Addition of Quantization Noise in COM

Supporting Document for Comments #360-370 Against Draft 1.3

Hossein Shakiba
Huawei Technologies Canada
January 2025

Background

- In contribution *shakiba_3dj_02_2405.pdf* effects of quantization noise in COM channel compliance verification was analyzed and significance of its inclusion was demonstrated
- Contribution *healey_3dj_01b_2405.pdf* considered using existing means (e.g. scaling eta_0) as a proxy to represent quantization noise
- At the time, there was more support for using the simpler proxy method
  - Still considerable Y's and a lot of undecideds

May 2024

Straw Poll #1

I support adding a new noise term (such as 'eta_1' in healey_3dj_01a_2405, slide 6) to the COM reference receiver.

Results (all) Y: 13, , N: 37 , A: 31

- The "N" outcome was mostly motivated by the argument of "reference receiver trap"
- While I generally understand the argument and agree with it where applicable, in my view the quantization noise is too important to be ignored or represented through a simple proxy
- Also, it would have helped if the theoretical basis of the modeling approach and its calculation overhead were better understood, justified, and quantified

Facts to Consider

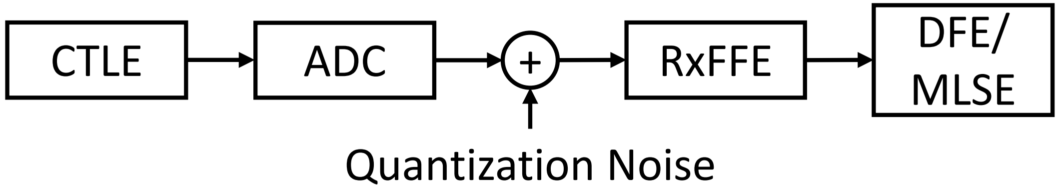
- A big part of the “reference receiver trap” was to avoid features that are implementation-specific and could cause unacceptable complication for the purpose
- Vast majority to almost all receiver implementations nowadays use ADC, making this architecture generic and de-facto, and the natural baseline for the reference receiver
- Shift in paradigm to consider the non-ADC-based receiver implementation-specific
- Quantization noise modeling stands on a solid theoretical foundation and can be simply embedded with very little overhead
- Quantization noise has some unique and specific attributes that makes it not a good candidate to be replaced by a proxy as simple as a fix scaled and uncorrelated eta_0 noise term
- Several other attributes of the current reference receiver and existing noise terms are likely less important and arguably more implementation-specific

Motivation

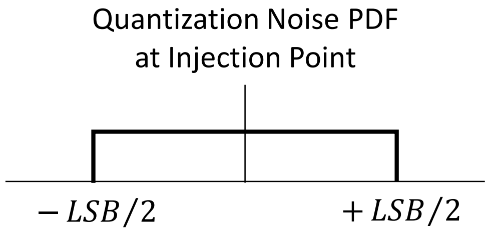
- Some observations and developments since then:
  - ❖ There was a lack of enough data and clarity on the extend of the overhead of adding the quantization noise model to the COM flow
  - ❖ Noticeable ongoing interests and requests to further follow up on this topic
  - ❖ Several direct requests for having access to the COM Matlab function with the capability
  - ❖ More data have been generated and some presented by others since then
  - ❖ The latest beta version (480beta2) of the COM Matlab function incorporates the feature
    - ■ Clearly quantifies a very reasonable calculation overhead for the added value
    - ■ Provides a wider access
    - ■ There will be a COM change request follow-up in the next COM ad hoc meeting
- Comments #360-370 along with this contribution is another effort to bring more awareness and the importance of the quantization noise to the audience's attention
- Hopefully consensus will be built and a move in the right direction will be made

Quantization Noise Model

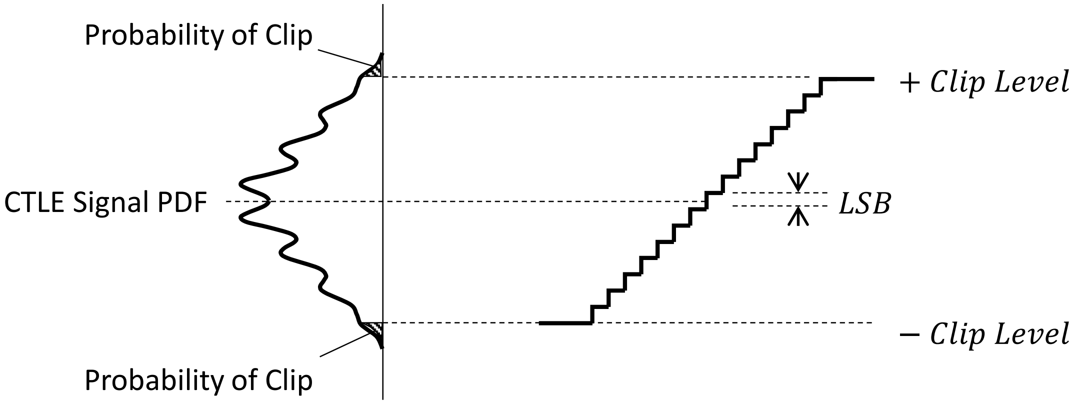
- Quantization noise is a new noise term added between CTLE and RxFFE



