MLSE Channel Estimation Parametric or Non-Parametric?

Optical Fiber Communication Conference 2008

Parametric versus Non-Parametric Branch Metrics for MLSE–based Receivers with ADC and Clock Recovery

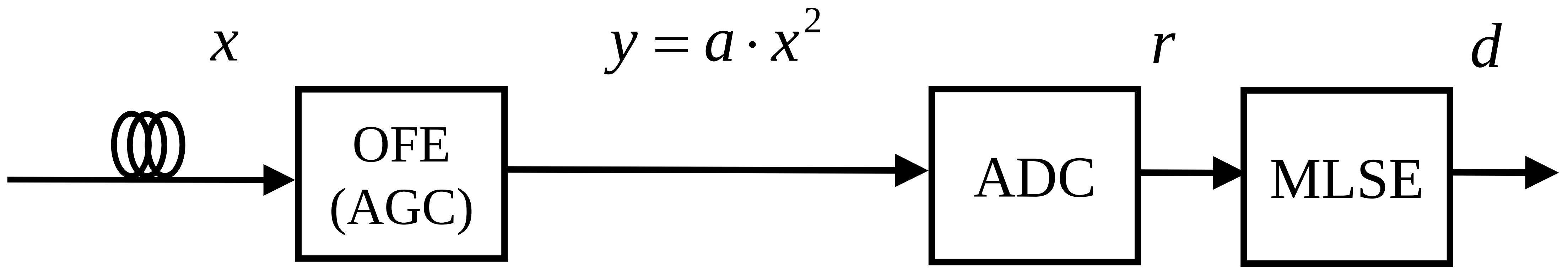
Stefan Langenbach [(1)], Gabriella Bosco [(2)], Pierluigi Poggiolini [(2)], Theo Kupfer [(1)]

1) CoreOptics, Nordostpark 12-14, D-90411 Nürnberg, Germany – {stefan,theo}@coreoptics.com

2) Politecnico di Torino, Corso Duca Abruzzi 24, Torino, Italy – {gabriella.bosco,pierluigi.poggiolini}@polito.it

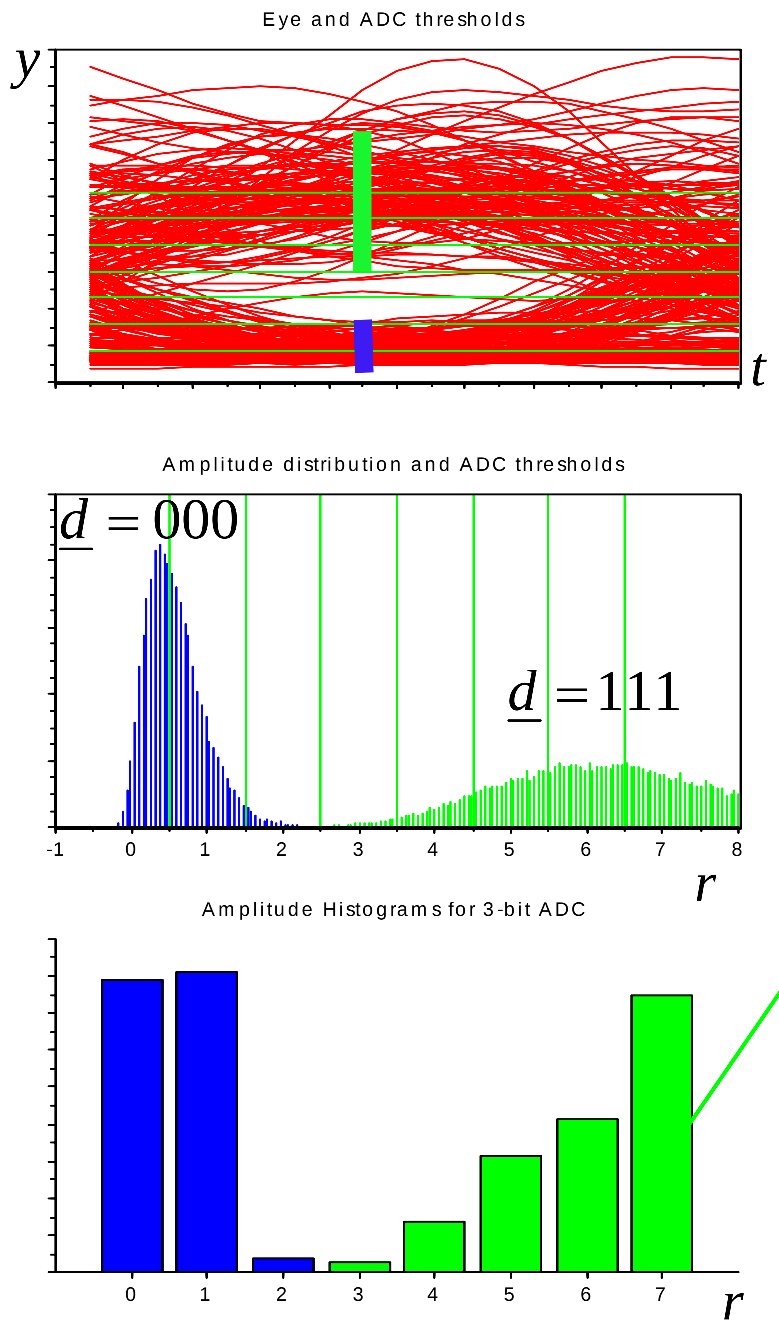
Part of this work was sponsored by the EU Integrated Project NOBEL II, WP5

Non-Parametric: Canonical Histogram Method – HM



Non-parametric Method

- Determine detected bit pattern d
- Associate observed quantized amplitude r
- Count observed amplitudes r for pattern d into a histogram N(d, r)
- Metrics for given d and r is ~ log N(d, r)



Pros

- Simple – just counting events
- Robust – insensitive to model mismatch

Cons

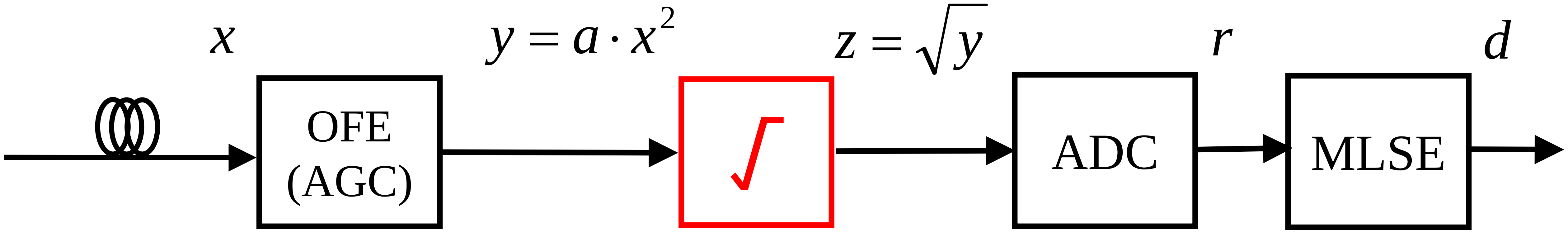
- Data collection time ~ ADC resolution
- Number of counters ~ ADC resolution
- Possibly more sensitive to error propagation? (decision errors translate into metrics errors)

Histogram based channel model
No parameters are estimated. But a full amplitude histogram needs to be „measured“ for each bit pattern.

