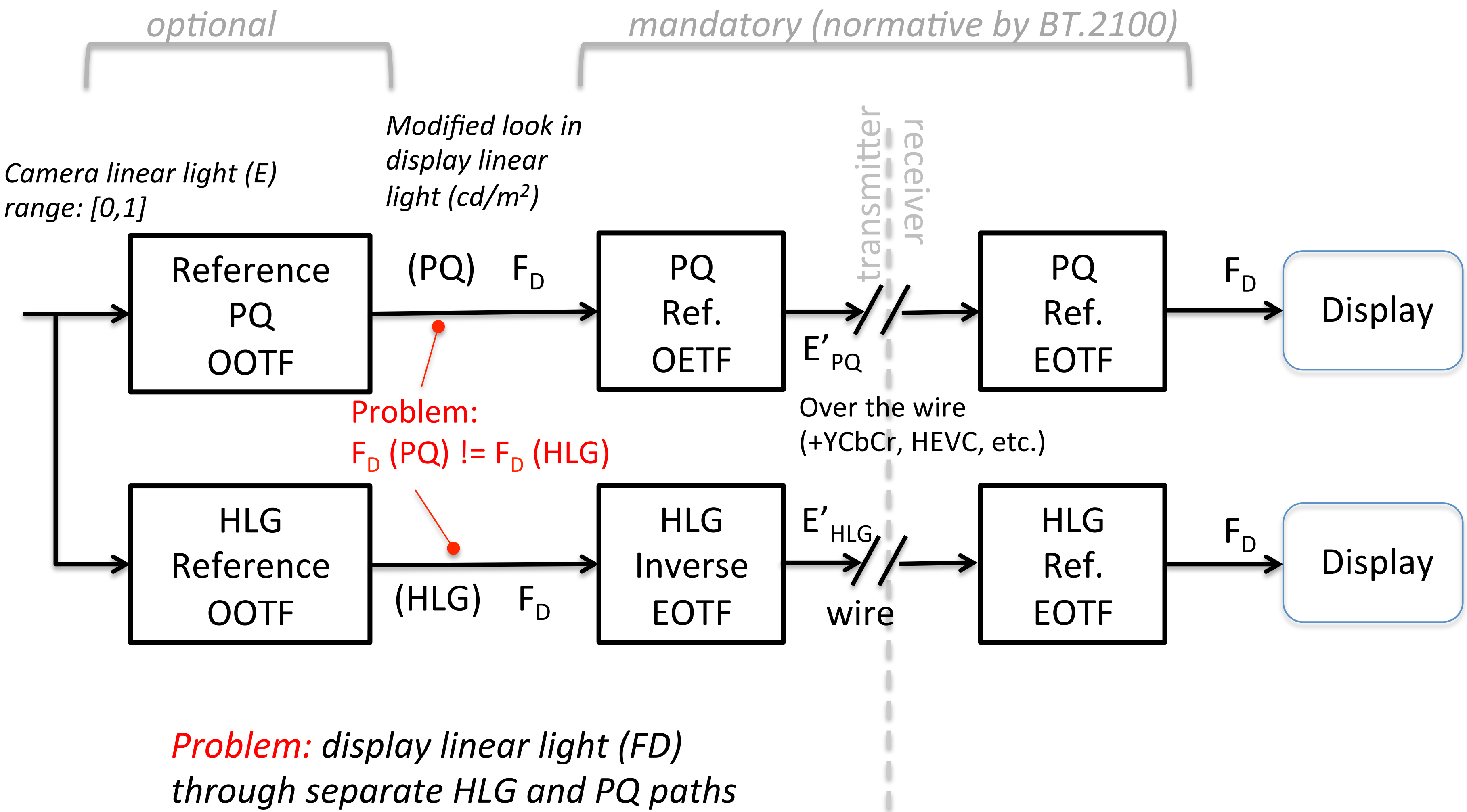
BT.2100 PQ and HLG conversions: towards fully invertible end-to-end signal conversion.

MovieLabs

chadfogg@gmail.com

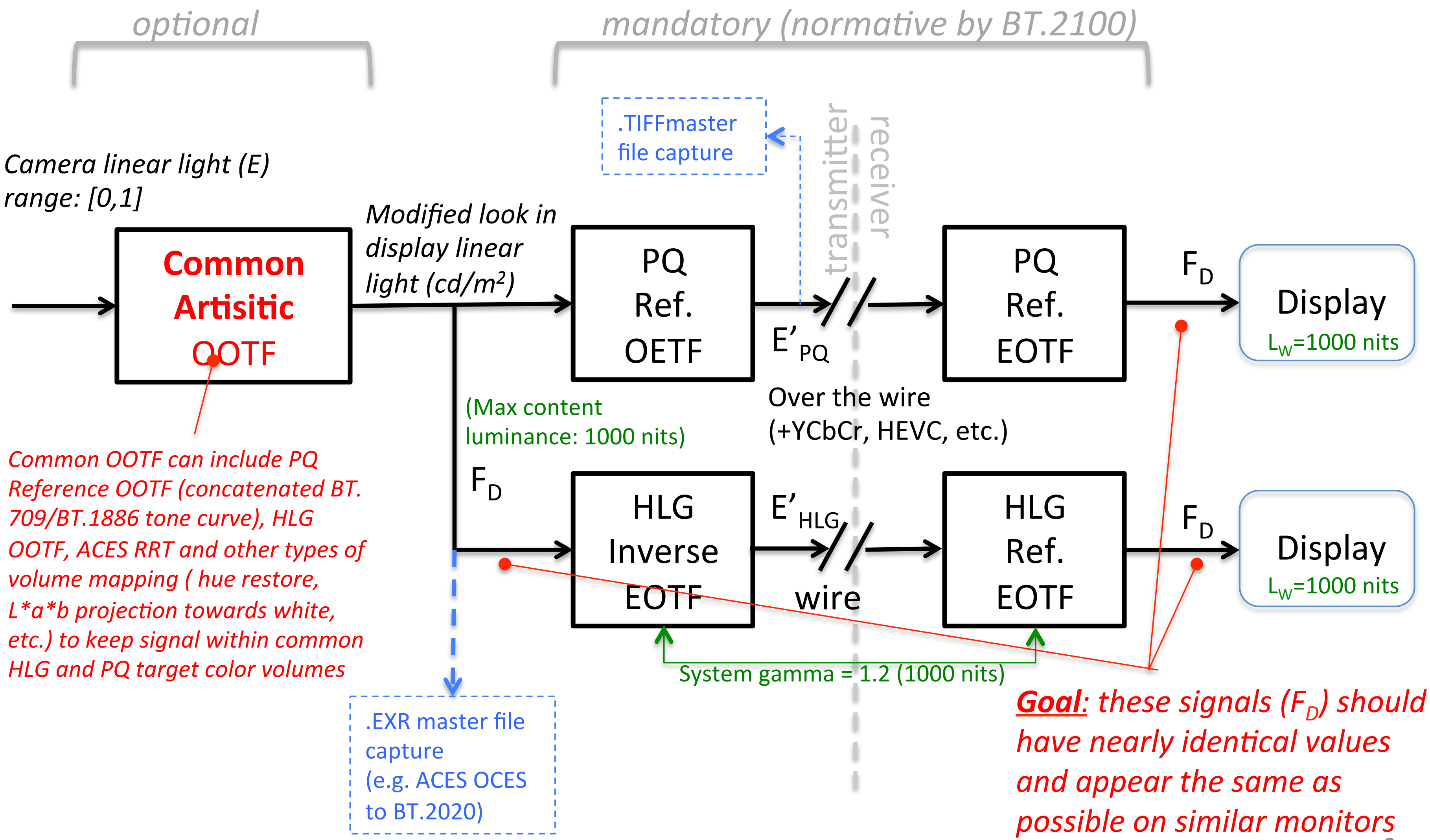
Case 1: Camera to simultaneous PQ,HLG



Problem: display linear light (FD) through separate HLG and PQ paths are not likely to appear the same

Case 1: Solution?