- It is modeled by a white random noise with uniform distribution over –LSB/2 to +LSB/2 at the injection point



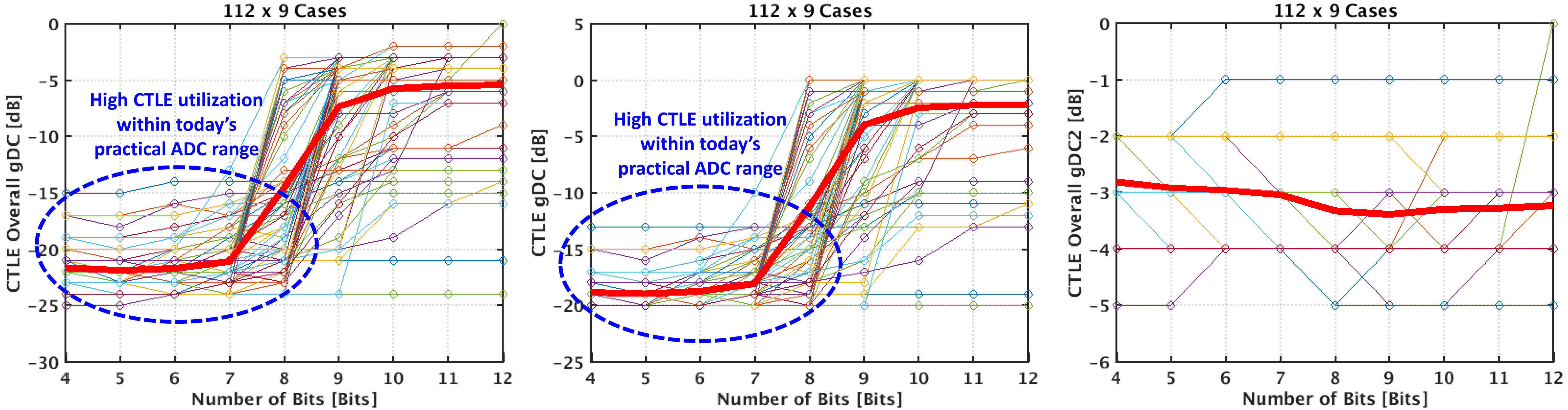
$$\sigma_{qn} = \frac{LSB}{\sqrt{12}}$$



- Quantization clip level is calculated from the desired probability of signal clipping
- LSB, the quantization step size, is calculated from the desired number of bits and clip level
- Note that modeling quantization functionality is outside the scope, it is only its noise

Impact of Quantization Noise on Equalization

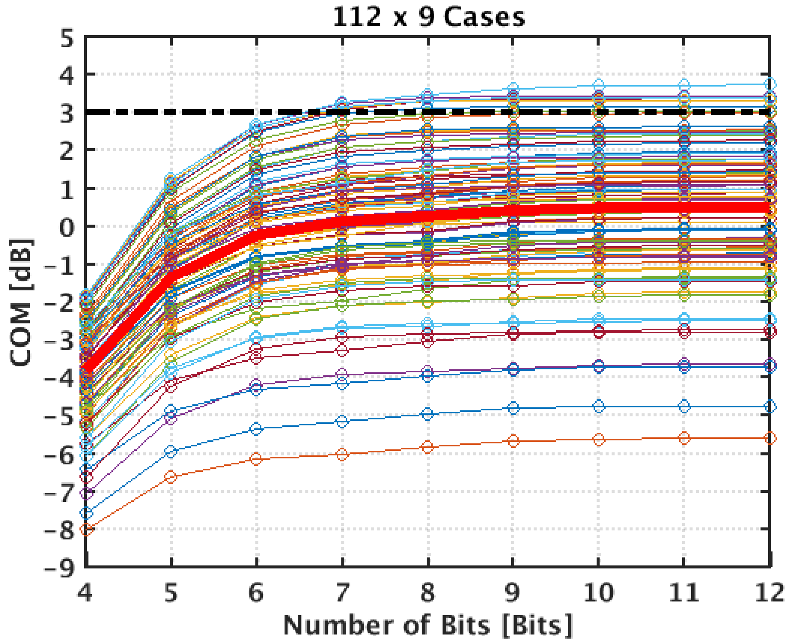
- Quantization noise has a prominent impact on the equalizer distribution and optimization