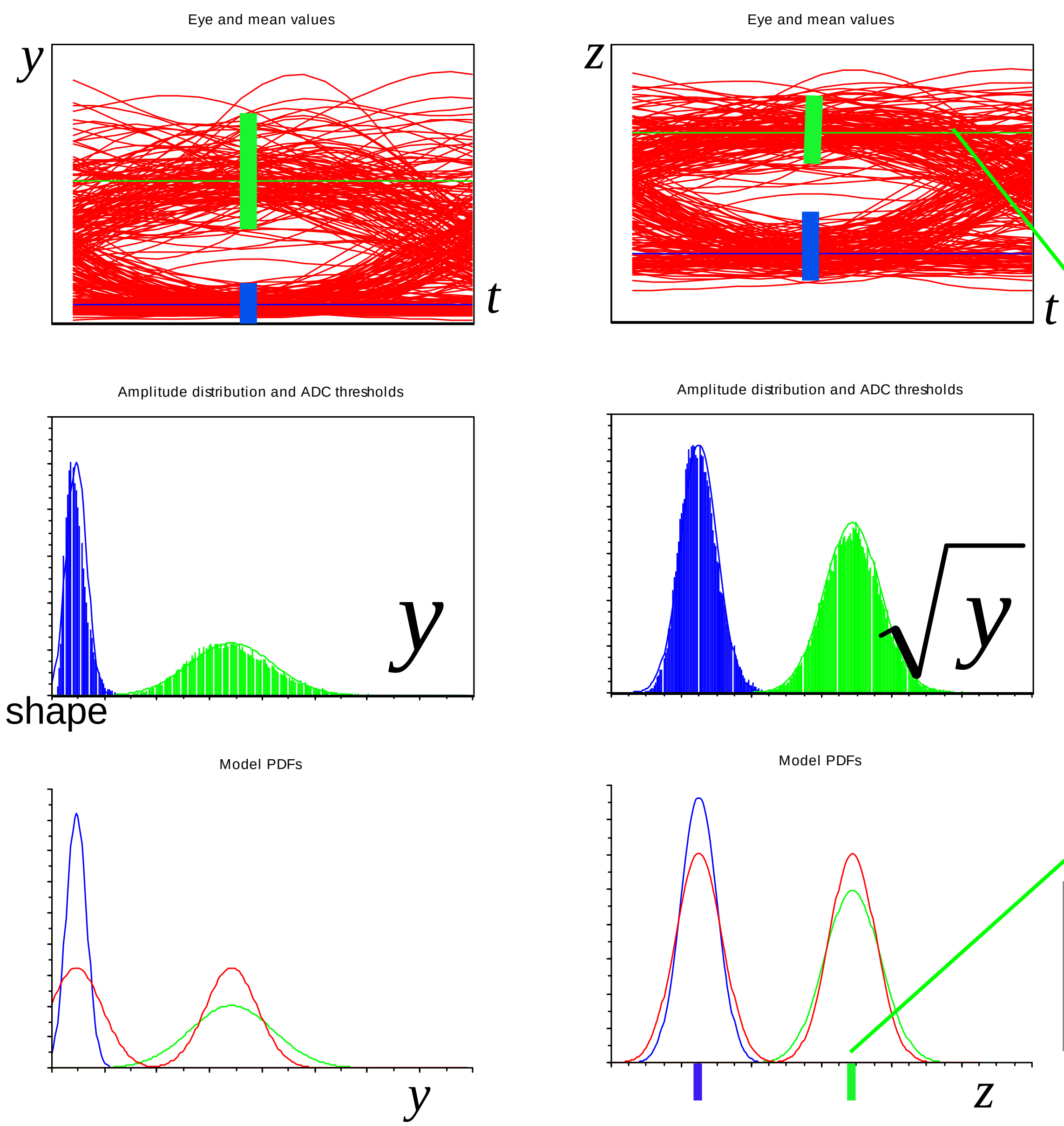
Canonical metrics
Metrics for given r is the logarithm of the observed relative frequency value.

Non-Parametric

Parametric: Square Root Method – SQRT



y: signal dependent noise
The electrical signal has more noise on ones than on zeros.



z: signal independent noise
The sqrt'ed electrical signal has roughly Gaussian noise and roughly signal independent noise.

Parametric Method

- Take square-root z of signal y
- Determine detected bit pattern d
- Determine mean z-amplitude m(d) for pattern d
- Metrics for given d and z is ~ (z - m(d)) ²

Pros

- Simple – just one parameter per PDF
- Fast – mean value can be estimated quickly
- Robust – decision errors do not corrupt PDF shape

Cons

- Model mismatch penalties are possible
- DC coupling

Mean based channel model
Only the mean value for each bit pattern needs to be estimated when signal independent noise is postulated (i.e. when the red PDFs are used)

Euclidean metrics
Metrics for given z is then the Euclidean distance from the mean sqrt'ed signal.

Note: The square root operation can also be applied implicitly by a non-uniform ADC, or explicitly after the uniform ADC, the latter with minor performance degradation

Possible Problems of Parametric Estimation

Method-Independent Problems

ISI Overload	Channel Memory exceeds State Memory
Quantization	Amplitude differences become invisible

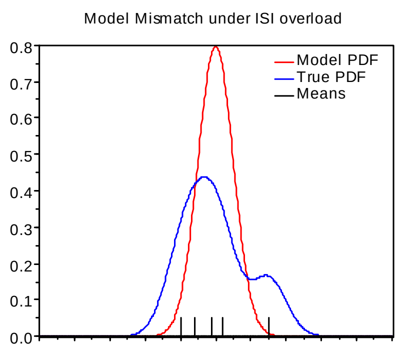
Some problems are specific for a parametric approach. All can lead to „wrong metrics“. In short:

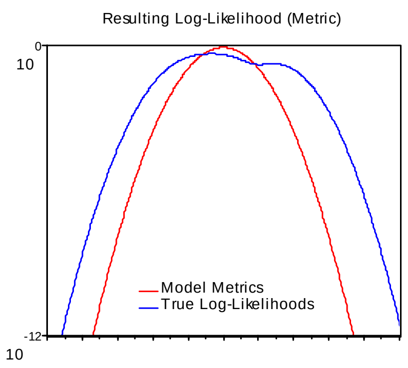
Model Mismatch

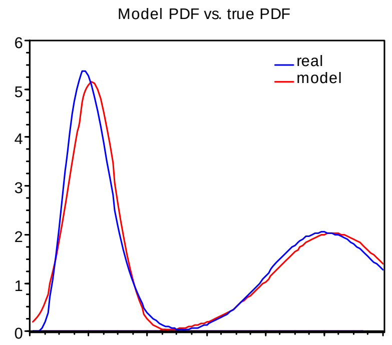
Unrealistic Noise Model → wrong PDF shape
When noise model does not sufficiently accurately model the „true“ noise PDF, metrics errors are introduced.

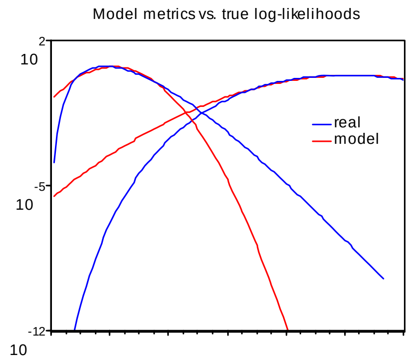
ISI Overload → wrong PDF shape and parameters
Using a model density instead of the mixture density.

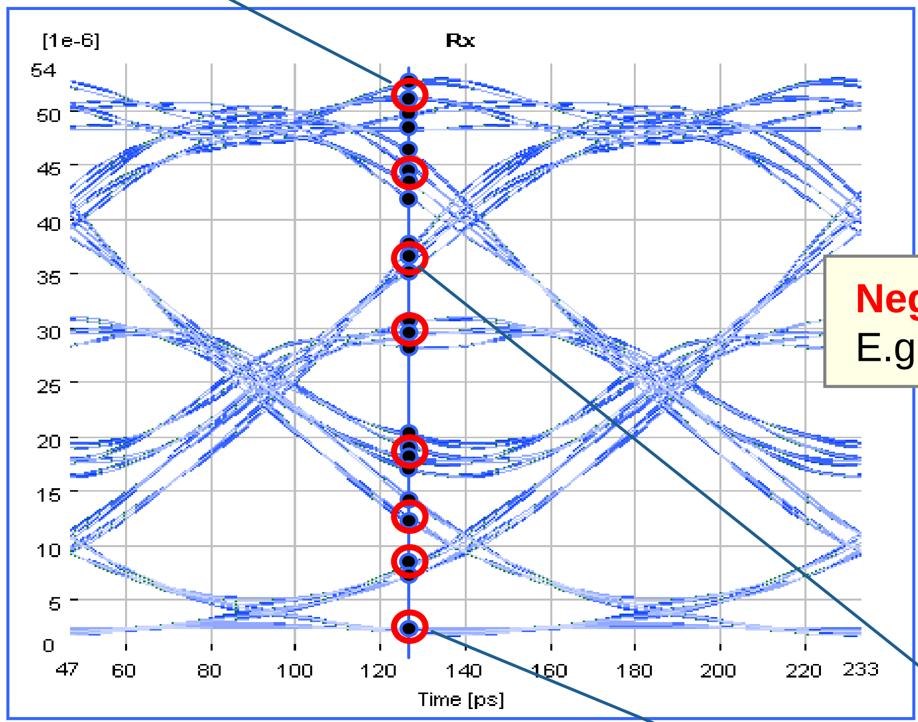
Note: Noise PDF for coarse state is a *mixture* density (i.e. a convex combination of several PDFs)





