Eye mask statistical significance and practicality

Piers Dawe
Avago Technologies

Supporters

John Petrilla	Avago Technologies
Tom Lindsay	ClariPhy
Kenneth Jackson	Emcore
Eddie Tsumura	ExceLight
Robert Coenen	Finisar
Jonathan King	Finisar
Mike Dudek	JDSU
Matt Traverso	Opnext

Problem statement

- Mask measurement is not fully defined
  - References used by 802.3 tell much of the story
  - The remainder is chosen by "custom and practice"
- Most measurements have poor reproducibility
- Most measurements do not correlate to useful performance
- Mask measurement can be more onerous than TDP*
  - Contrary to intention of 802.3ae
- Attempts to improve reproducibility lead to excessive test time
- Going to 802.3ae, large increase in measurement bandwidth and change from short block code to scrambled line code
  - Scope traces are less distinct than at lower speeds
  - Scope noise could be RSSd out but this is not well documented and adds measurement complexity
    - Measurement complexity = cost
  - In contrast to good/excellent robustness of 10G links
- Need to document what mask compliance means more completely

** See last slide for abbreviations*

Motivation

- Un-retimed crosstalk in SFP+ means transmitted eyes can be noisier than XENPAK, XFP
  - While still meeting TDP etc.
  - Opportunities for counter-propagating crosstalk
    - At IC in host
    - At optical module connector
    - Within module
  - They may "look fuzzier" but still deliver good links
- Motivated to plug the gaps in 52.9.7 to support SFP+ implementations of 802.3ae
- 4 and 10 lane 802.3ba links need much cheaper testing, expect SFP+ like or worse crosstalk
  - Opportunities for co-propagating and counter-propagating crosstalk
  - With 10G lanes or 25G lanes
- 10GEPON will also require well cost-optimised testing
- Motivated to plug the gaps in 52.9.7 for 10GEPON and 802.3ba
- Also wanted to compare product Tx spec (eye mask) with stressed eye generator spec
  - Analysis validates the 10^-3 vertical histogram in 52.9.9

Present situation: Clause 52

- When we wrote 52.9.7 we asked how mask dimensions were established and how the measurements were carried out. We were told that the dimensions were chosen by judgement, and the industry standard for measurements was 'zero hits in 200 waveforms' (waveforms not well specified).
  - Although the above is common belief, I have not found a documentation of this industry practice, aside from 802.3 clauses 58 and 68.
- This measurement has poor reproducibility (see later). Because of the huge noise bandwidth at 10 Gb/s, acceptable noise processes in the transmitter under test and in the oscilloscope cause scatter on the measured eye margin
  - This problem has been worse than expected and remains an obstacle to cost reduction for 10 Gb/s and its aspirations to high volume
- It is common to repeat the measurement a few times - but this is time consuming (adds cost).
- Another approach is to continue the measurement for an increased number of waveforms. This biases the result towards the pessimistic, attempts to measure a Tx noise that is drowned by Rx noise and obscured by oscilloscope noise and anyway, **does not cure the reproducibility issue**, and of course takes longer.
- We chose our mask by judgement, intending that **the mask criterion would be easier than the more thorough and representative TDP** specification. It turns out that unless the hit ratio is chosen wisely, the mask criterion becomes more demanding than TDP, failing interoperable transmitters and adding to requirements of cost, SERDES jitter, and thermals.
- For 10GBASE-L, the mask criterion IS more demanding than TDP
  - The opposite of what we intended when we wrote Clause 52