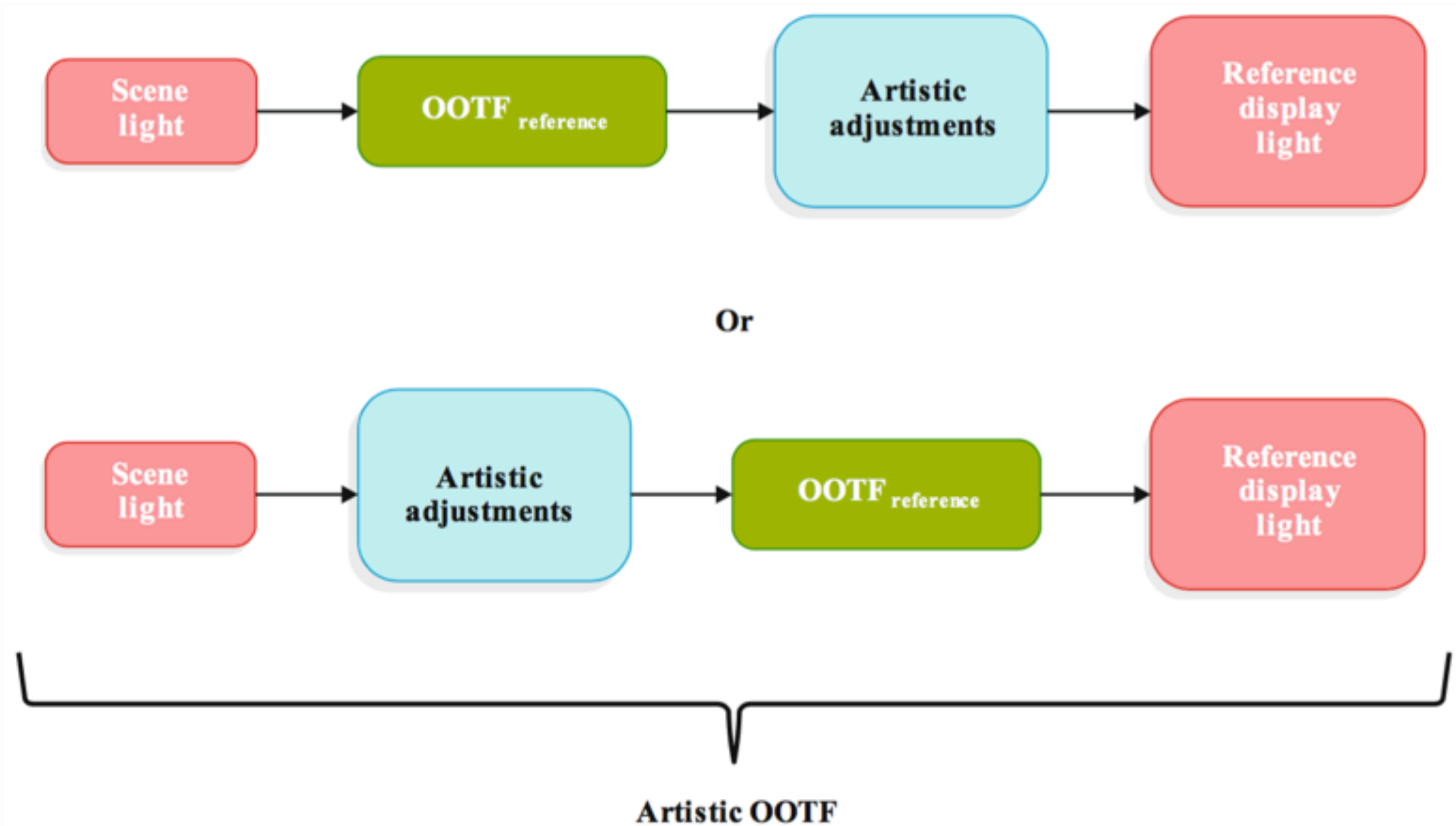
common OOTF is necessary before the OETF



Role of the OOTF: apply a look

(input is scene linear light [0,1], output is display light in cd/m_2)

from Annex 1 of ITU-R BT.2100 (July 2016)



OOTF can include artistic adjustments (grading, tone mapping, etc.) such as ACES, re-grades, etc. before or after the Reference OOTF.

Math

- The next set of slides document the formulae applied verbatim from BT.2100 July 2016 publication. BT.2100 July 2016 has all the necessary information, though much of it is obscure.
- Behind all derivations in this document is: the reversibility principle.
- Tested to be invertible end-to-end: F_D <-> E <-> E' <-> E <-> F_D for 1000 nits and 4000 nits graded content.
- most important rule: **F_D (display linear light) should have the same value everywhere in the processing chain, in all signal directions.**
- Content cannot exceed L_W associated with system gamma γ from BT.2100 NOTE 5e.
- Only the HLG Inverse EOTF is interpreted from natural language text (BT.2100 Annex 2), but formula is confirmed by BBC.
- All formulae assume signals have nominal range [0,1], except where noted (such as F_D in units of cd/m^2).
- ARIB STD-B67 is just the HLG OETF,
  - http://www.arib.or.jp/english/html/overview/doc/2-STD-B67v1_0.pdf
- ..while BT.2100 describes the HLG EOTF and OOTF that complement ARIB B67:
  - http://www.itu.int/rec/R-REC-BT.2100-0-201607-I/en
- HLG Inverse Reference EOTF supplied by BBC:
  - https://gitlab.com/standards/HDRTools/blob/master/common/src/DisplayGammaAdjustHLG.cpp

[Mandatory] Reference PQ EOTF

$$F_D = \text{EOTF}[E'] = 10000\,Y$$

$$Y = \left(\frac{\max\left[\left(E'^{1/m_2} - c_1\right), 0\right]}{c_2 - c_3 E'^{1/m_2}} \right)^{1/m_1}$$

where:

E' denotes a non-linear colour value $\{R', G', B'\}$ or $\{L', M', S'\}$ in PQ space [0,1]

F_D is the luminance of a displayed linear component $\{R_D, G_D, B_D\}$ or Y_D or I_D, in cd/m^2. [4b]

So that when $R'=G'=B'$, the displayed pixel is achromatic.