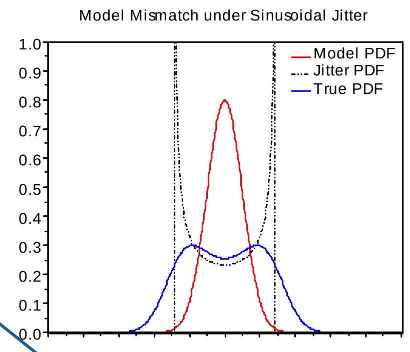


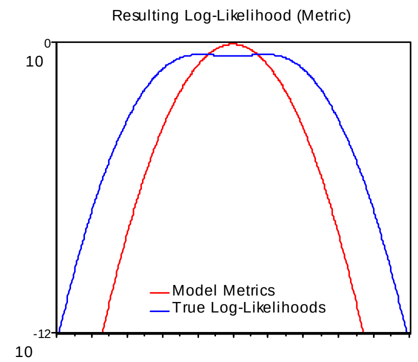




1500 ps/nm 0.6UI NRZ

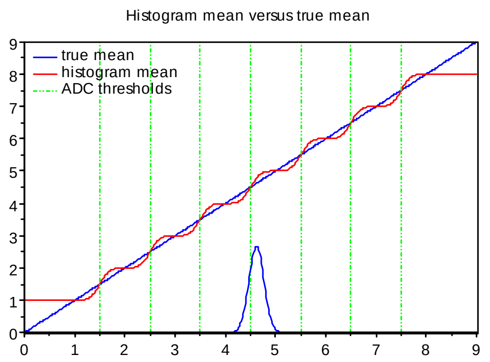
Neglected Effects → wrong PDF shape or parameters
E.g. imperfect clock (here: sinusoidal jitter)



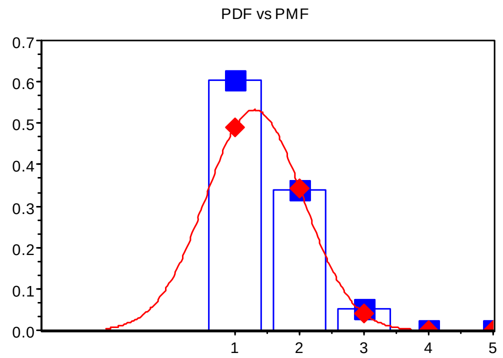


Note: Jitter impact on PDF depends on slope and is therefore pattern-dependent:
- Strong impact on edges.
- Little impact on rails

Quantization → wrong PDF parameters
Mean quantized ≠ true mean

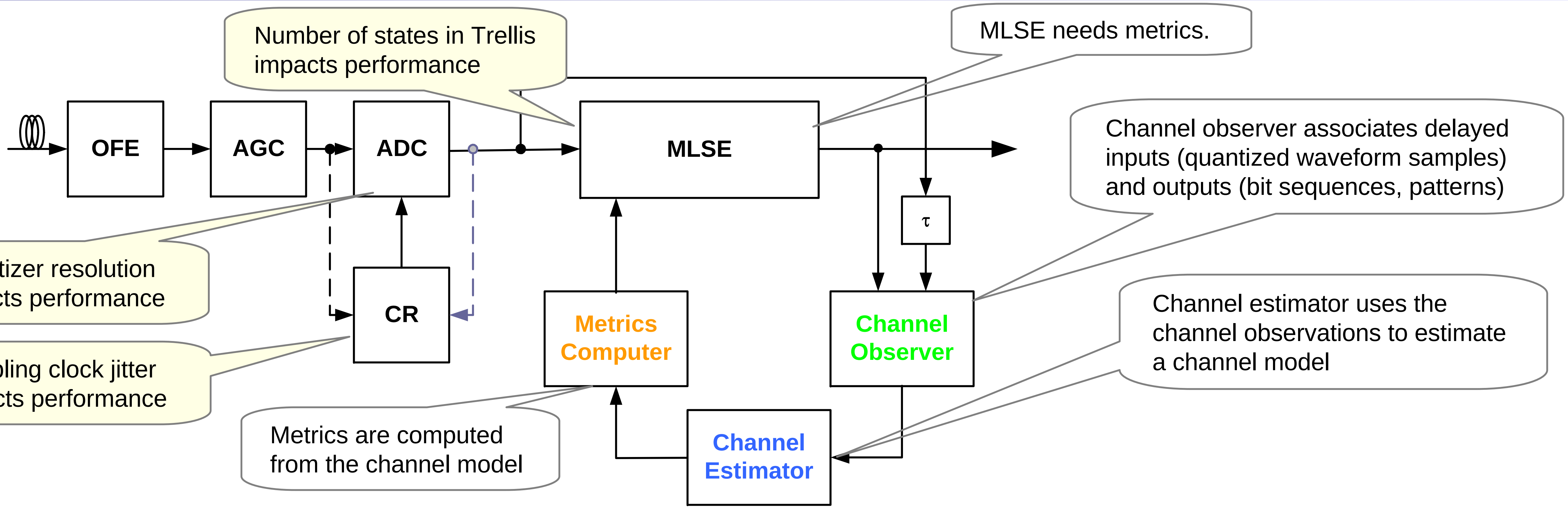


Using PDF values → wrong PMF value
Computing PDF is easy, but PMF is hard.



„Wrong“ Metrics?

Channel Estimation Methods for MLSE Metrics

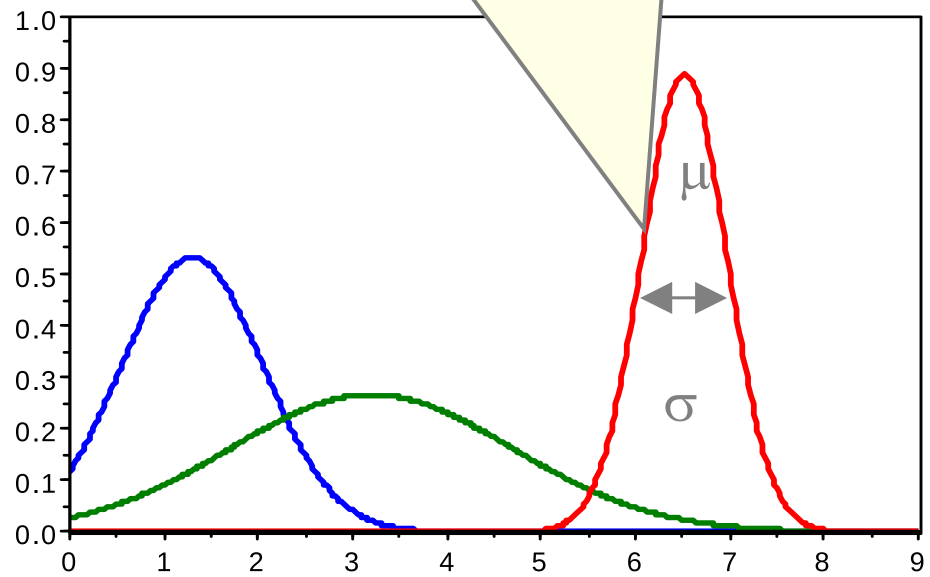


Parametric

There are two approaches of channel estimation

Non-parametric

Estimate **parameters** (e.g. μ, σ) of Probability Density Function (PDF) to *compute* log-likelihood metrics

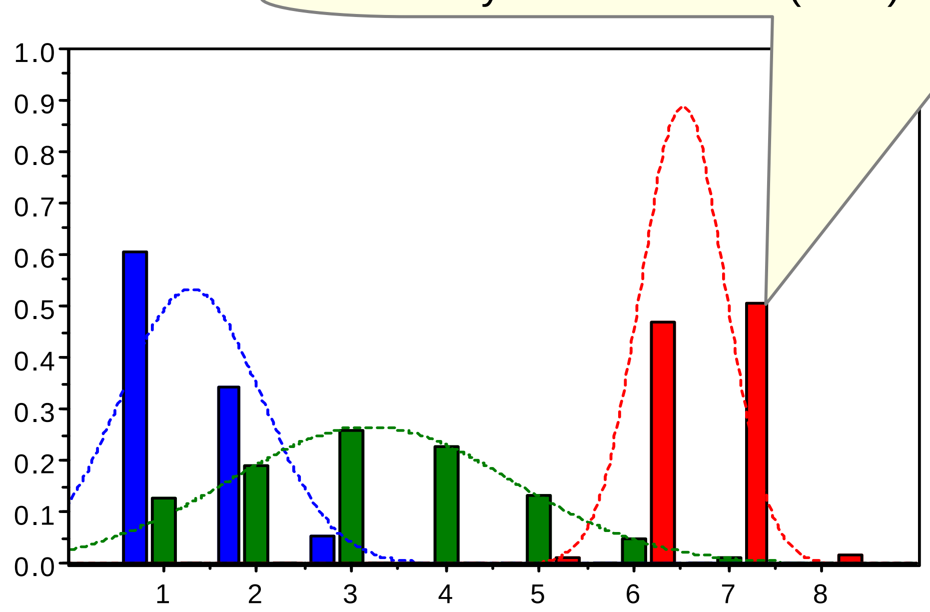


channel observations

estimate parameters → compute PDF → compute PMF → metrics

histograms → metrics

Estimate **probabilities**, i.e. values of Probablity Mass Function (PMF)



Any performance difference?