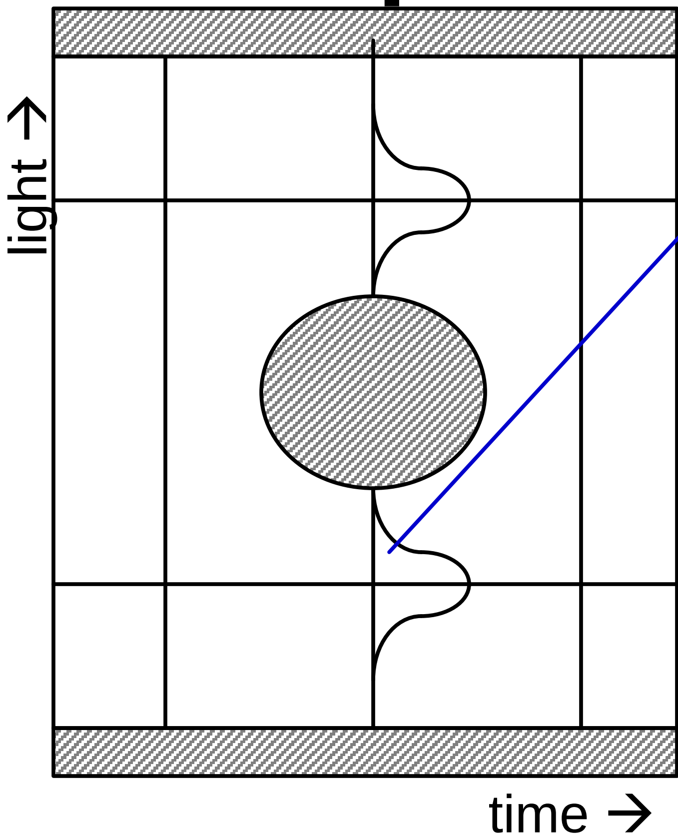
Present situation: references

- OFSTP-4 or IEC 61280-2-2
  - There was an example eye and mask in OFSTP-4 or IEC 61280-2-2
  - "200 waveforms" is taken mean 200 sweeps across the time window shown on a sampling scope synchronised to the symbol repetition rate
  - The screen might show 1.2 to 1.6 UI. There is an upper mask region, a lower mask region, and a central mask region.
    - The upper and lower regions are to constrain overshoot, a deterministic process
      - No need to be fussy about measuring the tail of the distribution for these
  - As I understand it, digital storage scopes of the vintage of OFSTP-4 referred to in 802.3ae development (OFSTP-4 vintage) and the IEC version of the same, had 451 sample points across the screen
  - So "zero hits in 200 waveforms" means that 1 UI was sampled about 451 * 200 / 1.5 = 6×10^4 times, and (perhaps after a repeat measurement) no samples fell in the masked regions. Implies that if the experiment were repeated multiple times, more often than not there would be no hits. An implied hit ratio of ~10^-5.
- EFM and 10GBASE-LRM
  - These define mask compliance at a defined hit ratio, 5×10^-5. The ratio was chosen to have the mask margin give the best feasible correlation to transmitter penalty (which isn't perfect, but can be made even worse by alternative choices).
  - See 58.7.8 and 68.6.5
  - See http://ieee802.org/3/efm/public/may03/optics/dawe_optics_2_0503.pdf
- Fibre Channel
  - 1/2/4GFC defines eyes at 10^-12
  - 8GFC optical eyes at 10^-3

"Zero hits" is bad

- A "zero hits" measurement has a measurement scatter (sigma) of ~0.4 times the noise of the combination of the signal under test and the oscilloscope.(see next several slides)
  - 10G optical signals might have noise due to RIN at -128 dB/Hz, which represents a Qsq of 28.3, although not all signals with this noise would pass TDP
    - Qsq is (OMA/2/(signal sigma). See 68.6.7.
  - For 10G optical signals, scope noise is significant
    - A 10G oscilloscope in a development lab might give a Qsq of 100
    - In production, optical switches attenuate the signal under test and degrade the observed SNR. Qsq from scope and switches might be 50
    - These are radically worse than an electrical scope viewing an electrical signal
  - We see mask performance quoted in "%": each 1% is OMA/400 because the mask is half the height of the eye and splits the passing region in two
- Quoted performance has considerable scatter
- Every vendor has to supply extra performance to cover his fear of his own and his customer's measurement irreproducibility
  - This is not compatible with aspirations of low cost high volume 10GE, and any 40G/100G
- **Increasing the number of samples in the measurement does not help reproducibility**
  - But it does make the measurement more pessimistic
  - But not consistently
- As with any experiment, one can repeat the measurement many times and average out some irreproducibility
  - Slow, clumsy
  - e.g. 10^6 samples at 40 ksamples/s/channel takes 25 s/channel; not low cost
- Rather than repeat a measurement 20 times looking for at least 10 trials with no hits, we can just let one measurement run for 20 times as long and look for less than 10 hits.
  - Works as well, neater, still consumes measurement time
- This approach improves measurement reproducibility by square root of the number of hits allowed
  - As long as the total number of samples is reasonable (>100)