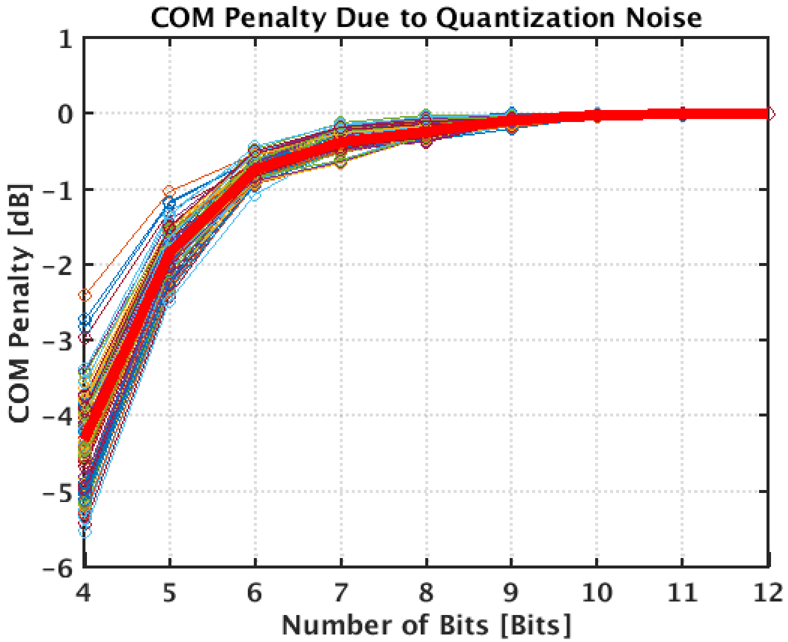


- CTLE high-frequency gain (gDC) utilization increases with increasing quantization noise
  - ❖ CTLE search range can not be generally reduced (fixing gDC to speed up optimization is not a choice)
- CTLE high-frequency utilization is unrealistically minimal when eta_0 is used as a proxy
- As expected, CTLE low-frequency boost (gDC2) utilization does not change

Impact of Quantization Noise on COM

- Impact of quantization noise on COM is quantified more accurately, predictively, and realistically

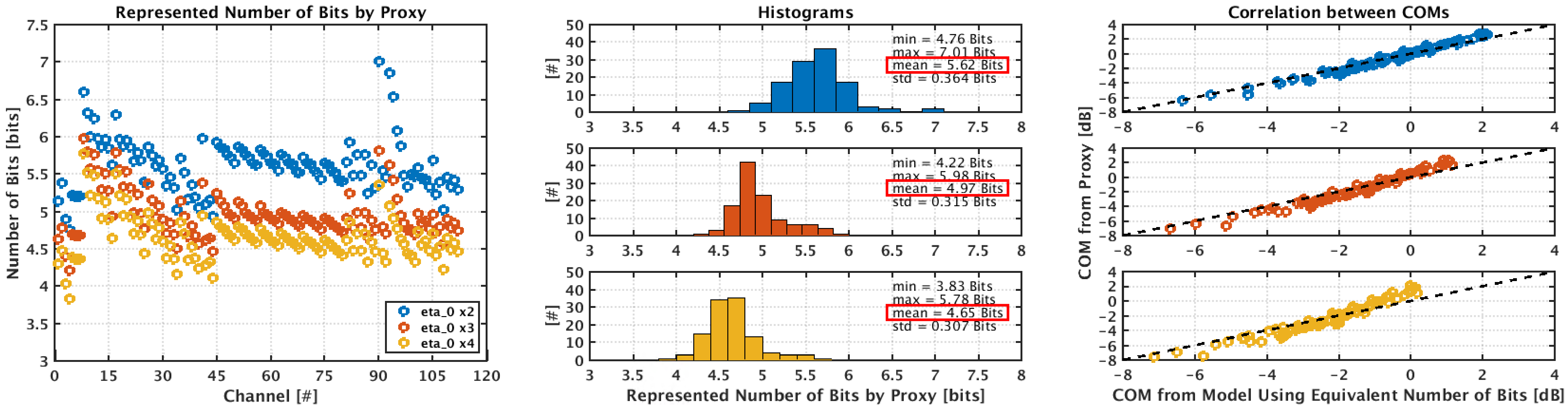




- For the test channels, at least 6 bits is recommended to contain the quantization noise
- With 6 bits, the test channels suffer anywhere between 0.44dB to 1.07dB of COM penalty