Abstract and Problem Statement

Abstract

*We compare the performance of MLSE-based receivers with **parametric and non-parametric channel estimation methods** and characterize their sensitivity against quantization, sampling jitter, and intersymbol interference (ISI) overload [(1)]*

MLSE needs branch metrics

Branch metrics are log-likelihoods

Two approaches to estimate likelihoods from observations:

Non-Parametric	Likelihoods are estimated directly (from observed relative **frequencies)**
Parametric	Likelihoods are estimated indirectly (parameters of a probabilistic model are estimated from observations)

(1) **ISI Overload:** The physical channel memory exceeds the state memory of the MLSE

Problem Statement

Do parametric models suffer from effects not covered in the model?
Are there relevant "model mismatch" penalties ?

Simulation Approach

Histogram Method "**HM**"	a practice-proven canonical method of non-parametric channel estimation
SQRT method "**SQRT**"	a particularly efficient example of a parametric method

*Compare **ultimately** and **practically** achievable performance of HM and of SQRT.*

Results and Conclusions

SQRT Method compared to Histogram Method (@ BER 10^{-3}) in a Nutshell

Ultimate Performance?	identical	without complexity limitations, i.e. for unlimited ADC resolution and unlimited number of states
Practical Performance?	slightly worse, but ...	penalty is not very relevant – achieves the same dispersion limit (e.g. 5000 ps/nm at 15 dB)
Model Mismatch Penalty?	yes, but ...	only at low dispersion and for PMD – and outside of useable operation range
Quantization Penalty?	yes, but ...	significant only for 3-bit ADC
Jitter Penalty?	no	not for relevant jitter magnitudes

Conclusions

We compared *ultimately* and *practically* achievable performance

We *assumed* that SQRT suffers more from “quantization” and “model mismatch”

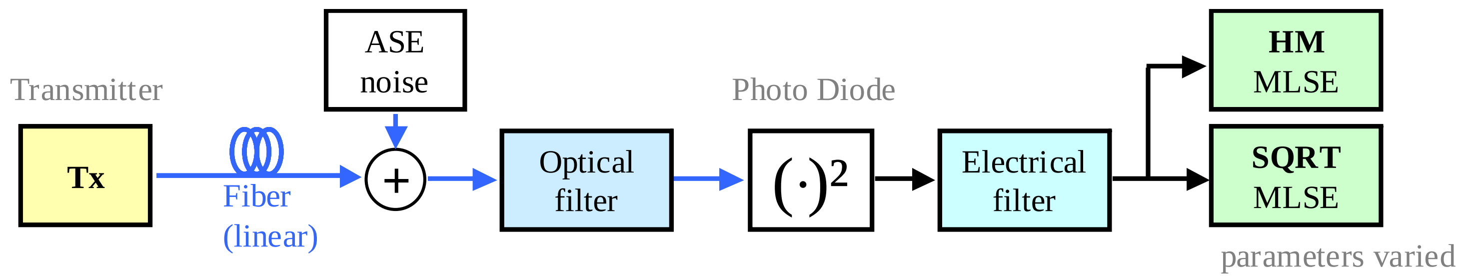
We *found* such penalties but they are not very significant

The HM channel estimator has practical performance advantage for 3-bit ADC

The SQRT channel estimator has speed & complexity advantages for N-bit ADC

For further study: Model mismatch penalties at lower BER?

Simulation Setups



Two setups were used

	Setup 1 „Bad Tx“	Setup 2 „Good Tx“
	for CD, PMD, Jitter	*for unconstrained complexity*
Data	DeBruijn-15 (2^{19} bits, 32 samples/bit)	PRBS-18 (2^{18} bits, 20 samples/bit)
Format	NRZ @ 10.7 Gbit/s	
Shaping Filter	0.3 UI rise-time erfc shaped + 1-pole Bessel (10.7 GHz)	5-pole Bessel (7.5 GHz)
Extinction Ratio	11.8 dB	infinite
Fiber	SSMF (D=16 ps/nm), linear propagation	
Opt. filter	Flat Top (40GHz)	SuperGauss 2nd Order (35 GHz)
El. Filter	4-pole Bessel (7.5 GHz)	5-pole Bessel (7.5 GHz)
AGC / ADC	gain optimized(1) *roughly*, best sampling phase, varied quantizer resolution	
MLSE	2 samples per bit, self-training, varied number of states	

(1) For HM, gain was not optimized. Mean rectified value was maintained at a constant level.

References

Non-parametric Channel Estimation

suggested early for MLSE usage in non-linear channel

W. Sauer-Greff et al., "Modified Volterra Series and State Model Approach …", Proc. IEEE Sig Proc 99, 19-23

H. F. Haunstein et al., "Design of near optimum electrical equalizers for optical transmission …", OFC 2001, WAA 4-1

implemented in real systems, e.g.

A. Färbert et al., "Performance of a 10.7 Gb/s Receiver with Digital Equalizer using …", ECOC 2004, Th4.1.5

(many) experimental data available, e.g.

J.P. Elbers et al., „Measurement of the dispersion tolerance of optical duobinary …", OFC 2005, OThJ4

S. Chandrasekhar et al., "Chirp-managed laser and MLSE-RX …", PTL, Vol. 18, No. 14, 1560-1562, 2006

S. Chandrasekhar et al., "Performance of MLSE Receiver …", PTL, Vol. 18, No. 23, 2448-2450, 2006

J. M. Gené et al., "Joint PMD and Chromatic Dispersion Compensation Using an MLSE", ECOC 2006 , We2.5.2

I. L. L. Polo et al., "Comparison of Maximum Likelihood Sequence Estimation equalizer …", ECOC 2006, We2.5.3

J. D. Downie et al., "Experimental Measurements of the Effectiveness of MLSE …", OFC 2007, OMG4

C. Xie et. al., "Performance Evaluation of Electronic Equalizers for Dynamic PMD …", OFC 2007, OTuA7

Parametric Channel Estimation

studied since long (for perfomance analysis), e.g.

P. A. Humblet, M. Azizoglu, "On the Bit Error Rate …", JLT 9/11 p.1577 (3), 1991 (and predecessors)

A. Weiss, "On the Performance of Electrical Equalization in Optical Fiber Transmission Systems", PTL, Vol. 15, 1225-1227, 2003

well covered in recent MLSE literature, e.g.

D. E. Crivelli et al., "On the Performance of Reduced-State Viterbi Receivers …", ECOC 2004, We4.P.083

N. Alic et al., "Signal statistics and maximum likelihood sequence estimation …", Optics Express, Vol.13, No. 12, 4568-4578, 2005

G. Bosco et al., "Long-Distance Effectiveness of MLSE IMDD Receivers", PTL, vol. 18, pp.1037-1039, 2006

T. Freckmann, J. Speidel, "Viterbi Equalizer with Analytically Calculated Branch Metrics …"PTL, Vol. 18, 277-279, 2006

T. Foggi et al., "Maximum-likelihood sequence detection with closed-form metrics …", JLT, Vol. 24, No. 8, 3073-3087, 2006

P. Poggiolini et al., "Branch Metrics for Effective Long-Haul MLSE", ECOC 2006, We2.5

M. R. Hueda et al., "Parametric Estimation of IM/DD Optical Channels …", JLT, Vol. 25, No. 3, 957-975, 2007

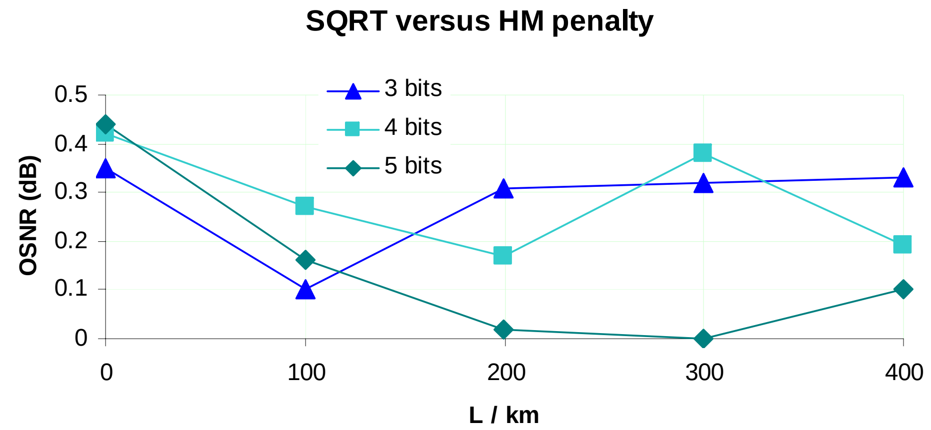
(some) experimental data from offline simulations, e.g.