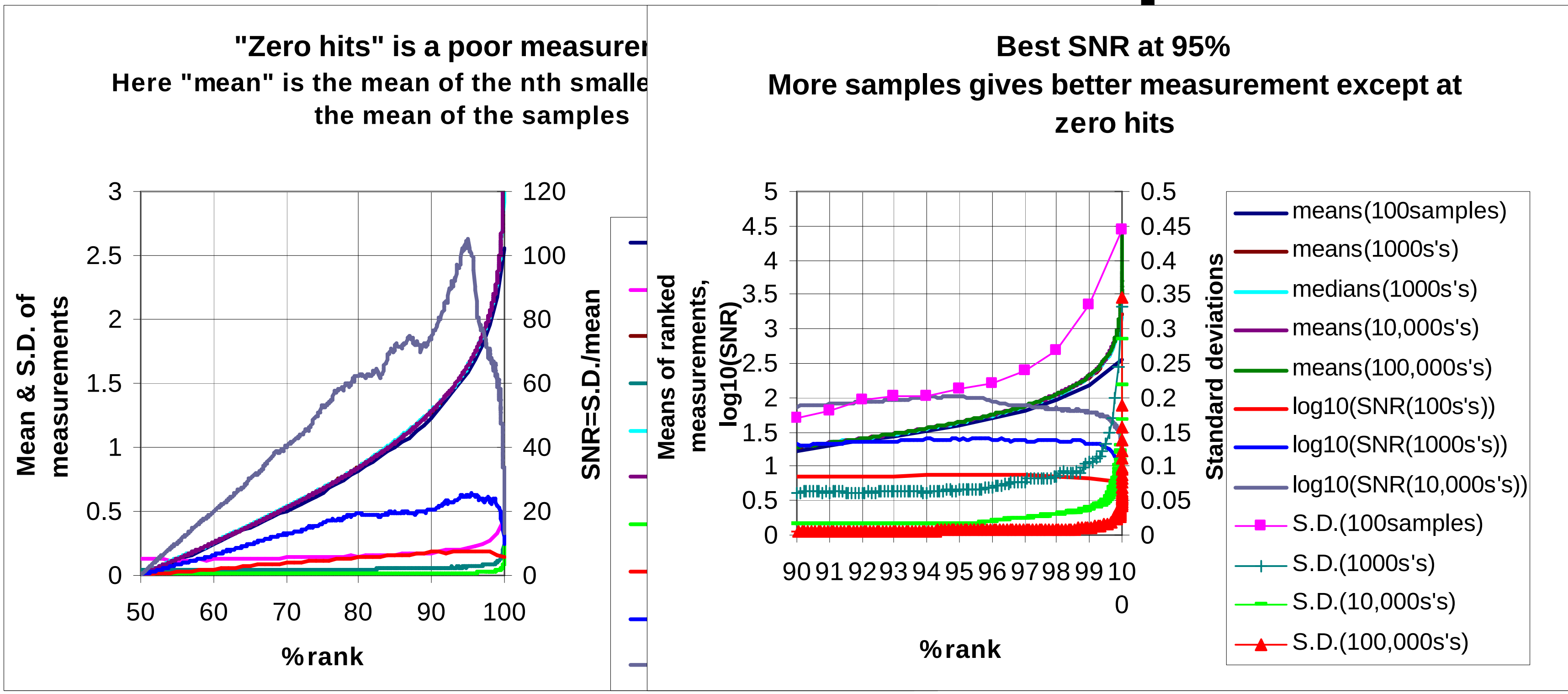
Simple statistical experiment



Oscilloscope eye mask

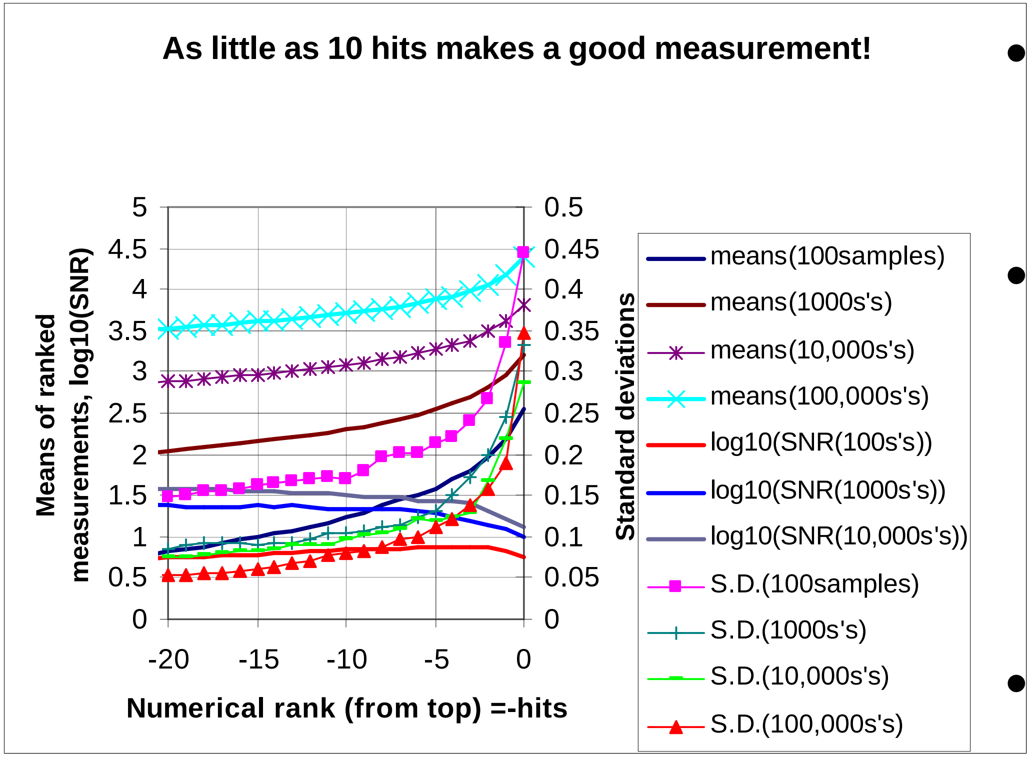
- Representing a mask measurement
- Take *n* samples from a Normal distribution, mean 0, standard deviation 1
- Find the highest sample, next highest, ...
- Repeat for 100 trials
- Find the mean of the highest samples, of the next highest...
  - (not the same mean as above)
- Find the standard deviations (SD) of the highest samples, of the next highest...
- Define "Measurement SNR" (visibility?) as mean/SD
- Repeat whole experiment with another sample size *n*

Results of statistical experiment



- The means follow the characteristic error function curve
- "Measurement SNR" increases with sample size as expected (standard deviation reduces)
- "Measurement SNR" peaks around 95%, independent of sample size, falls off steeply at highest rank