The Issue with eta_0 Proxy Approach

- Mapping COM data from when eta_0 is scaled (proxy) to COM data from sweeping number of bits (previous slide) reveals the number of bits that scaling represents for each channel



- Equivalent number of bits that each scaling value represents is the average over channels and can be mapped to the COM value that quantization noise model would have given for each channel

		eta_0 x2	eta_0 x3	eta_0 x4
Equivalent Number of Bits		5.62	4.97	4.65
COM Error [dB]	min	-1.03	-1.49	-1.64
	max	0.87	1.53	2.17
	std	0.36	0.55	0.69

Less Accurate

- For the test channels, correlation is weak

Description of Comments

- Comment #360:
  - Page 757, after Section 178A.1.7.5
  - Add a new sub-section “178A.1.7.6 Quantization noise”

178A.1.7.6 Quantization noise

The power spectral density of the quantization noise at the input of the quantized time receiver equalizer is defined by Equation (178A-X0).

$$S_{qn}(\theta) = \sigma_q^2 / f_b \qquad (178\text{A} - \text{X0})$$

where σ_q^2 is the power of the quantization noise at the output of the quantizer defined by Equation (178A-X1).

$$\sigma_q^2 = LSB^2/12 \qquad (178\text{A} - \text{X1})$$

where LSB is the quantization step size defined by Equation (178A-X2).

$$LSB = 2\ CL/\left(2^{N_{qb}} - 1\right) \qquad (178\text{A} - \text{X2})$$

Description of Comments

where CL is the quantization clip level defined by Equation (178A-X3).

$$CL = -P^{-1}\left(P_{qc}/2\right) \qquad (178\text{A} - \text{X3})$$

where P is the cumulative distribution function of the signal prior to quantization defined by Equation (178A-X4).

$$P(y) = \int_{-\infty}^{y} p(u)du \qquad (178\text{A} - \text{X4})$$

where p is the probability density function of the signal prior to quantization defined by Equation (178A-X5).

$$p(y) = p_{sig}(y) * p_{noise}(y) \qquad (178\text{A} - \text{X5})$$

where

p_{sig} is the probability density function of the noiseless signal prior to quantization obtained by following the procedure defined in 93A.1.7.1 with pulse response $h(n)$ replaced by the pulse response from the input of the transmitter FFE to the input of the quantizer.

Description of Comments

p_{noise} is the probability density function of the noise prior to quantization estimated by a Gaussian distribution* defined by Equation (178A-X6).

$$p_{noise}(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2\sigma_{noise}^2}} \qquad (178\mathrm{A} - \mathrm{X}6)$$

where σ_{noise}^2 is the power of the noise prior to quantization defined by Equation (178A-X7).

$$\sigma_{noise}^2 = \sigma_{rn}^2 + \sum_{k=1}^{K-1} {\sigma_{xn}^{(k)}}^2 + \sigma_{tn}^2 + \sigma_{jn}^2 + \sigma_{in}^2 \qquad (178\mathrm{A} - \mathrm{X}7)$$

where each of the terms in the Equation (178A-X7) is calculated from the general Equation (178A-X8) using their corresponding power spectral densities obtained from sections 178A.1.7.1 to 178A.1.7.5.

$$\sigma^2 = \int_{-\pi}^{\pi} S(\theta) d\theta \qquad (178\mathrm{A} - \mathrm{X}8)$$

* The actual noise probability distribution is not necessarily Gaussian. However, for the purpose of equalizer optimization a Gaussian assumption helps optimization algorithm run time by avoiding to calculate the noise PDF every time. The final COM will be based on the actual noise PDF.

Description of Comments

- Comment #361:
  - Page 754, Section 178A.1.7, Figure 178A-7
  - Add quantization noise to the figure