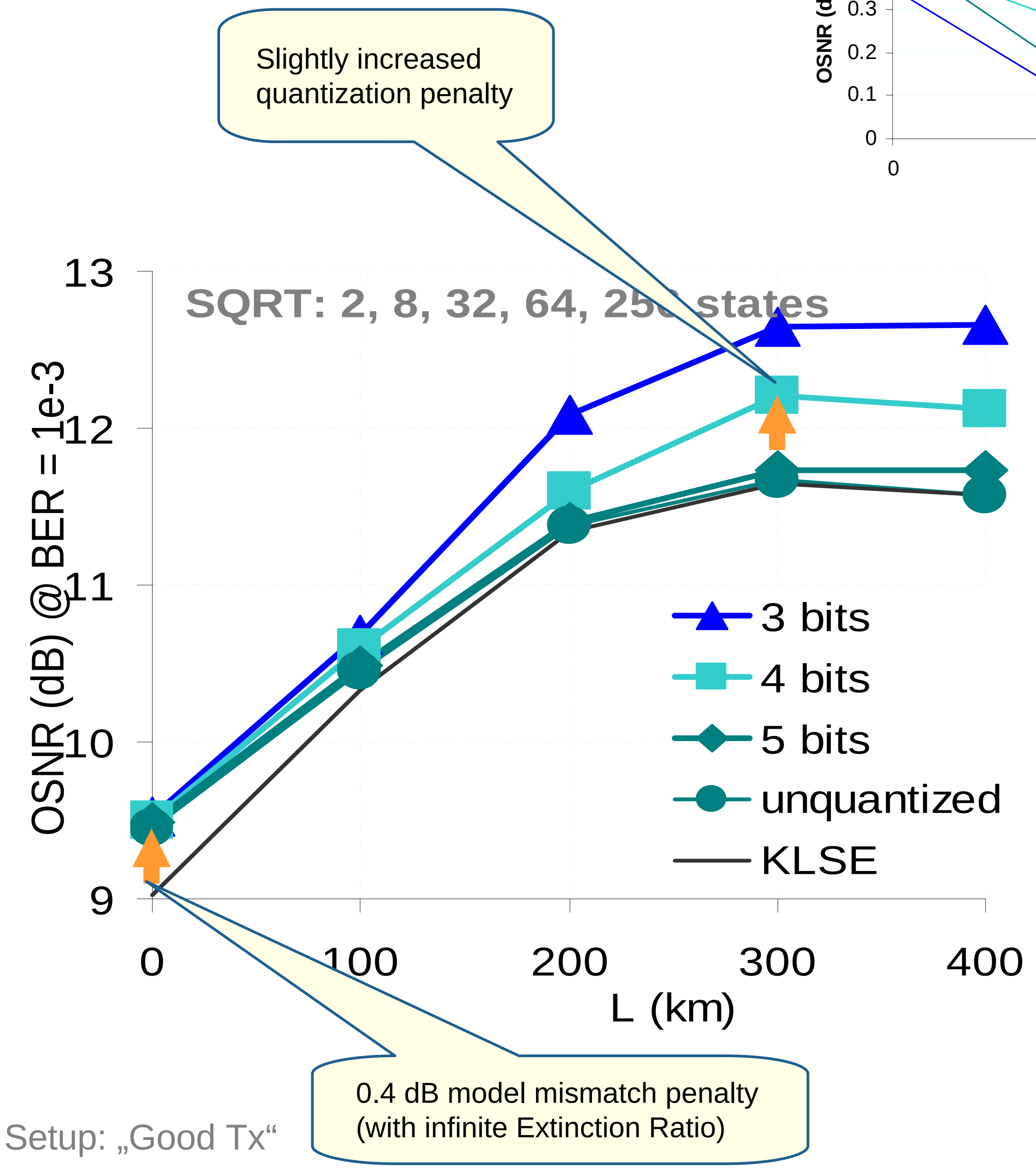
P. Poggiolini et al., "1,040 km uncompensated IMDD transmission …", ECOC 2006, post-deadline Th4.4.6

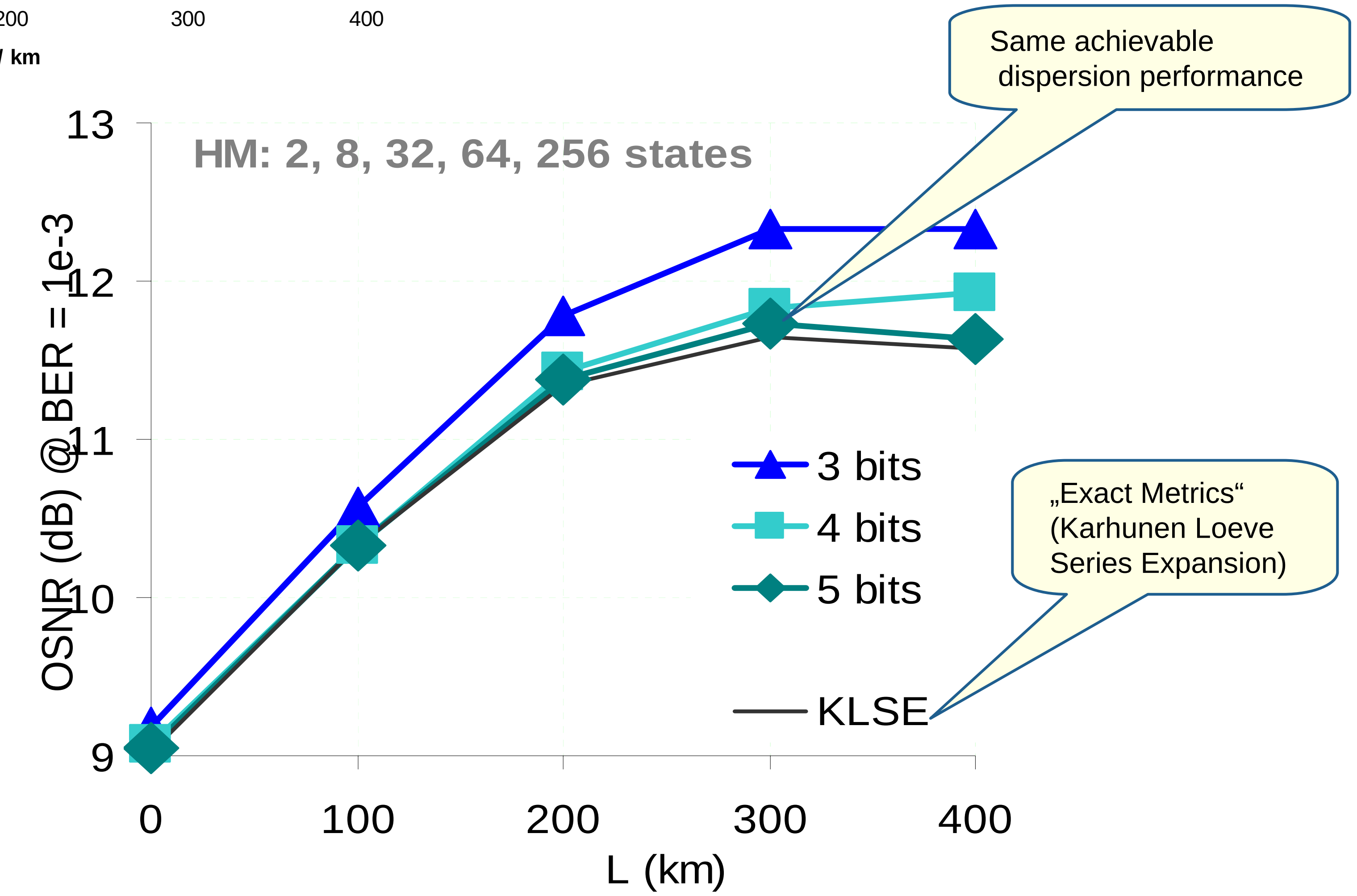
Performance with Unconstrained MLSE and ADC