Y denotes the normalized linear colour value, in the range [0:1]

$m_1 = 2610/16384 = 0.1593017578125$

$m_2 = 2523/4096 \times 128 = 78.84375$

$c_1 = 3424/4096 = 0.8359375 = c_3 - c_2 + 1$

$c_2 = 2413/4096 \times 32 = 18.8515625$

$c_3 = 2392/4096 \times 32 = 18.6875$

Table 4 from ITU-R BT. 2100 (July 2016)

[Mandatory] Reference PQ OETF

$$E' = \mathrm{OETF}[E] = \mathrm{EOTF}^{-1}[\mathrm{OOTF}[E]] = \mathrm{EOTF}^{-1}[F_D]$$

where

$$\mathrm{EOTF}^{-1}[F_D] = \left(\frac{c_1 + c_2 Y^{m_1}}{1 + c_3 Y^{m_1}} \right)^{m_2}$$

$$Y = F_D / 10000$$

E' is the resulting non-linear signal (R', G', B') in the range [0:1]

F_D, E, are as specified in the opto-optical transfer function

m_1, m_2, c_1, c_2, c_3 are as specified in the electro-optical transfer function

from Table 4 from ITU-R BT.2100 (July 2016)

[Mandatory] HLG Reference OOTF

Rewritten Table 5 from ITU-R BT.2100 (July 2016) for nominal range [0,1]

$$R_D = \alpha \cdot Y_S^{\gamma-1} \cdot R_S + \beta$$

$$G_D = \alpha \cdot Y_S^{\gamma-1} \cdot G_S + \beta$$

$$B_D = \alpha \cdot Y_S^{\gamma-1} \cdot B_S + \beta$$

$$Y_S = 0.2627 \cdot R_S + 0.6780 \cdot G_S + 0.0593 \cdot B_S$$

Where:
F_D is the luminance of a display linear component {R_D, G_D, or B_D}, in cd/m^2.
E is the signal for each color component {R_S, G_S, B_S} proportional to scene linear light and scaled by camera exposure, normalized to range [0:1]

Y_S is the normalized linear scene luminance

α, β, γ are defined for the EOTF.

[Mandatory] HLG Reference EOTF

$$F_D = OOFT[E] = \alpha \cdot Y_S^{\gamma-1} \cdot E + \beta$$

$$R_D = \alpha \cdot Y_S^{\gamma-1} \cdot R_S + \beta$$

$$G_D = \alpha \cdot Y_S^{\gamma-1} \cdot G_S + \beta$$

$$B_D = \alpha \cdot Y_S^{\gamma-1} \cdot B_S + \beta$$

$$Y_S = 0.2627 \cdot R_S + 0.6780 \cdot G_S + 0.0593 \cdot B_S$$

$$E = OETF^{-1}[E'] = \begin{cases} E'^2/3 & 0 \le E' \le 1/2 \\ \exp((E'-c)/a)+b & 1/2 < E' \end{cases}$$

NOTE 5e: system gamma (γ) = 1.2 + 0.42*Log10(L_W / 1000)

Where:

R_S, G_S, B_S are the scene linear light signals, *E*, for each color component normalized in the range [0:1].

$\alpha = L_W - L_B$ $\beta = L_B$

F_D is the luminance of a displayed linear component { R_D, G_D, B_D } in cd/m²

E' is the non-linear signal {*R'*, *G'*, *B'*} as defined for the OETF

R_D, G_D, B_D are the displayed light for each color component, in cd/m²

a=0.17883277, b=0.02372241, c=1.00429347 from BT.2100 Note 5b

L_W is the nominal peak luminance of the display in cd/m²

L_B is the nominal luminance for black in cd/m²

The nominal signal range of *E*, R_S, G_S, B_S, and Y_S is [0,1]

Rewritten Table 5 from ITU-R BT. 2100 (July 2016) for nominal range [0,1]

[Mandatory] HLG OETF and Inverse OETF

This presentation assumes HLG OETF normalized to [0,1] range as per Note 5b

NOTE 5b – If E is normalized to the range [0:1] then the equivalent equation for the OETF is:

$$E' = \mathrm{OETF}[E] = \begin{cases} \sqrt{3E} & 0 \le E \le 1/12 \\ a \cdot \ln(E - b) + c & 1/12 < E \end{cases}$$

where $a = 0.17883277$, $b = 0.02372241$, $c = 1.00429347$

This presentation assumes HLG Inverse OETF normalized to [0,1] range as per Note 5c

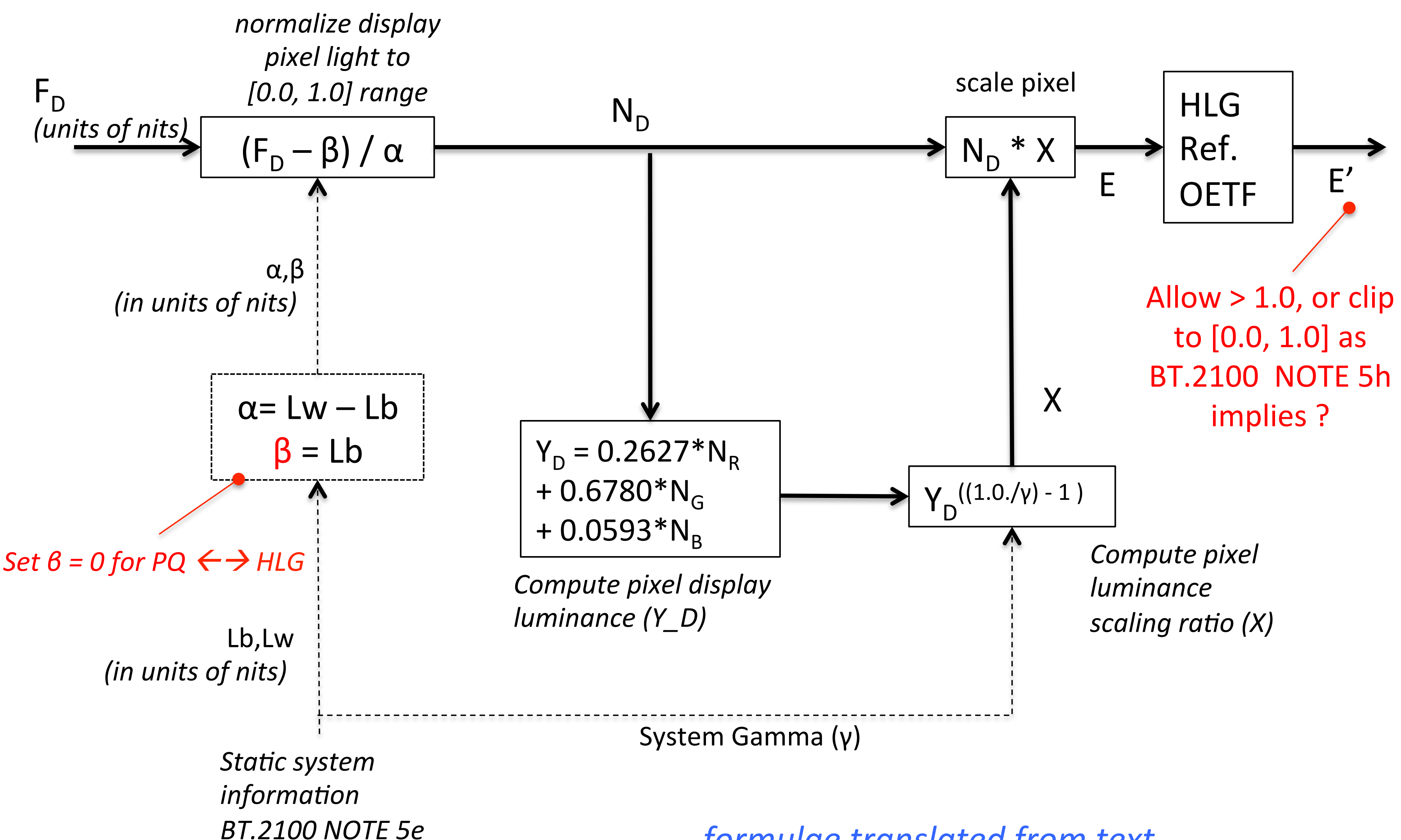
NOTE 5c – If E is normalized to the range [0:1] then the equivalent equation for the E is:

$$E = \mathrm{OETF}^{-1}[E'] = \begin{cases} E'^2/3 & 0 \le E' \le 1/2 \\ \exp((E' - c)/a) + b & 1/2 < E' \end{cases}$$

where a, b and c are as defined in Note 5b.