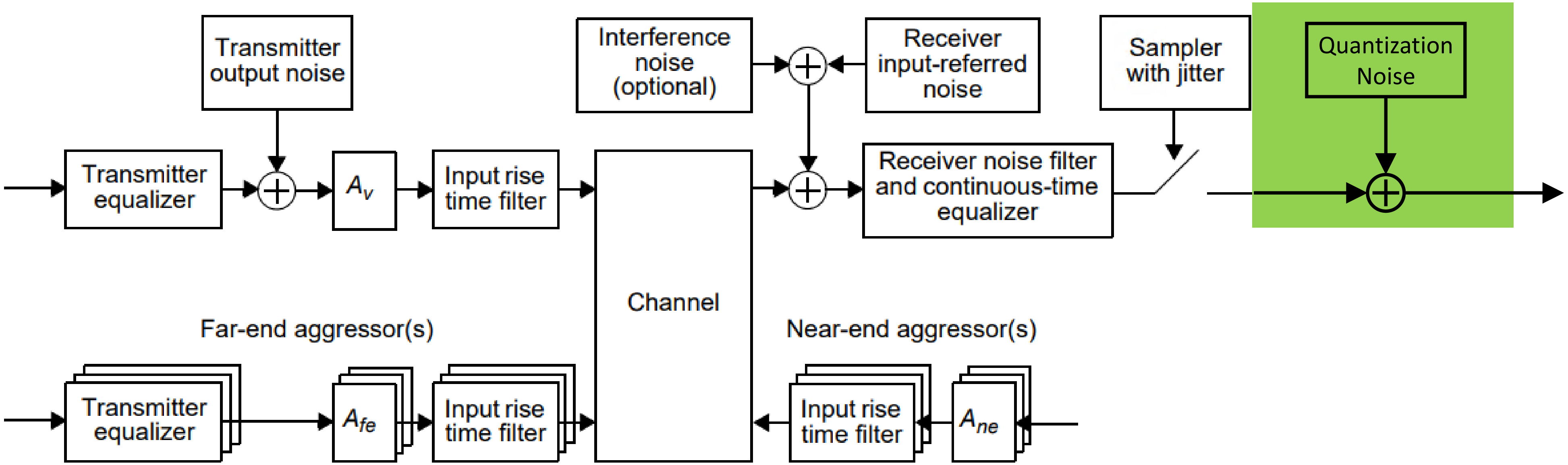


Figure 178A–7—Sources of noise considered in calculation of COM

Description of Comments

- Comment #362:
  - Page 755, Section 178A.1.7, Table 178A-9
  - Add number of quantization bits and clipping rate to the table

Table 178A–9—Summary of noise parameters

Parameter	Symbol	Units
Number of signal levels	L	—
One-sided noise spectral density at receiver input	η_0	V^2/GHz
Transmitter signal-to-noise ratio	SNR_{TX}	dB
Random jitter, RMS	σ_{RJ}	UI
Dual-Dirac jitter, peak	A_{DD}	UI
Number of quantization bits	N_{qb}	–
Quantization clip rate	P_{qc}	–

Description of Comments

- Comment #363:
  - Page 755, Section 178A.1.7, Equation (178A-14)
  - Add Quantization Noise PSD

$$S_n(\theta) = S_{rn}(\theta) + \sum_{k=1}^{K-1} S_{xn}^{(k)}(\theta) + S_{tn}(\theta) + S_{jn}(\theta) + S_{in}(\theta) + S_{qn}(\theta) \quad (178A\text{–}14)$$

where

θ	is normalized frequency in the range $[-\pi, \pi)$ where $\pi = f_b / 2$
$S_{rn}(\theta)$	is the receiver input-referred noise power spectral density defined in 178A.1.7.1
$S_{xn}^{(k)}(\theta)$	is the crosstalk power spectral density for signal path k defined in 178A.1.7.2
$S_{tn}(\theta)$	is the transmitter output noise power spectral density defined in 178A.1.7.3
$S_{jn}(\theta)$	is the power spectral density of the noise due to jitter defined in 178A.1.7.4
$S_{in}(\theta)$	is the interference noise power spectral density defined in 178A.1.7.5
$S_{qn}(\theta)$	is the quantization noise power spectral density defined in 178A.1.7.6

Description of Comments

- Comment #364:
  - Page 754, Section 178A.1.7, Line 32
  - Change “sampler” to “quantizer”
- Comment #365:
  - Page 755, Section 178A.1.7, Line 15
  - Change “sampler” to “quantizer”
- Comment #366:
  - Page 757, Section 178A.1.8.1, Line 43
  - Change “sampler” to “quantizer”

Description of Comments

- Comment #367:
  - Page 758, Section 178A.1.8.1, Figure 178A-9
  - Add quantization noise after sampler