SQRT Method

Histogram Method







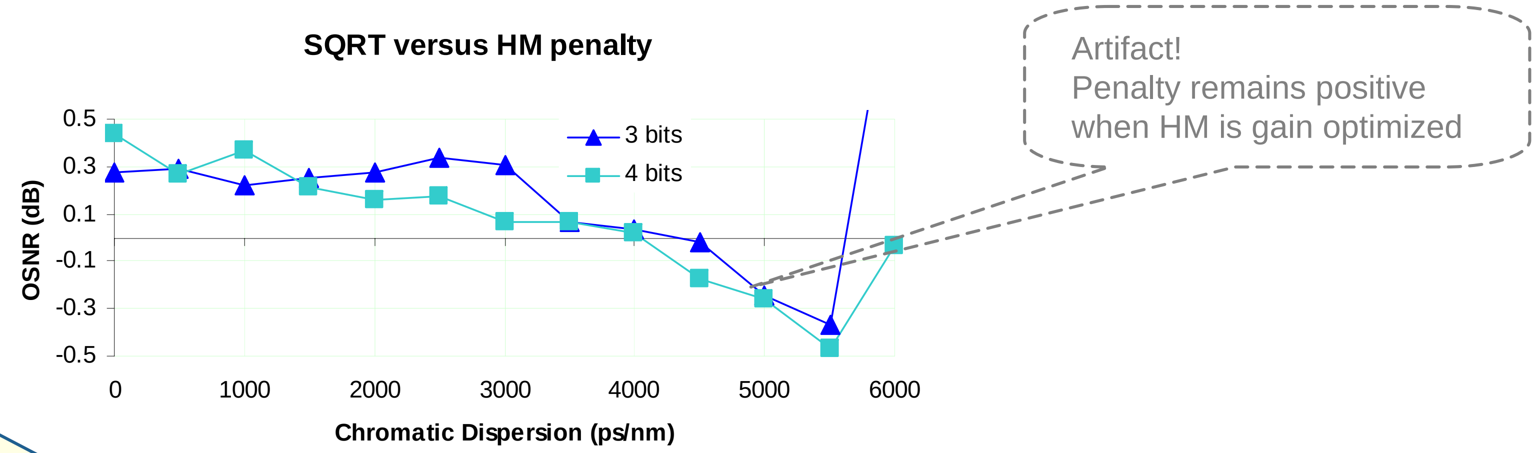
Setup: „Good Tx“

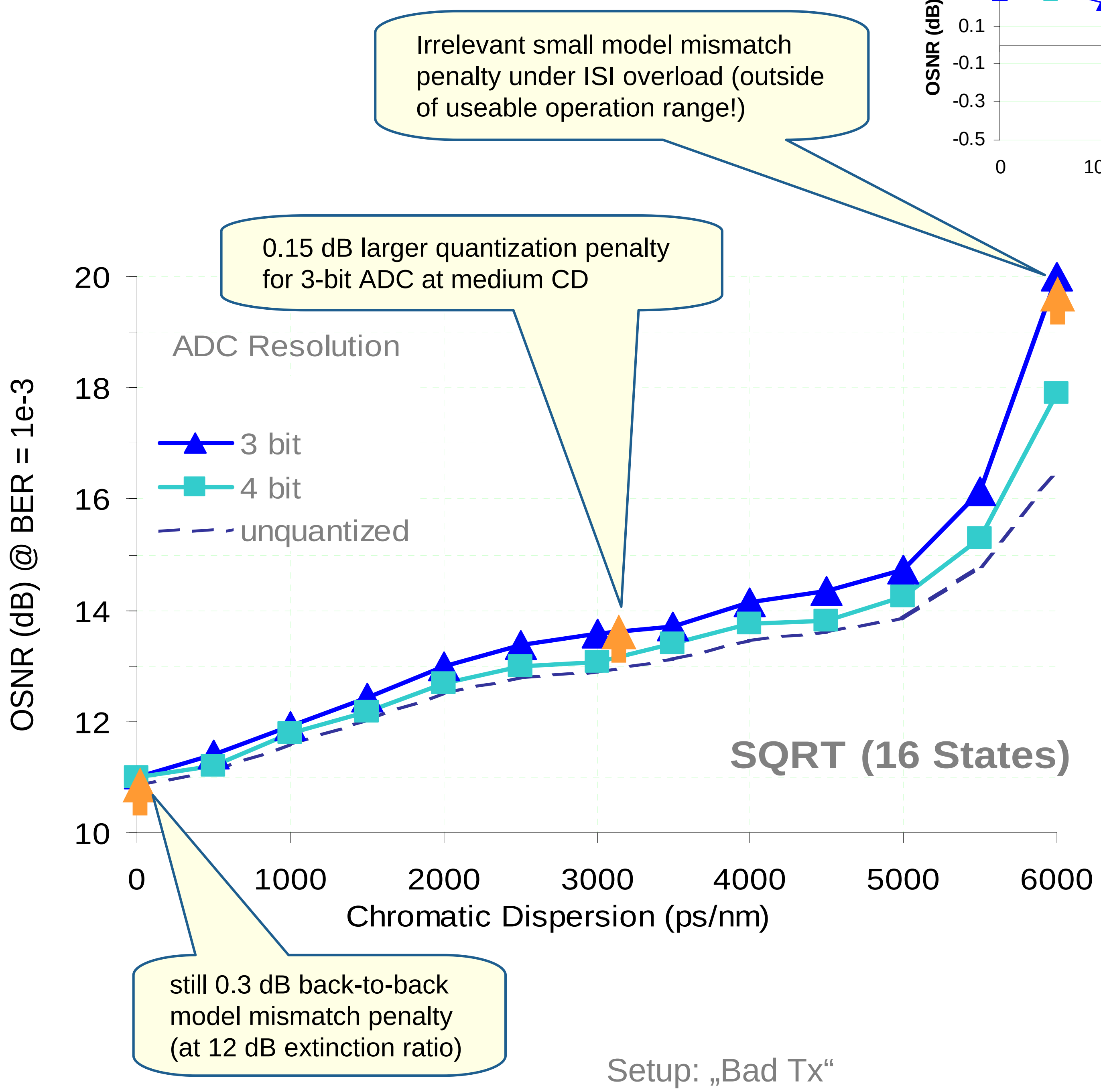
Achievable Performance? No difference for fine ADC

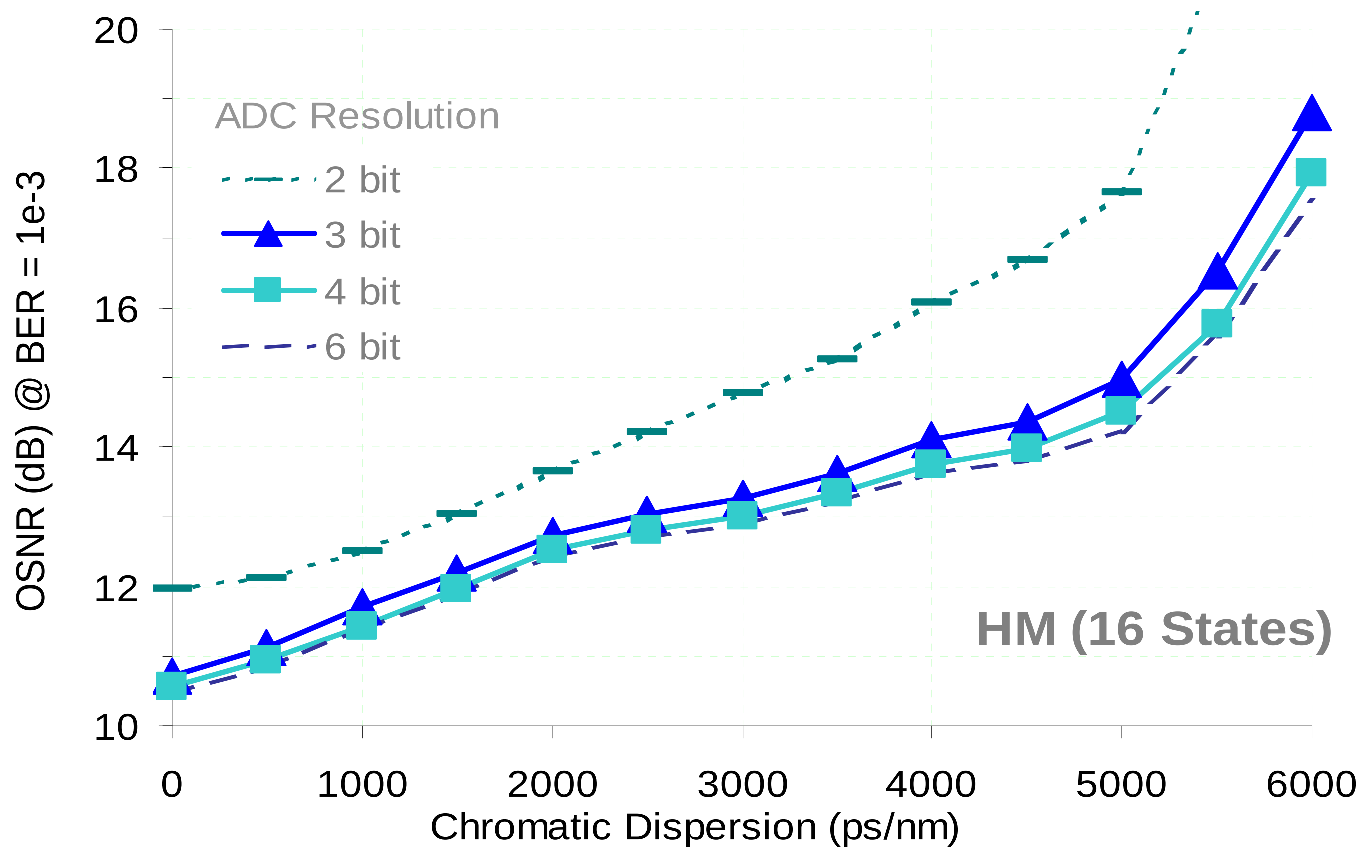
Dispersion Tolerance with 16-states MLSE and ADC