from NOTES for Table 5 from ITU-R BT.2100 (July 2016)

[Mandatory for Case 1] **HLG Inverse OETF**



formulae translated from text instructions of Annex 2 of ITU-R BT.2100 (July 2016)

[Optional] Reference PQ OOTF

$F_D = \mathrm{OOTF}[E] = \mathrm{G}_{1886}\,[\mathrm{G}_{709}[E]]$

where

$E = \{R_S, G_S, B_S; Y_S;$ or $I_S\}$ is the signal determined by scene light and scaled by camera exposure

E' is a non-linear representation of E

F_D is the luminance of a displayed linear component (R_D, G_D, B_D; Y_D; or I_D)

The values E, R_S, G_S, B_S, Y_S, I_S are in the range [0:1]

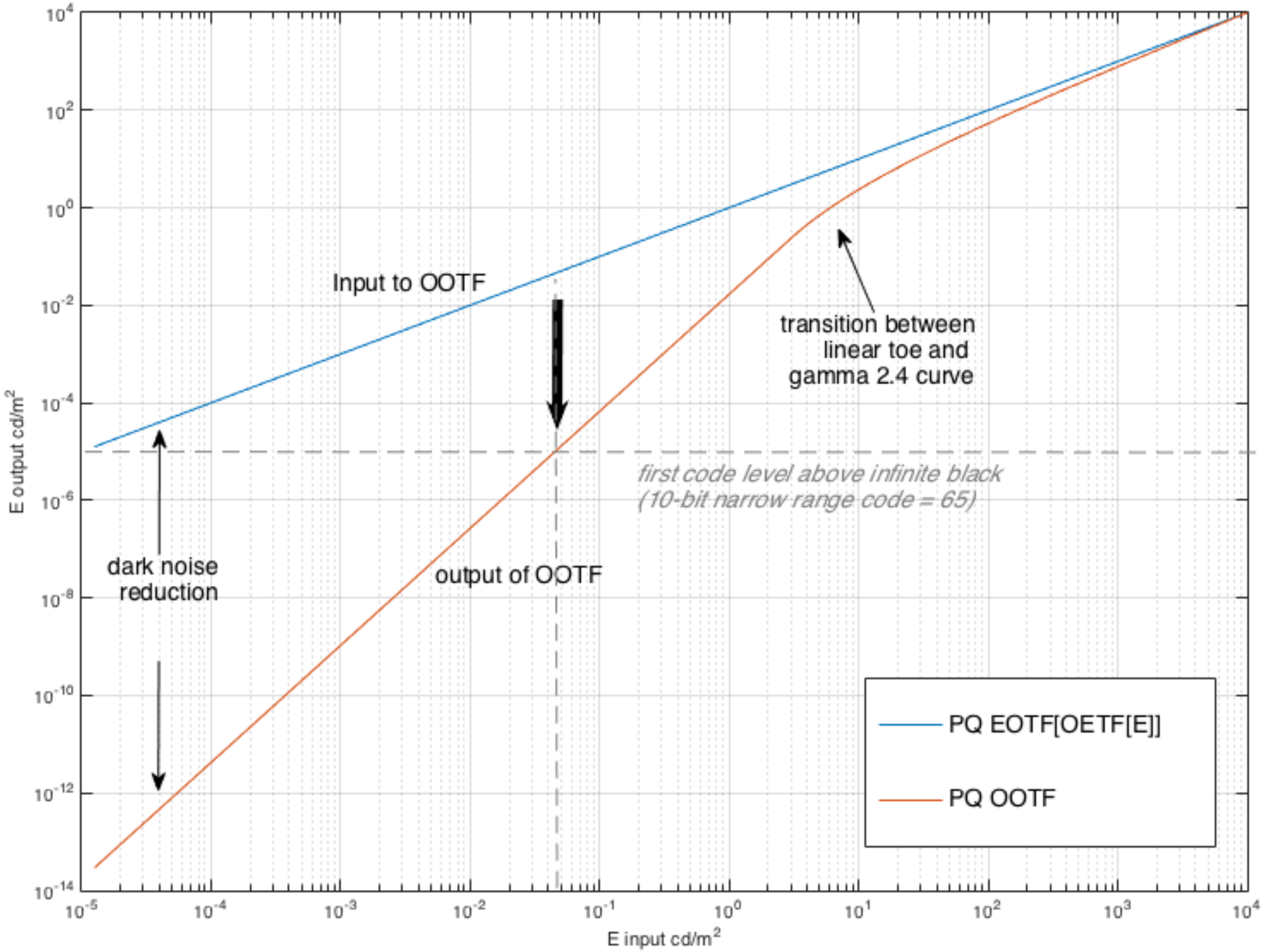
$F_D = \mathrm{G}_{1886}\,[\,\mathrm{G}_{709}[E]] = \mathrm{G}_{1886}\,E'$

$$E' = \mathrm{G}_{709}[E] = 1.099\,(59.5208\,E)^{0.45} - 0.099 \quad \text{for } 1 > E > 0.0003024$$
$$= 267.84\,E \quad \text{for } 0.0003024 \geq E \geq 0$$

