 10^{-4}$	13.8	6.9
0	1	1	$(1-p_1)p_2p_3 = 9.9990 \times 10^{-6}$	9.2	11.5
1	0	0	$p_1(1-p_2)(1-p_3) = 9.8901 \times 10^{-5}$	11.5	9.2
1	0	1	$p_1(1-p_2)p_3 = 9.9900 \times 10^{-7}$	6.9	13.8
1	1	0	$p_1p_2(1-p_3) = 9.9000 \times 10^{-8}$	4.6	16.1
1	1	1	$p_1p_2p_3 = 1.0000 \times 10^{-9}$	0.0	20.7

$$P(E|0) = 9.9990\times10^{-6} + 9.9900\times10^{-7} + 9.9000\times10^{-8} + 1.0000\times10^{-9} = 1.1098\times10^{-5}$$

$$P_e = \sum_{i \in \mathcal{I}_+} P[i]$$

P(E) = 1.1098e-5

MLBD probability of error

MLBD calculation with p estimation error

p_1	p_2	p_3
1e-4	1e-3	1e-2

Actual p_1,p_2,p_3

y_1	y_2	y_3	probability	$\sum_{\mathcal{N}_0}$	$\sum_{\mathcal{N}_1}$
0	0	0	$(1-p_1)(1-p_2)(1-p_3) = 9.8891 \times 10^{-1}$		
0	0	1	$(1-p_1)(1-p_2)p_3 = 9.9890 \times 10^{-3}$		
0	1	0	$(1-p_1)p_2(1-p_3) = 9.8990 \times 10^{-4}$		
0	1	1	$(1-p_1)p_2p_3 = 9.9990 \times 10^{-6}$		
1	0	0	$p_1(1-p_2)(1-p_3) = 9.8901 \times 10^{-5}$		
1	0	1	$p_1(1-p_2)p_3 = 9.9900 \times 10^{-7}$		
1	1	0	$p_1p_2(1-p_3) = 9.9000 \times 10^{-8}$		
1	1	1	$p_1p_2p_3 = 1.0000 \times 10^{-9}$		

Probability of (y_1,y_2,y_3) symbols are based on the actual channel transition probabilities

Estimated $\hat{p}_1,\hat{p}_2,\hat{p}_3$

$$log\left(\frac{1-\hat{p}_1}{\hat{p}_1}\right),\ log\left(\frac{1-\hat{p}_2}{\hat{p}_2}\right),\ log\left(\frac{1-\hat{p}_3}{\hat{p}_3}\right)$$

$$\sum_{n\epsilon \mathrm{N}_0[i]} \log\left(\frac{1-\hat{p}_n}{\hat{p}_n}\right) \quad \sum_{n\epsilon \mathrm{N}_1[i]} \log\left(\frac{1-\hat{p}_n}{\hat{p}_n}\right)$$

Weighting of sources is based on *estimated* channel error probabilities from receiver

$$\hat{P}_e = \sum_{i\epsilon\hat{\mathcal{I}}_+} P[i]$$

MLBD probability of error with $\hat{p}$ estimate

$$\log\left(\frac{1-\hat{p}_n p_{bias}}{\hat{p}_n p_{bias}}\right) \xrightarrow{\hat{p}_n p_{bias} \ll 0} -\log(\hat{p}_n) - \log(p_{bias})$$

Systematic estimation bias offsets all terms

Can this bound be shown to apply for all N and p tuples?

- When all else fails, dig into the math…
- What does it take to degrade P(E)?
  - Answer: Adjust p_1 to flip the smallest positive M[i] negative

N=3, $p_1 = 10^{-4}$, $p_2 = 10^{-3}$, $p_3 = 10^{-2}$					Ideal Metrics ($p_1 = 10^{-4}$)			Estimated Metrics ($\hat{p}_1 = 10^{-5}$)		
i	y1	y2	y3	Probability P[i]	$\sum_{N_0}$	$\sum_{N_1}$	$M[i]$	$\sum_{N_0}$	$\sum_{N_1}$	$\hat{M}[i]$
0	0	0	0	9.8891 x 10^{-1}	20.7	0	-20.7	23.0	0	-23.0
1	0	0	1	9.9890 x 10^{-3}	16.1	4.6	-11.5	18.4	4.6	-13.8
2	0	1	0	9.8990 x 10^{-4}	13.8	6.9	-6.9	16.1	6.9	-9.2
3	0	1	1	* 9.9991 x 10^{-6}	9.2	11.5	* **2.3**	11.51	11.50	-0.01
4	1	0	0	9.8901 x 10^{-5} *	11.5	9.2	-2.3	11.50	11.51	**0.01** *
5	1	0	1	* 9.9900 x 10^{-7} *	6.9	13.8	* **6.9**	6.9	16.1	**9.2** *
6	1	1	0	* 9.9000 x 10^{-8} *	4.6	16.1	* **11.5**	4.6	18.4	**13.8** *
7	1	1	1	* 1.0000 x 10^{-9} *	0	20.7	* **20.7**	0	23.0	**23.0** *
P[i]'s do not change due to estimation					P(E | 0)=1.1098 x 10^{-5}			$\hat{P}(E \mid 0)$=1.0000 x 10^{-4}		