SQRT Method

Histogram Method







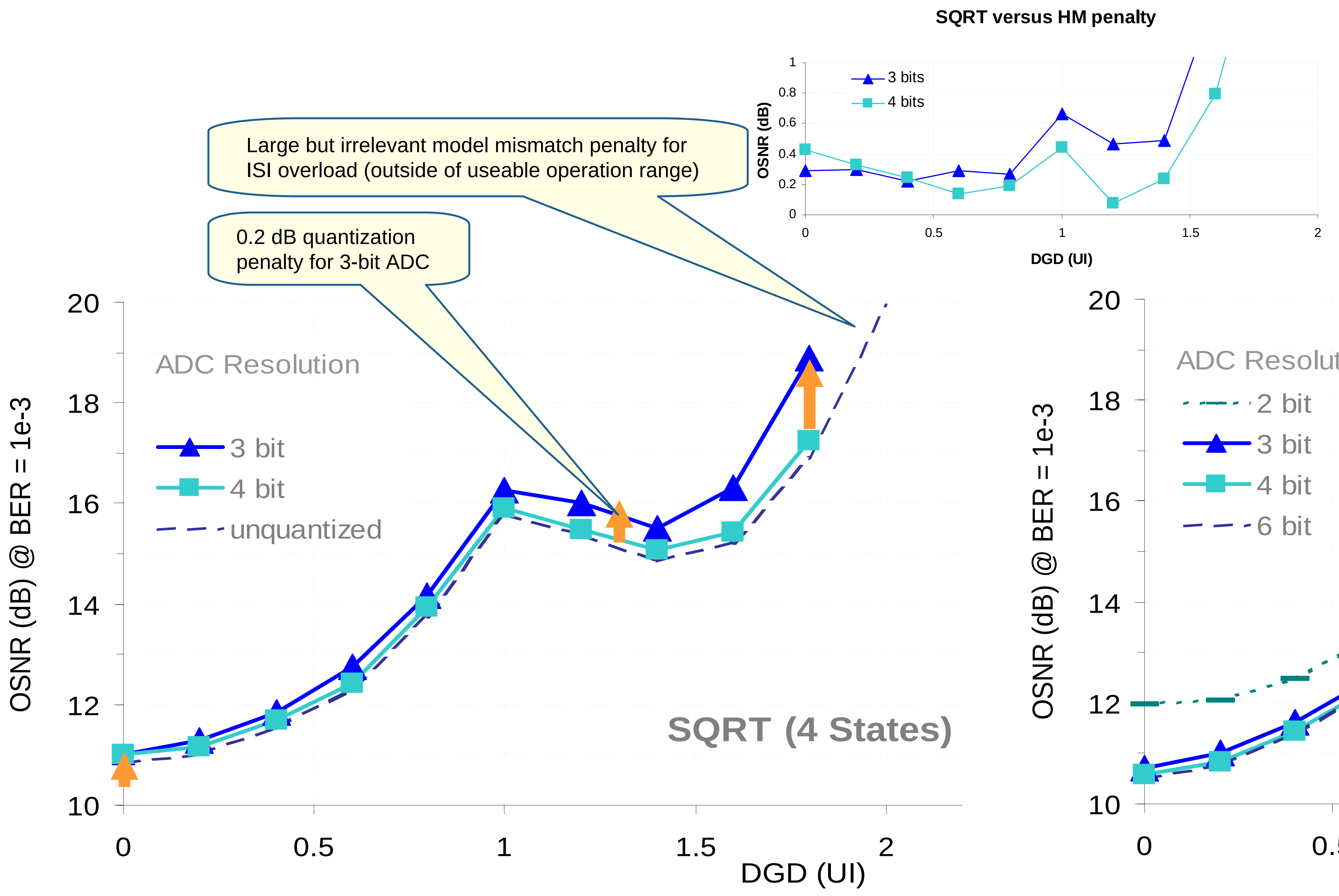
Setup: „Bad Tx“

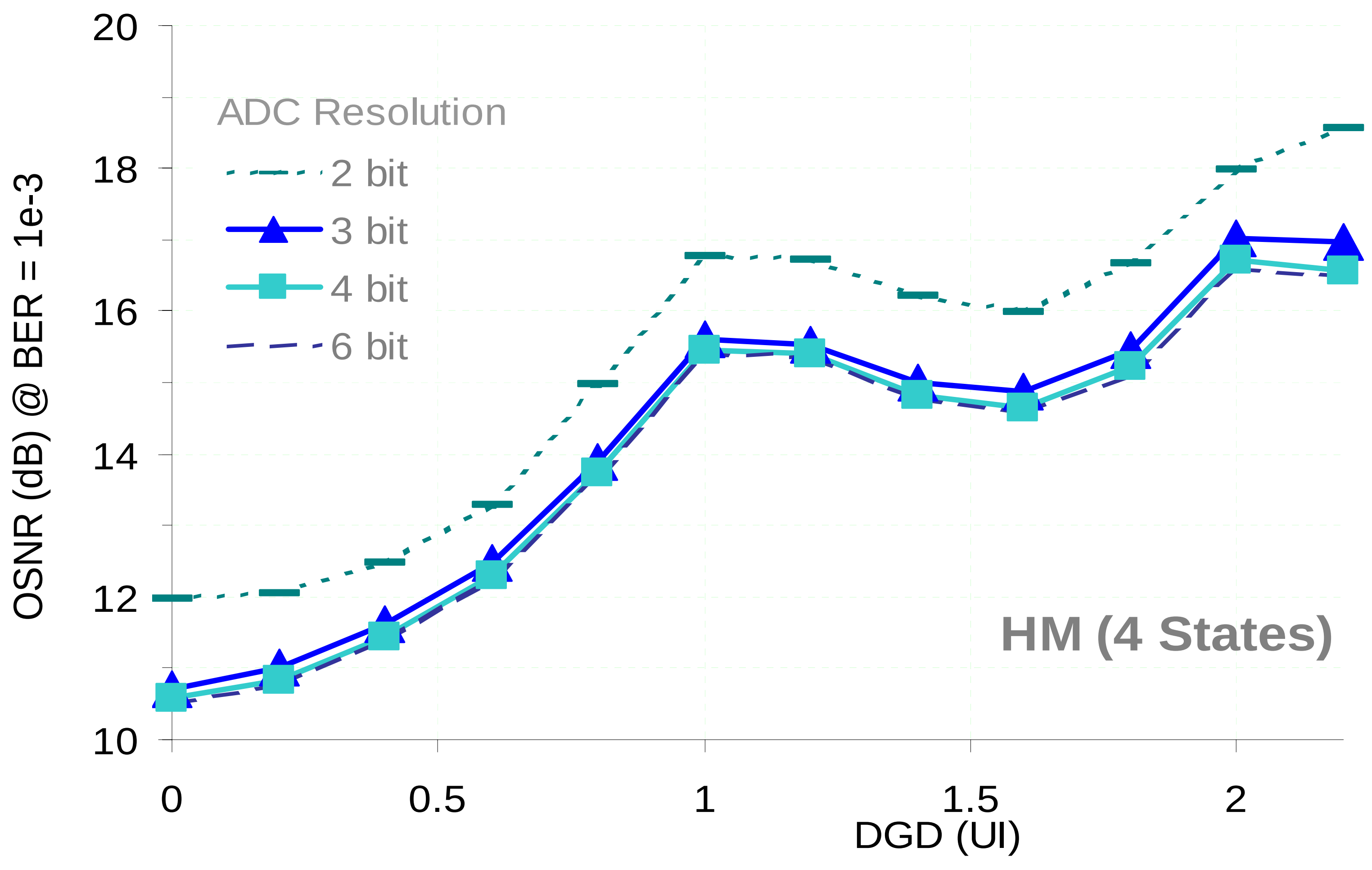
Dispersion? Relevant differences are small