Results of statistical experiment



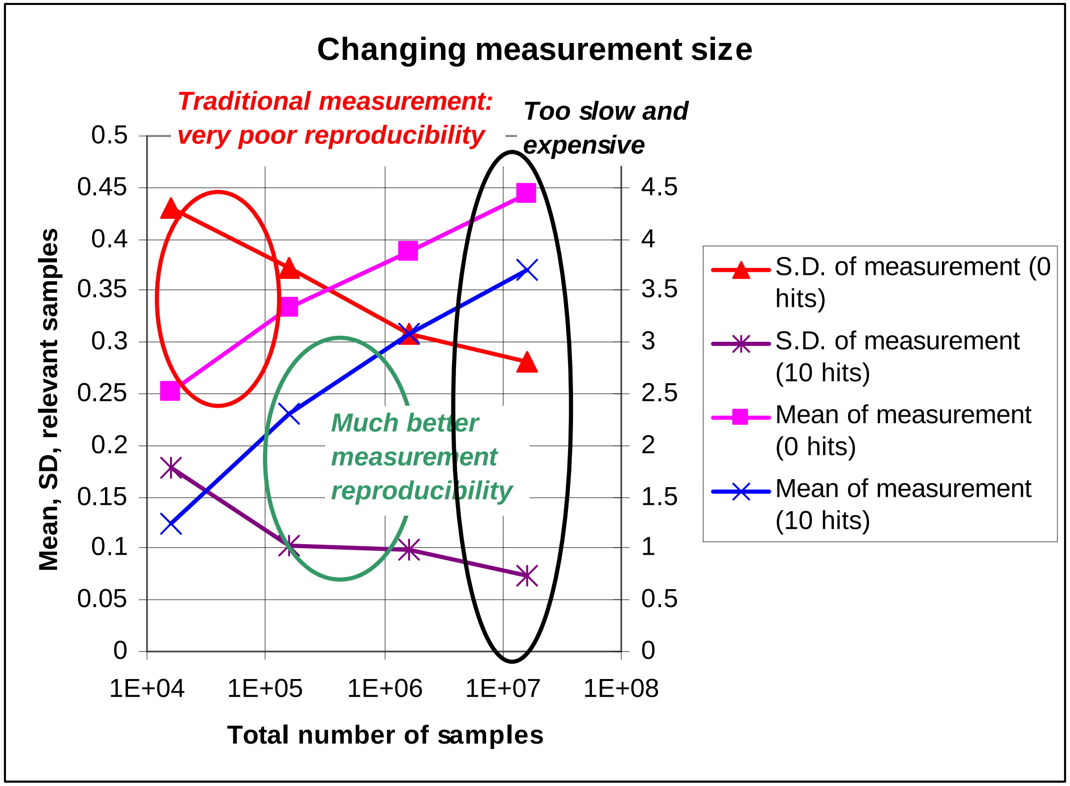
- The scatter on the highest sample in each trial is remarkable
- The scatter on the 10th highest sample is typically 1/3 as much
  - Even e.g. the 5th highest sample is much more reproducible
- A bigger sample size does NOT significantly reduce the scatter on the highest sample, but does reduce the scatter on less extreme ranks

- A bigger sample size explores "more sigmas" of the distribution
- How many sigmas is relevant? It's not 7 for TP2 mask test! It's 2 to 5 depending on circumstance

How much is one sample?

- For eye mask measurements
- The oscilloscope samples evenly in time e.g. from –0.7 UI to +0.7 UI
- Typical rate 40 ksamples/s
- Assuming the waveform is nearer the inner polygon than the outer,
- Samples are nearest to the inner polygon for only ~ 1/20 UI
  - (as I estimated. May be larger, up to 1/5)
- A marginal transmitter will have patterning:
  - If simply slow, 101 and 010 sequences are worst
  - For direct modulated transmitter, these aren't equivalent
  - For SRS tester, they should be, but trajectories can be raised or lowered by SI, early or late by SJ
- Hence only 1 in 8 of (1/20 to 1/5) UI are samples relevant to the test
  - (1 in 4 of ... for SRS tester because eye should be symmetrical)

Results of statistical experiment



- If 1/8 of (1/20 to 1/5) of total samples are relevant
- Measurement times up to 0.4 to 400 (!) s
- Traditional "200 waveforms" at ~451 points/screen ~60,000 samples in 1 UI: approx. 3 sigma
  - Not 4.3 sigma because only 1 in 8 in (5 to 20) samples would fail anyway