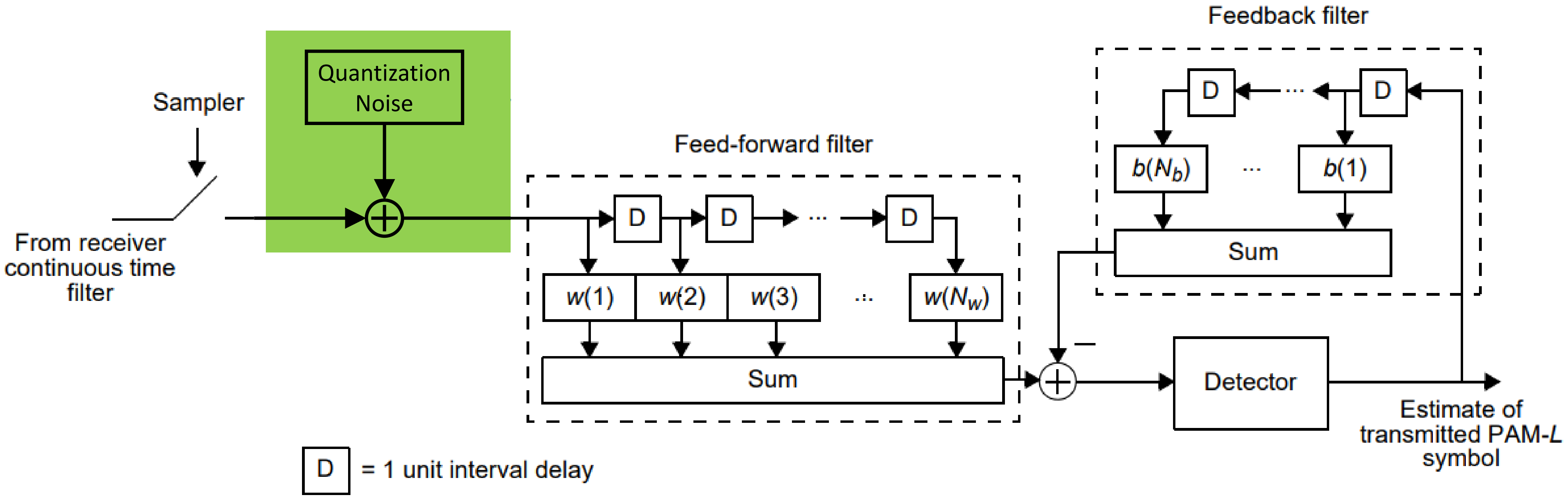


Figure 178A–9—Receiver discrete-time equalizer

Description of Comments

- Comment #368:
  - Page 761, Section 178A.1.9, Equation (178A-34)
  - Add Quantization Noise PSD

$$\sigma_G^2 = f_b \int_{-\pi}^{\pi} [S_{tn}(\theta) + S_{jn}^{(RJ)}(\theta) + S_{rn}(\theta) + S_{in}(\theta) + S_{qn}(\theta)] |H_{rxffe}(\theta)|^2 d\theta \qquad (178A\text{–}34)$$

Description of Comments

- Comment #369:
  - Page 761, Section 178A.1.10.2, Line 51
  - Add the following before the last sentence

178A.1.9. The corresponding cumulative distribution function is defined by Equation (178A–38).

Add quantization noise by convolving the resulting probability density function with the quantization noise probability distribution function defined by evaluating Equation (178A-X9) sequentially for integer values $j = 1$ to $j = N_w$ with p_{qn} initialized to a Dirac delta function.

$$p_{qn}(y) = p_{qn}(y) * \frac{1}{w_{lim}(j)} p_{qi}\left(\frac{y}{w_{lim}(j)}\right)$$

$(178\text{A} - \text{X}9)$

where

N_w is the total number of taps in the feed-forward filter of the receiver equalizer defined in 178A.1.8.1.

w_{lim} is the optimized vector of taps in the feed-forward filter of the receiver equalizer defined in 178A.1.8.1.

Description of Comments

p_{qi} is the probability density function of the quantization noise at its injection point defined by Equation (178A-X10).

$$p_{qi}(y) = \begin{cases} 1/LSB & -LSB/2 \leq y < LSB/2 \\ 0 & Otherwise \end{cases} \qquad (178\text{A} - \text{X}10)$$

where LSB is defined by Equation (178A-X2) with replacing p_{noise} with the actual probability density function of the noise prior to quantization and after optimization during execution of the equation.

Description of Comments

- Comment #370:
  - Page 762, Section 178A.1.11, Figure 178A-10
  - Add quantization noise after sampler