1st order PMD with 4-states MLSE and ADC

SQRT Method

Histogram Method





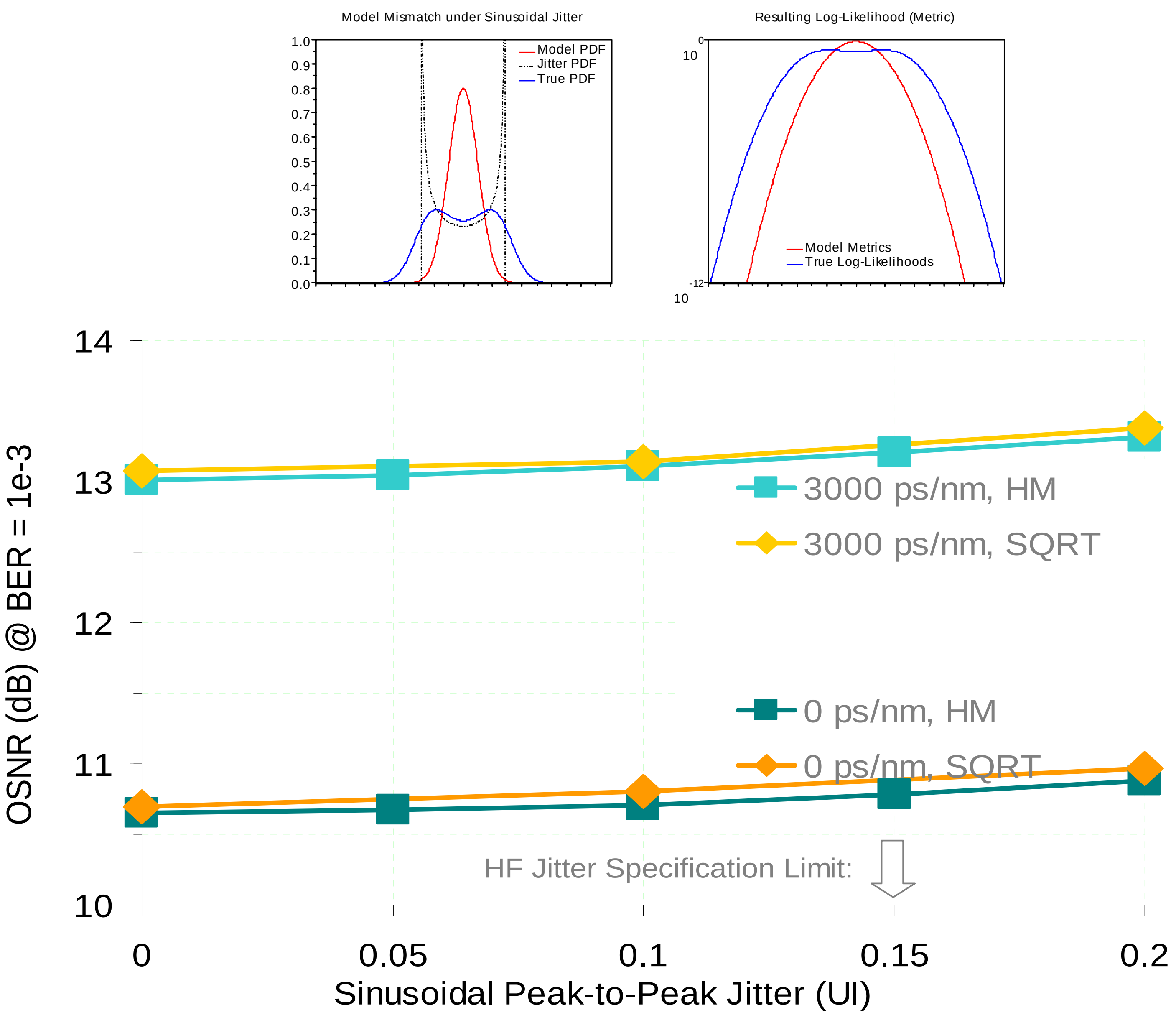
Setup: „Bad Tx“

PMD? Relevant differences remain small

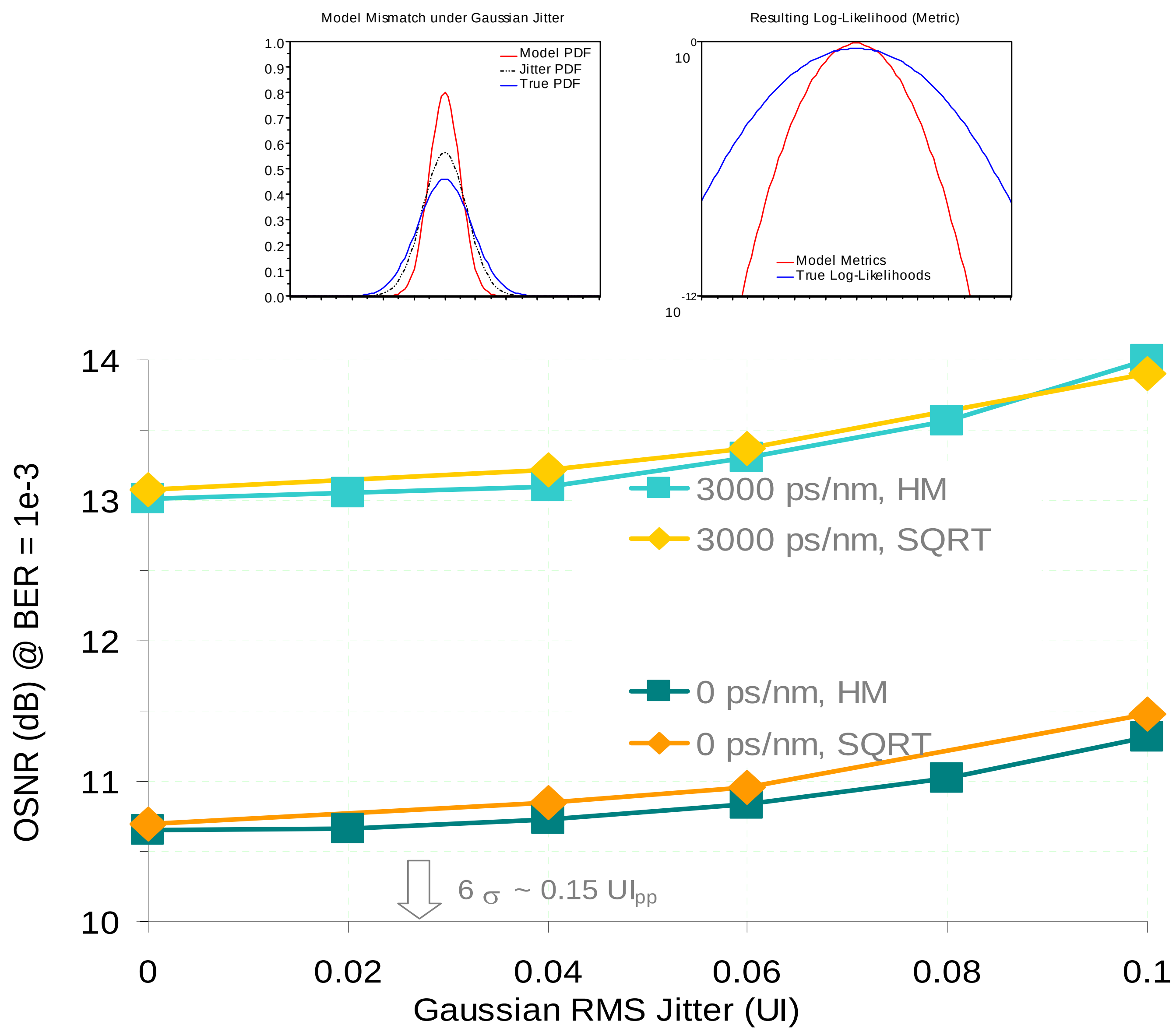
Clock Recovery with Jitter (16-states and 4-bit ADC)

Sinusoidal Jitter

(Test Signal)



Gaussian Jitter



Setup: „Bad Tx“

Jitter? No relevant difference