- One measurement of 10n samples, 10 hits is about the same as averaging 10 measurements of n samples, 0 hits
- **Recommend 10 samples in 1E5 or 2E5, or hit ratio of 1E-4 or 5E-5. But see later...**

Measurement times (per lane)

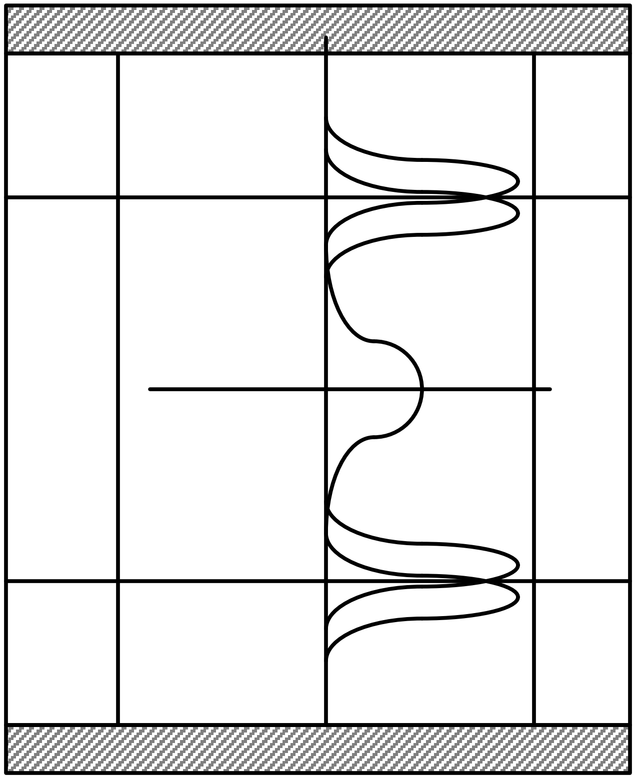
Total samples	Relevant samples	Measure-ment time (s)	"sigmas"
16000	100	0.4	1.3
2E+05	1000	4	2.3
2E+06	10000	40	3.1
2E+07	1E+05	400	3.7

- ***1 million samples is onerous. Spec should enable 1e5 samples or less***
- 2 to 2.5 "sigmas" is practical

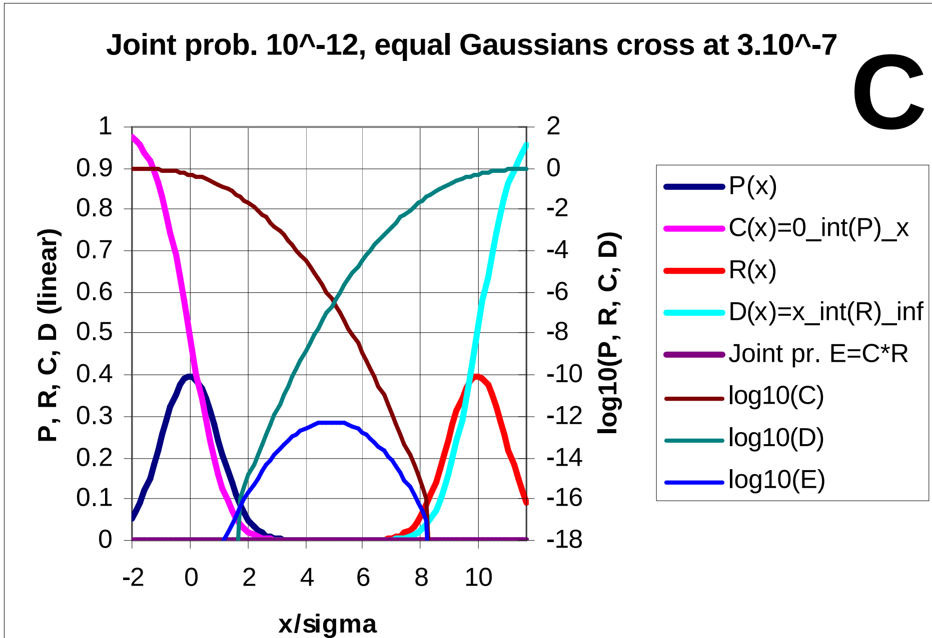
How appropriate are these hit ratios?