$F_D = \mathrm{G}_{1886}[E'] = 100\,E'^{2.4}$

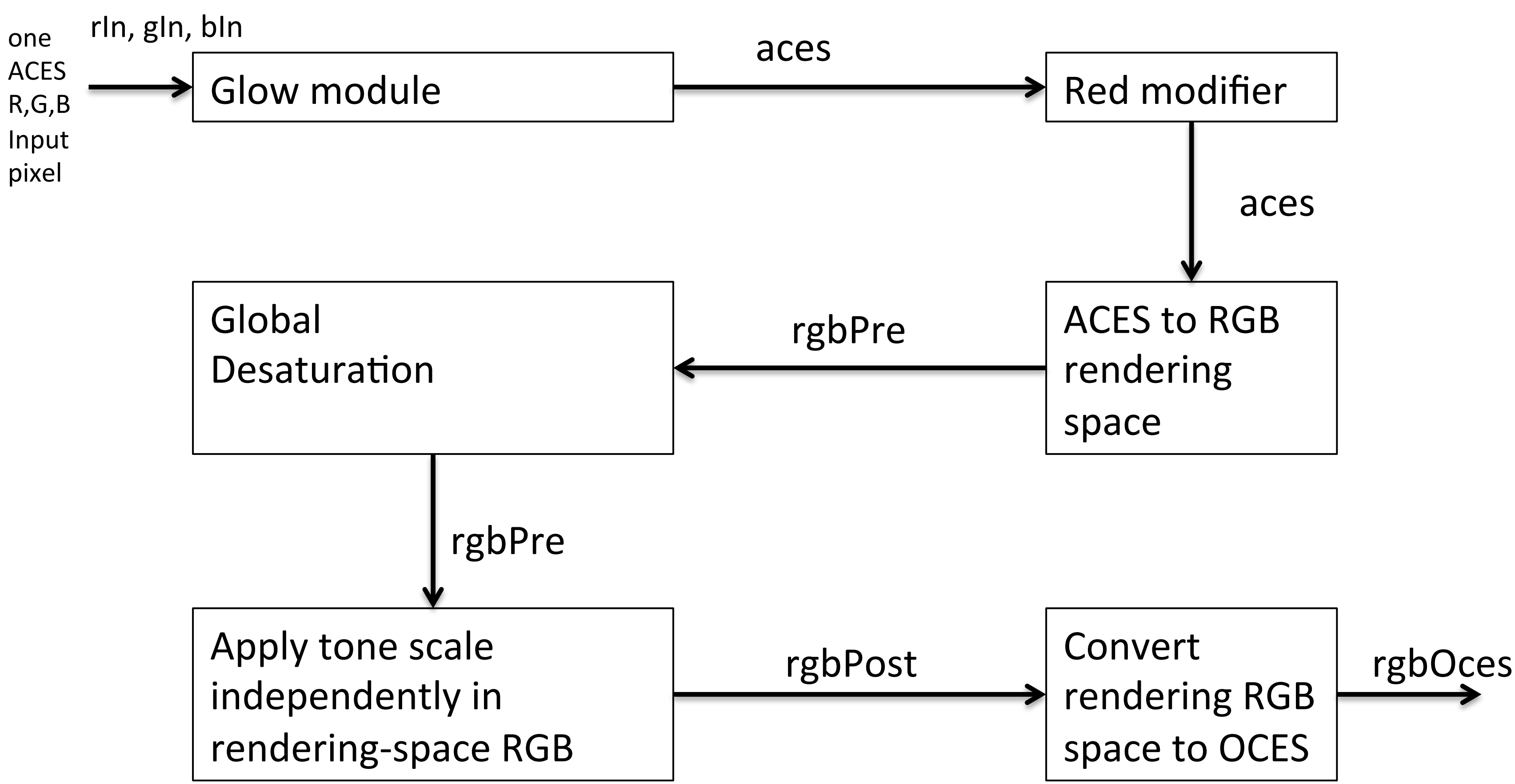
from Table 4 from ITU-R BT.2100 (July 2016)

[optional] PQ OOTF (plotting BT.2100 Table 4)



[Optional] Example artistic OOTF

https://github.com/ampas/aces-dev/blob/master/transforms/ctl/rrt/RRT.a1.0.1.ctl



Color conversion to Y'CbCr

TABLE 6

Non-Constant Luminance $Y'C'_BC'_R$ signal format[6a]

Parameter	Values PQ	Values HLG
Derivation of R', G', B'	$\{R', G', B'\}=\mathrm{EOTF}^{-1}(F_D)$ where $F_D = \{R_D, G_D, B_D\}$	$\{R', G', B'\}=\mathrm{OETF}(E)$ where $E = \{R_S, G_S, B_S\}$
Derivation of Y'	$Y' = 0.2627R' + 0.6780G' + 0.0593B'$	
Derivation of colour difference signals	$C'_B = \dfrac{B'-Y'}{1.8814}$ $C'_R = \dfrac{R'-Y'}{1.4746}$	

NOTE 6a – For consistency with prior use of terms, Y', C'_B and C'_R employ prime symbols indicating they have come from non-linear Y, B and R.

Alternatively, the above can be expressed in matrix form:

$$Y'CbCr = \begin{bmatrix} 0.2627 & 0.6780 & 0.0593 \\ -0.1396 & -0.3604 & 0.5000 \\ 0.5000 & -0.4598 & -0.0402 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix}$$

inverse

$$R'G'B' = \begin{bmatrix} 1.0000 & -0.0000 & 1.4746 \\ 1.0000 & -0.1646 & -0.5714 \\ 1.0000 & 1.8814 & 0 \end{bmatrix} \begin{bmatrix} Y' \\ Cb \\ Cr \end{bmatrix}$$

The whole of Annex 2 of ITU-R BT.2100 (July 2016)

Annex 2
(Informative)