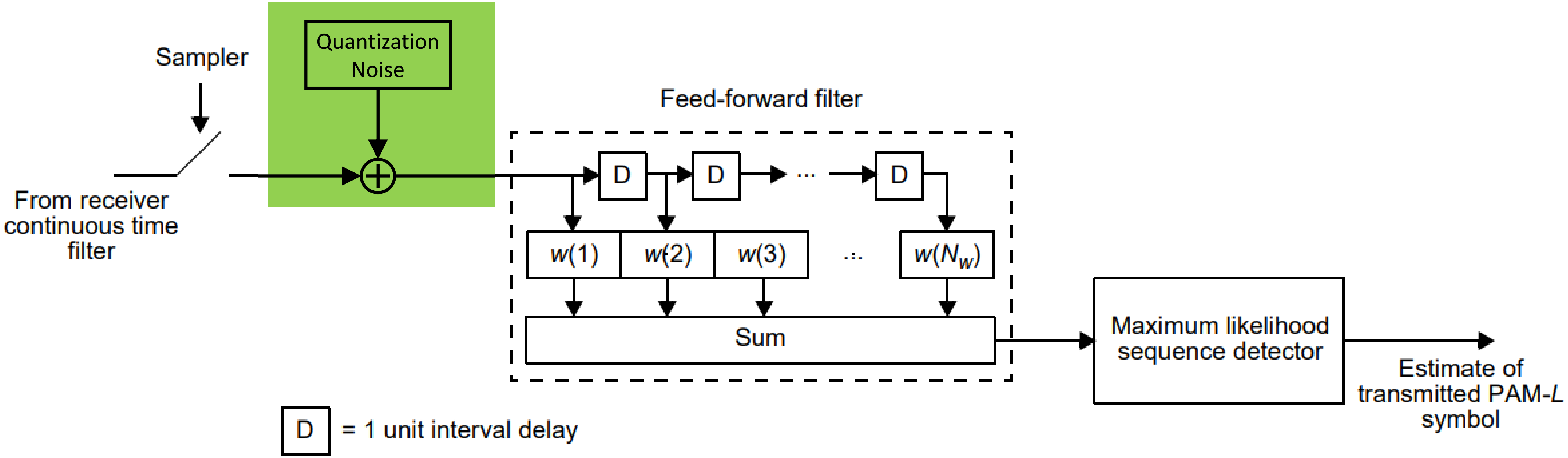


Figure 178A–10—Receiver discrete-time equalizer with MLSD

Thank You ☺

Hossein Shakiba
Huawei Technologies Canada

COM Info

- Version
  - COM version 480beta2_hs1p0
  - Customization _hs1p0 applies changes for which a follow-up request will be presented in the next COM ad hoc
- Configuration

Table 93A-1 parameters			
Parameter	**Setting**	**Units**	**Information**
f_b	106.25	GBd	
f_min	0.05	GHz	
Delta_f	0.01	GHz	
C_d	[0.4e-4 0.9e-4 1.1e-4 ;0.4e-4 0.9e-4 1.1e-4]	nF	[TX RX]
L_s	[0.13 0.15 0.14; 0.13 0.15 0.14]	nH	[TX RX]
C_b	[0.3e-4 0.3e-4]	nF	[TX RX]
R_0	5.00E+01	Ohm	
R_d	[50 50]	Ohm	[TX RX]
PKG_NAME	PKG_HiR_CLASSB PKG_Module		TX RX
A_v	0.413	V	
A_fe	0.413	V	
A_ne	0.608	V	
z_p select	[4]		
L	4		
M	32		
filter and Eq			
f_r	0.58	*fb	
c(0)	0.55		min
c(-1)	0		[min:step:max]
c(-2)	0		[min:step:max]
c(-3)	0		[min:step:max]
c(-4)	0		[min:step:max]
c(1)	0		[min:step:max]
N_b	1	UI	
b_max(1)	0.85		As/dffe1
b_max(2..N_b)	0.3		As/dfe2..N_b
b_min(1)	0		As/dffe1
b_min(2..N_b)	-0.15	S	As/dfe2..N_b
g_DC	[-20:1:0]	dB	[min:step:max]
f_z	42.50	GHz	
f_p1	42.50	GHz	
f_p2	106.25	GHz	
g_DC_HP	[-6:1:0]		[min:step:max]
f_HP_PZ	1.328125	GHz	
Butterworth	1	logical	include in fr

I/O control		
DIAGNOSTICS	0	logical
DISPLAY_WINDOW	0	logical
CSV_REPORT	0	logical
RESULT_DIR	.\results\C2M_{date}\	
SAVE_FIGURES	0	logical
Port Order	[1 3 2 4]	
RUNTAG	C2M TP1a_COM_model	
COM_CONTRIBUTION	1	logical
TDR and ERL options		
TDR	1	logical
ERL	1	logical
ERL_ONLY	0	ns
TR_TDR	0.01	
N	4000	logical
TDR_Butterworth	1	
beta_x	0	
rho_x	0.618	
TDR_W_TXPKG	0	UI
N_bx	20	
fixture delay time	[0 0]	
Tukey_Window	1	
Noise, jitter		UI
sigma_RJ	0.01	UI
A_DD	0.02	V^2/GHz
eta_0	1.25E-08	dB
SNR_TX	33	
R_LM	0.95	
DER_CDR	1.00E-02	
ENOB	0	
adc_clip_rate	5.00E-05	
trunc	128	
N_tc	128	
BREAD_CRUMBS	1	logical
baseline		
new		
relevant for RxFFE		
adjusted in experiment		