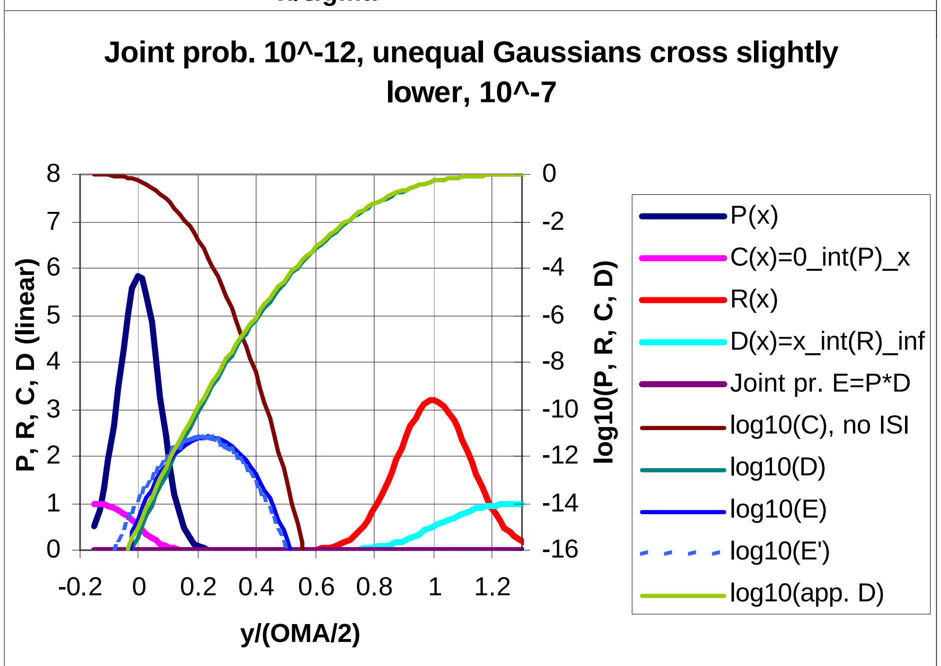
- Consider dual-Dirac distribution
  - Mean of the two Dirac functions is at 0 or 1 on the eye diagram
  - When the received signal convolved with the receiver noise gives tails of the distribution crossing the middle of the eye >10^-12, we have a bad link
- We know the spec TDP, so by definition we know the receiver noise
- As spec TDP ~ 3 dB, receiver noise tails fill ½ the eye
  - This is much more than the transmitter noise
- No point looking for the extreme tail of the transmitter's noise
  1. It will be drowned by the convolution of receiver noise and the more-probable parts of the transmitter's PDF
  2. It cannot be found directly by a scope because of scope noise

...How appropriate are these hit ratios?



- Consider 10GBASE-L first
  - For traditional DFBs, dispersion penalty is negligible (can be slightly negative)
  - If transmitter noise is low, the skirts are steep
    - We can define the mask by any level of statistical significance at which the measurement is not compromised by scope noise, OR define a method to RSS out the scope noise: obvious in principle
  - If transmitter noise is high and transmitter is marginal, deterministic ISI must tend to 0
    - But this case would fail RIN and TDP
    - The following pages show two simple overlap-integral examples then a reasonable range of 10GBASE-L cases