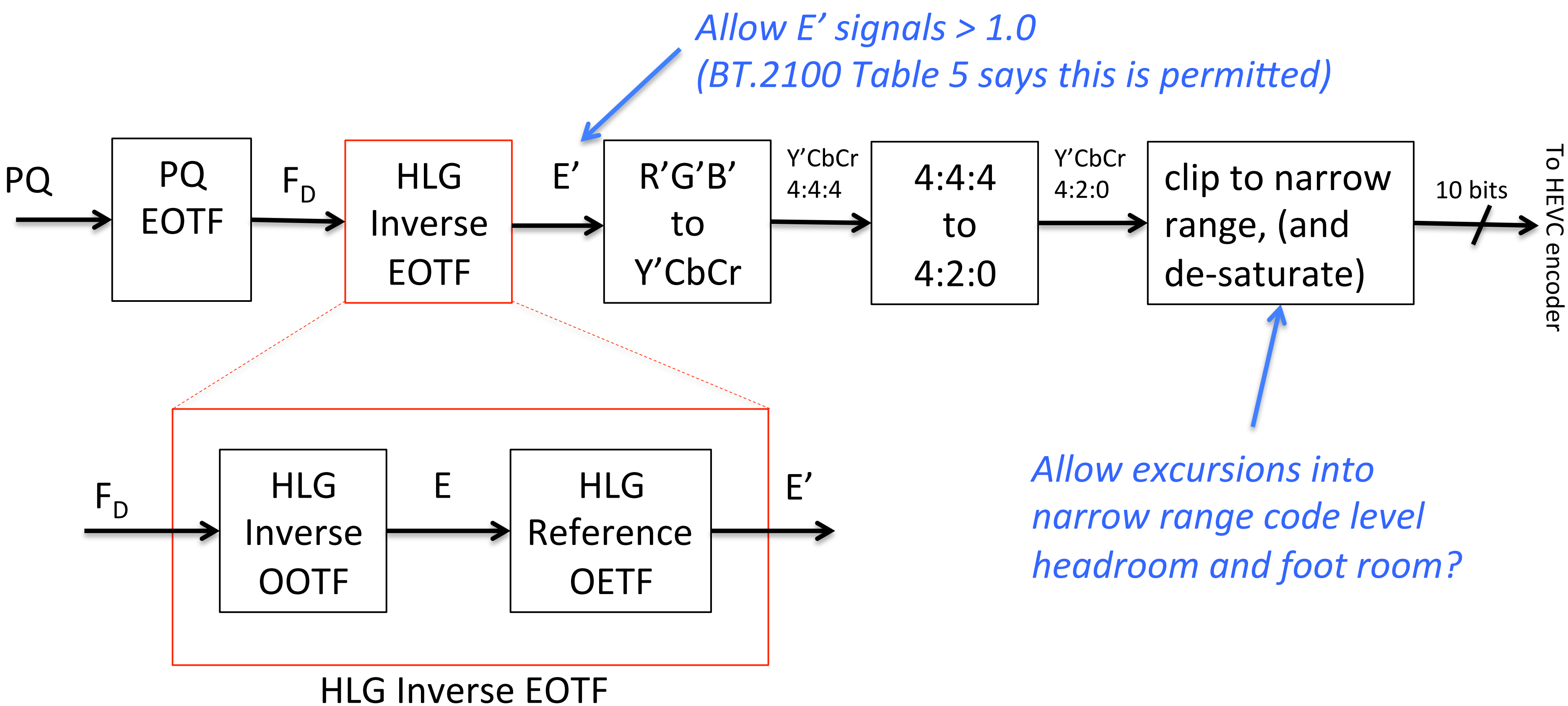
Conversion between HLG and PQ signals

The following diagram illustrates conversion from the PQ signal to the HLG signal. The signal processing is that the PQ signal is decoded by the PQ EOTF to yield a signal that represents linear display light. This signal is then encoded by the HLG inverse EOTF to produce an equivalent HLG signal. When this HLG signal is subsequently decoded by the HLG EOTF in the display, the result will be the same display light that would be produced by decoding the original PQ signal with the PQ EOTF. The HLG inverse EOTF is the HLG inverse OOTF followed by the HLG OETF. For the HLG inverse OOTF, black level should be zero, and the gamma parameter is determined by the peak level of the PQ signal.

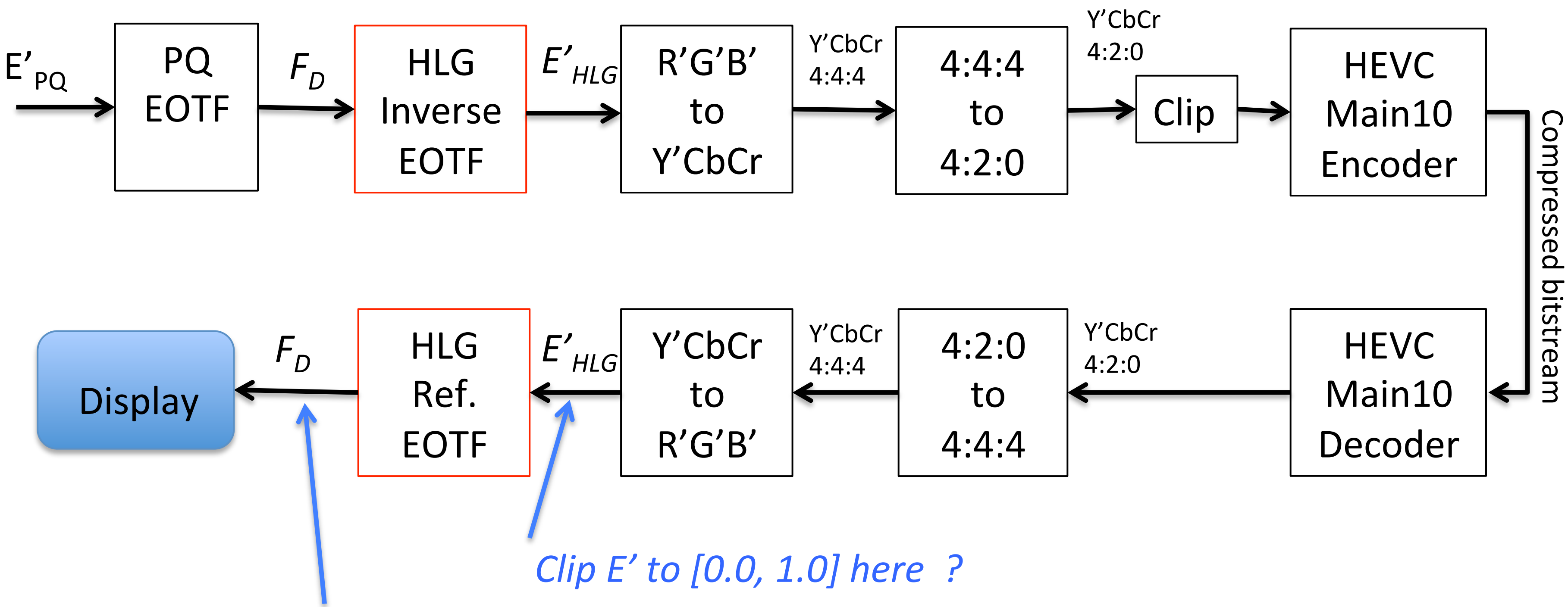
BT.2100-Ann2-01

The following diagram illustrates conversion from the HLG signal to the PQ signal. The signal processing is that the HLG signal is decoded by the HLG EOTF to yield a signal that represents linear display light. This signal is then encoded by the PQ inverse EOTF to produce an equivalent PQ signal. When this PQ signal is subsequently decoded by the PQ EOTF in the display, the result will be the same display light that would be produced by decoding the original HLG signal with the HLG EOTF. For the HLG EOTF, black level should be zero, and the gamma may be set to the value specified in Table 5 (presuming peak luminance of 1 000 cd/m^2).

Case 2: PQ to HLG



..or only clamp E'->F_D only before display as the BT.2100 spec. suggests



Clip E' to [0.0, 1.0] here ?

..or clip F_D output to [0, L_W] here ?
(this is the current recommended solution from BBC)

PQ to HLG

Problem: HLG Reference OOTF normalizes scene signal (E) components (R_S,G_S,B_S) by system gamma-scaled scene luminance (Y_S)

$$F_D = \mathrm{OOTF}[E] = \alpha Y_S^{\gamma-1} E + \beta$$
$$R_D = \alpha Y_S^{\gamma-1} R_S + \beta$$
$$G_D = \alpha Y_S^{\gamma-1} G_S + \beta$$
$$B_D = \alpha Y_S^{\gamma-1} B_S + \beta$$
$$Y_s = 0.2627 R_S + 0.6780 G_S + 0.0593 B_S$$

BT.2100 Table 5: HLG Reference OOTF

…which tends to compresses the scene referred input signal (Y_S).

..but if the source is display-referred…

… instead of camera scene-referred as assumed by the normal HLG model, then the *Inverse* Reference EOTF that converts display light (F_D) to non-linear scene light (E') ends up expanding the signal by a ratio determined the inverse of the luminance system gamma pixel scale factor in the reference EOTF

The inverse of the Reference EOTF scale factor $Y_S^{\gamma-1}$ is: $Y_D^{((1.0./\gamma)\ -\ 1\ )}$

Extreme case: for max blue (F_B = 1.0, F_R = F_G = 0.0), Y_D = 0.0593 (the ~6% BT.2020 blue contribution to CIE-Y), and the over-gain factor is therefore:

System gamma 1.2000 (Lw = 1000 nits): E_B = **1.6014** = (0.0593)^((1.0/1.2)-1)
System gamma 1.3264 (Lw = 2000 nits): E_B = **2.0042** *(200% overshoot)*
System gamma 1.4529 (Lw = 4000 nits): E_B = **2.4124** *(241% overshoot)*