Table 93A-3 parameters			
Parameter	**Setting**	**Units**	Information
package_tl_gamma0_a1_a2	[5e-4 0.00065 0.0003]		
package_tl_tau	0.006141	ns/mm	
package_Z_c	92 ; 70 70; 80 80; 100	Ohm	
z_p (TX)	1 1 1 1; 1 1 1 1 ; 0.5	mm	[test cases to run]
z_p (NEXT)	1 1 1 1; 1 1 1 1 ; 0.5	mm	[test cases]
z_p (FEXT)	1 1 1 1; 1 1 1 1 ; 0.5	mm	[test cases]
z_p (RX)	1 1 1 1; 1 1 1 1 ; 0.5	mm	[test cases]
C_p	[0.4e-4 0.4e-4]	nF	[test cases]
Operational			
ERL Pass threshold	10	dB	
COM Pass threshold	3	db	
DER_0	2.50E-05		
T_r	4.00E-03	ns	
FORCE_TR	1	logical	
PMD_type	C2C		
T_O	0	mUI	
samples_for_C2M	100	samples/UI	
EW	0		
MLSE	3	logical	
ts_anchor	1		
sample_adjustment	[-32 32]		
Local Search	0		
Filter: RxFFE			
ffe_pre_tap_len	5	UI	
ffe_post_tap_len	12	UI	
ffe_pre_tap1_max	1	(normalized)	
ffe_post_tap1_max	1	(normalized)	
ffe_tapn_max	1	(normalized)	FV-LMS or MMSE
FFE_OPT_METHOD	MMSE		
num_ui_RXFF_noise	2048		FOM or ISI
RX FFE FLOAT CTL	FOM		
Floating Tap Control			
N_bg	2	0 1 2 or 3 groups	
N_bf	4	taps per group	
N_f	50	UI span for floating taps	
bmaxg	0.2	max DFE value for floating taps	
B_float_RSS_MAX	1	rss tail tap limit	
N_tail_start	13	(UI) start of tail taps limit	

SAVE_CONFIG2MAT	0	
Receiver testing		
RX_CALIBRATION	0	logical
Sigma BBN step	5.00E-03	V
ICN parameters		
f_v	0.278	Fb
f_f	0.278	Fb
f_n	0.278	Fb
f_2	61.625	GHz
A_ft	0.450	V
A_nt	0.450	V
Parameter	**Setting**	
board_tl_gamma0_a1_a2	*[0 6.44084e-4 3.6036e-05]*	1.4 db/in @ 53.125G
board_tl_tau	*5.790E-03*	ns/mm
board_Z_c	100	Ohm
z_bp (TX)	*32*	mm
z_bp (NEXT)	*32*	mm
z_bp (FEXT)	*32*	mm
z_bp (RX)	*32*	mm
C_0	[0.2e-4 0]	nF
C_1	[0.2e-4 0]	nF
Include PCB	**0**	**logical**
Seletions (rectangle, gaussian,dual_rayleigh,triangle		
Histogram_Window_Weight	gaussian	selection
Qr	*0.02*	UI

Test Channels Info

Channel #	Channel Source
1	https://www.ieee802.org/3/dj/public/tools/CR/lim_3dj_03_230629.zip
2	https://www.ieee802.org/3/dj/public/tools/CR/lim_3dj_04_230629.zip
3 – 7	https://www.ieee802.org/3/dj/public/tools/CR/kocsis_3dj_02_2305.zip
8 – 34	https://www.ieee802.org/3/dj/public/tools/KR/mellitz_3dj_02_elec_230504.zip
35 – 40	https://www.ieee802.org/3/dj/public/tools/CR/shanbhag_3dj_01_2305.zip
41 – 44	https://www.ieee802.org/3/dj/public/tools/KR/shanbhag_3dj_02_2305.zip
45 – 80	https://www.ieee802.org/3/dj/public/tools/KR/weaver_3dj_02_2305.zip
81 – 88	https://www.ieee802.org/3/dj/public/tools/KR/weaver_3dj_elec_01_230622.zip
89	https://www.ieee802.org/3/dj/public/tools/CR/lim_3dj_07_2309.zip
90 – 96	https://www.ieee802.org/3/dj/public/tools/KR/akinwale_3dj_01_2310.zip
97 – 100	https://www.ieee802.org/3/dj/public/tools/CR/akinwale_3dj_02_2311.zip
101 – 112	https://www.ieee802.org/3/dj/public/tools/CR/weaver_3dj_02_2311.zip