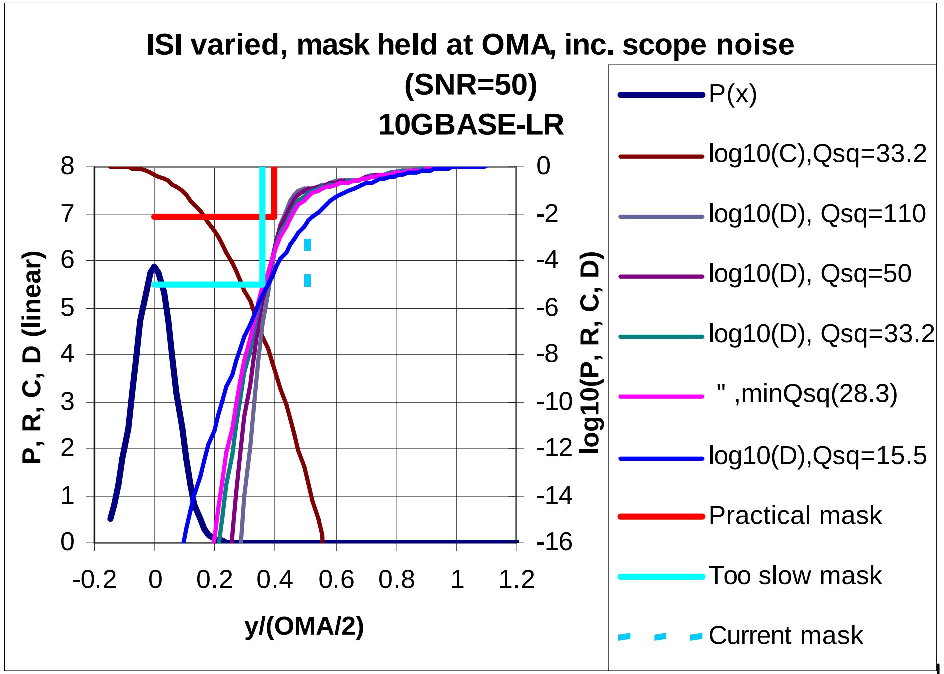




CDFs cross ~10^-7

- If two Gaussians overlap for 10^-12 BER, the CDFs overlap at ~10^-7
- First example: equal sigma, horizontal scale of sigma
- Second example: red distribution on left is receiver set to discriminate at 3.2 dB TDP, horizontal scale 0 to 1
- But, 10^-7 isn't the right criterion either...

One-dimensional analysis



- "8-Dirac" transmitters, all with 3.2 dB TP
- The waterfall lines on the left (log scale) cross each other, but
- Imperfect "focus point" around 10^-4 for transmitters with less noise than RIN max
- Remember the mask test is meant to be permissive: unjustified fails are bad, unjustified passes are OK because TDP spec will catch those transmitters
- Mask "1" can remain near OMA "1" even when ISI is present because of laser overshoot
- The grey line would be vertical without scope noise

- The 802.3 10G mask (Y1=0.25, equivalent to 0.5 on this scale) is simply too tall for 3.2 dB TP, whatever the level of statistical significance
- These probabilities still have to be divided by 5 to 20 (test time multiplied by 5 to 20) because hits occur first in a narrow time-window on the UI scale ...

...but two-dimensional analysis leads to a different conclusion

Next steps

- More study on the factor of 5 to 20
  - Needs a two-dimensional analysis of mask hits – under way
- Depending on above, define a hit ratio in the range 10^-4 to 10^-3 for Clause 52 TP2
- Revise the mask coordinates Y1, Y2 for 10GBASE-L only
  - from 0.25, 0.28, to 0.30, 0.33
  - Compare 8GFC optical Tx mask 0.32, 0.35
- Raise an 802.3 maintenance request for 10GE
- Provide recommendations for P802.3ba

- (And for SFP+
  - Investigate if can apply same analysis to SFP+ jitter mask at B
    - expect an intermediate hit ratio would be appropriate
  - Investigate if can apply same analysis to SFP+ jitter mask at C
    - expect a hit ratio much nearer 10^-12 would be appropriate)

Abbreviations

CDF	cumulative distribution function
DFB	distributed feedback (laser)
ISI	inter-symbol interference
OMA	Optical Modulation Amplitude
PDF	probability distribution function
RIN	relative intensity noise
RSS	root sum of squares
SFP+, XENPAK, XFP	Fiber optic module form factors
SI	sinusoidal interference
SJ	sinusoidal jitter
SRS	stressed receiver sensitivity
TDP	transmitter and dispersion penalty
UI	unit interval