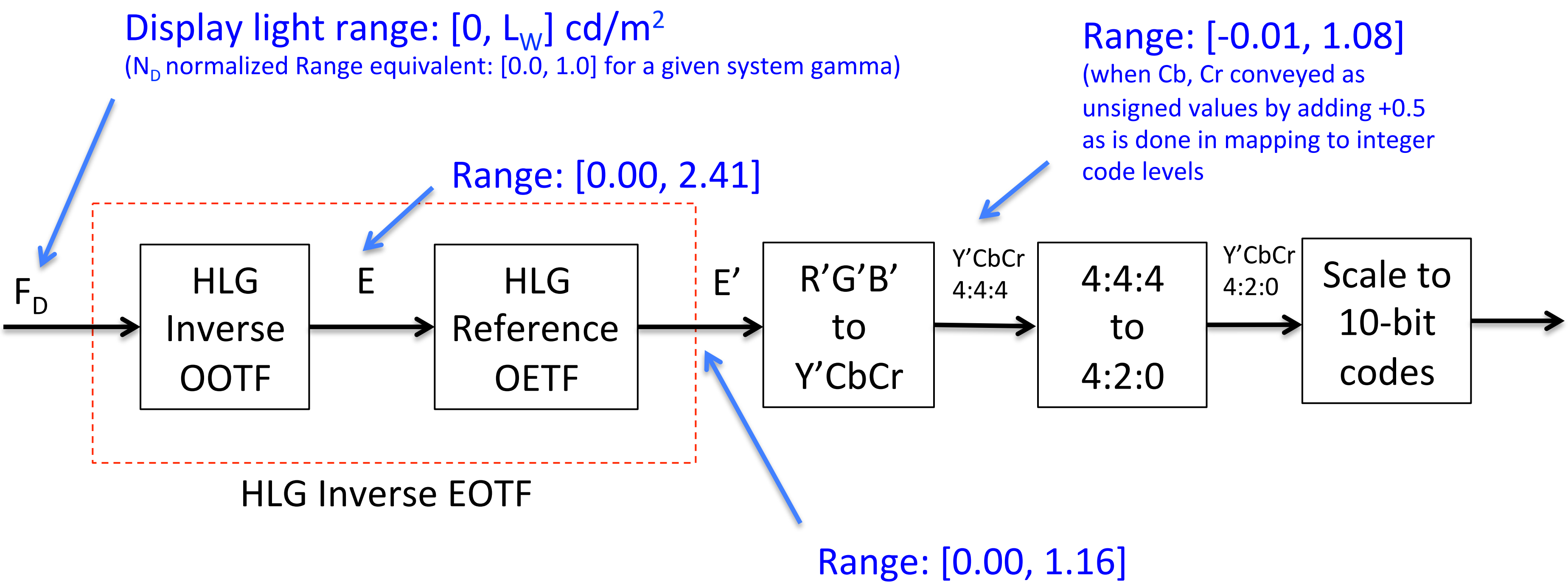
But after conversion to E', and then RGB → Y'CbCr, the corner cases signal is compressed back to almost [0.0, 1.0] range

γ=1.2000: E'_B = Reference_EOTF(1.6014) = **1.085829** *(~9% overshoot)*
γ=1.3264: E'_B = Reference_EOTF(2.0042) = **1.126500** *(~13% overshoot)*
γ=1.4529: E'_B = Reference_EOTF(2.4124) = **1.160008** *(~16% overshoot)*

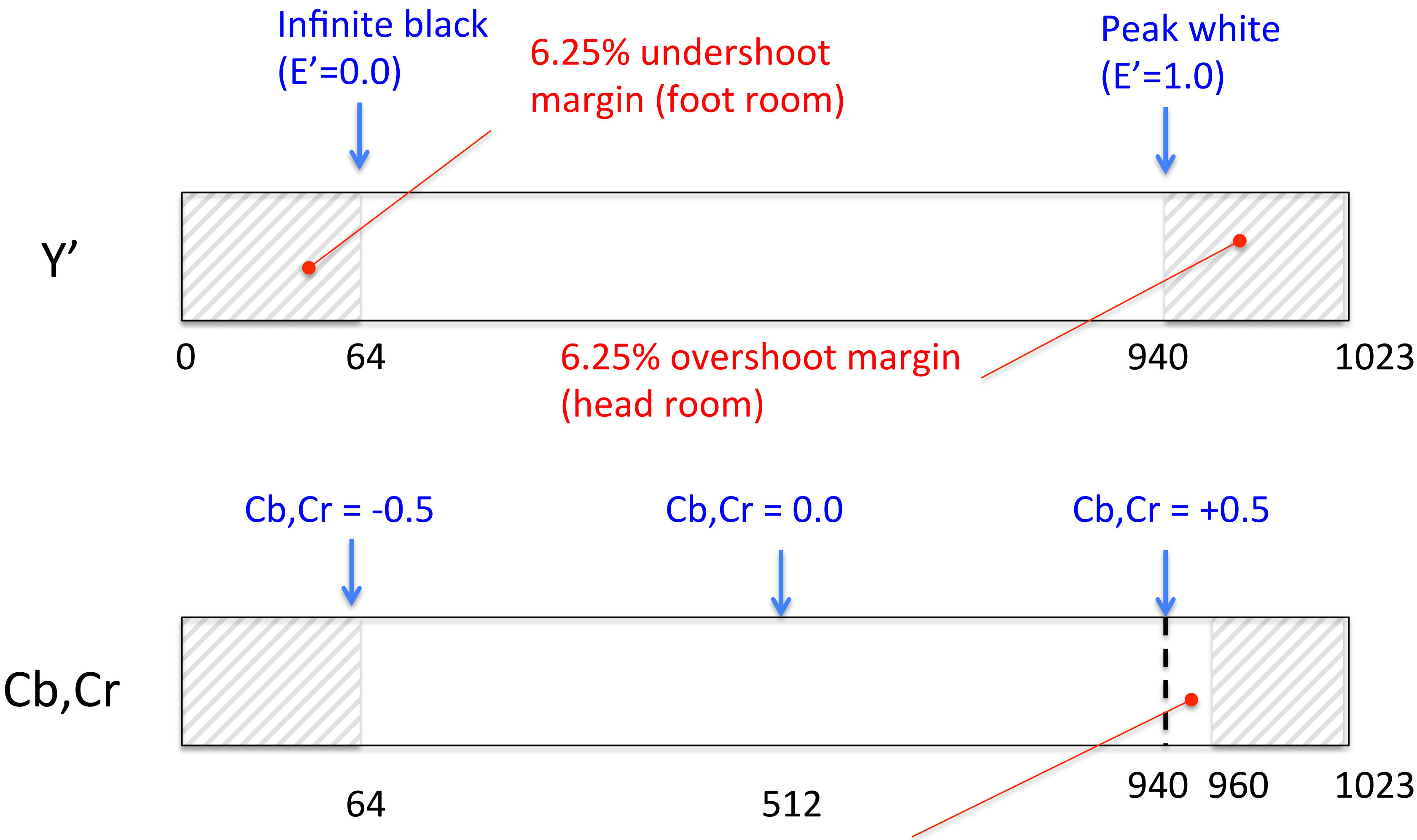
After conversion from R'G'B' to Y'CbCr, the values become:

	Real-value			10-bit narrow range code			
System gamma	Y'	Cb	Cr	Y	Cb	Cr	
γ=1.2000	0.06439	0.542915	-0.043666	120	998	473	*(~4% overshoot)*
γ=1.3264	0.066801	0.56325	-0.045301	123	1017	471	*(~6% overshoot)*
γ=1.4529	0.068789	0.580004	-0.046649	124	1032	470	*(~8% overshoot)*

Range expansion, then contraction



10-bit narrow range codes



A little extra legal range for Cb,Cr positive signals designed by original 1980's ITU-R (CCIR Rec.) BT.601 team to accommodate for saturated colors

Saturated primary color combos

F_D (Input to Inverse EOTF)

E (output of Inverse EOTF)

E’ (output of Reference OETF)

R’G’B’ to Y’CbCr

Scale real-value to 10-bit narrow range integer code level

Color combo	E_R	E_G	E_B	E'_R	E'_G	E'_B	Y’	Cb	Cr	Y Code	Cb code	Cr code
Magenta 1000 nits	1.207881	0	1.207881	1.034522	0	1.034522	0.333116	0.372811	0.475658	356	846	938
Magenta 2000 nits	1.321648	0	1.321648	1.050927	0	1.050927	0.338399	0.378723	0.483201	360	851	945
Magenta 4000 nits	1.423652	0	1.423652	1.064457	0	1.064457	0.342755	0.383598	0.489422	364	856	951
Yellow 1000 nits	1.010241	1.010241	0	1.001866	1.001866	0	0.942455	-0.500933	0.040289	890	63	548
Yellow 2000 nits	1.015158	1.015158	0	1.002755	1.002755	0	0.943292	-0.501378	0.040325	890	63	548
Yellow 4000 nits	1.019238	1.019238	0	1.00349	1.00349	0	0.943983	-0.501745	0.040355	891	62	548
Cyan 1000 nits	0	1.052106	1.052106	0	1.009299	1.009299	0.744156	0.140928	-0.504649	716	638	60
Cyan 2000 nits	0	1.077885	1.077885	0	1.013726	1.013726	0.74742	0.141547	-0.506863	719	639	58
Cyan 4000 nits	0	1.099654	1.099654	0	1.017382	1.017382	0.750115	0.142057	-0.508691	721	639	56

Red marks code values outside 10-bit narrow range [64,940] Y or [64,960] Cb,Cr

..continued

Only one of the RGB-PQ corner cases (Blue) cannot be represented in 10-bits without clamping

Color combo	E_R	E_G	E_B	E'_R	E'_G	E'_B	Y'	Cb	Cr	Y Code	Cb code	Cr code
Red 1000 nits	1.249559	0	0	1.040708	0	0	0.273394	-0.145314	0.520354	303	382	978
Red 2000 nits	1.389536	0	0	1.060045	0	0	0.278474	-0.148014	0.530022	308	379	987
Red 4000 nits	1.516901	0	0	1.075989	0	0	0.282662	-0.15024	0.537994	312	377	994
Green 1000 nits	0	1.066911	0	0	1.011855	0	0.686038	-0.364642	-0.465236	665	185	95
Green 2000 nits	0	1.100358	0	0	1.017499	0	0.689864	-0.366676	-0.467831	668	183	93
Green 4000 nits	0	1.128773	0	0	1.022157	0	0.693023	-0.368355	-0.469973	671	182	91
Blue 1000 nits	0	0	1.601367	0	0	1.085829	0.06439	0.542915	-0.043666	120	998	473
Blue 2000 nits	0	0	2.004236	0	0	1.1265	0.066801	0.56325	-0.045301	123	1017	471
Blue 4000 nits	0	0	2.412375	0	0	1.160008	0.068789	0.580004	-0.046649	124	1032	470

Yellow marks code values outside (full) 10-bit possible range [0,1023]

BT.2100 HLG OETF does not force clamping to [0,1]

NOTE 5b – If E is normalized to the range [0:1] then the equivalent equation for the OETF is:

$$E' = \mathrm{OETF}\left[E\right] = \begin{cases} \sqrt{3E} & 0 \le E \le 1/12 \\ a \cdot \ln\left(E - b\right) + c & 1/12 < E \end{cases}$$

where $a = 0.17883277$, $b = 0.02372241$, $c = 1.00429347$

..and neither does the *HLG Inverse OETF*:

$$E = OETF^{-1}[E'] = \begin{cases} E'^2/3 & 0 \le E' \le 1/2 \\ \exp((E' - c)/a) + b & 1/2 < E' \end{cases}$$

log,exp segments not bounded to 1.0

.. BT.2100 NOTE 5h does suggests E' values > 1.0 or < 0.0 should not be displayed (probably means clipped to [0.0, 1.0], rather than leaving 'undefined' pixels):

NOTE 5h – During production, signal values are expected to exceed the range $E' = [0.0 : 1.0]$. This provides processing headroom and avoids signal degradation during cascaded processing. Such values of E', below 0.0 or exceeding 1.0, should not be clipped during production and exchange. Values exceeding 1.0 should not be shown on reference displays. Values below 0.0 should not be clipped in reference displays (even though they represent "negative" light) to allow the black level of the signal (L_B) to be properly set using test signals known as "PLUGE" (see Recommendation ITU-R BT.814).

Clipping & desaturation: R'G'B'

- Clipping in E (RGB primary) or E' (R'G'B') can use hue restore algorithms in direct RGB space such as dw3:
  - https://github.com/ampas/aces-dev/blob/v0.7.1/transforms/ctl/utilities/transforms-common.ctl

```
int inds[3] = order3( pre_tone[0], pre_tone[1], pre_tone[2]);
float orig_chroma = pre_tone[ inds[0]] - pre_tone[ inds[2]];
float hue_factor = 0;
if (orig_chroma != 0.) hue_factor = ( pre_tone[ inds[1] ] - pre_tone[ inds[2] ]) / orig_chroma;
float new_chroma = post_tone[ inds[0] ] - post_tone[ inds[2] ];
out[ inds[ 0] ] = post_tone[ inds[0] ];
out[ inds[ 1] ] = hue_factor * new_chroma + post_tone[ inds[2] ];
out[ inds[ 2] ] = post_tone[ inds[2] ];
```

Clipping & desaturation: Y’CbCr

- Clipped Y’CbCr signals can restore ~hue by preserving CIECAM’02 style ratio before and after clipping. Example:

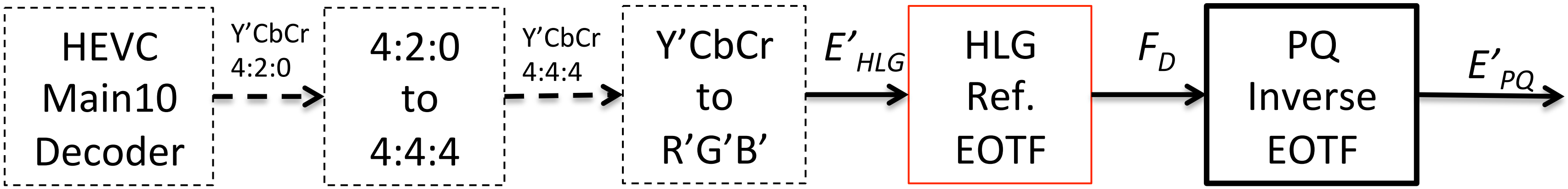
$$Hue = \tan^{-1}\left(\frac{Cr}{Cb}\right)$$

$$Saturation = \sqrt{Cb^2 + Cr^2}$$

Many other methods are possible, including working in more perceptual uniform spaces than YCbCr or RGB.

(note that Cb,Cr are not uniformly perceptual color channels)

Case 3: HLG to PQ

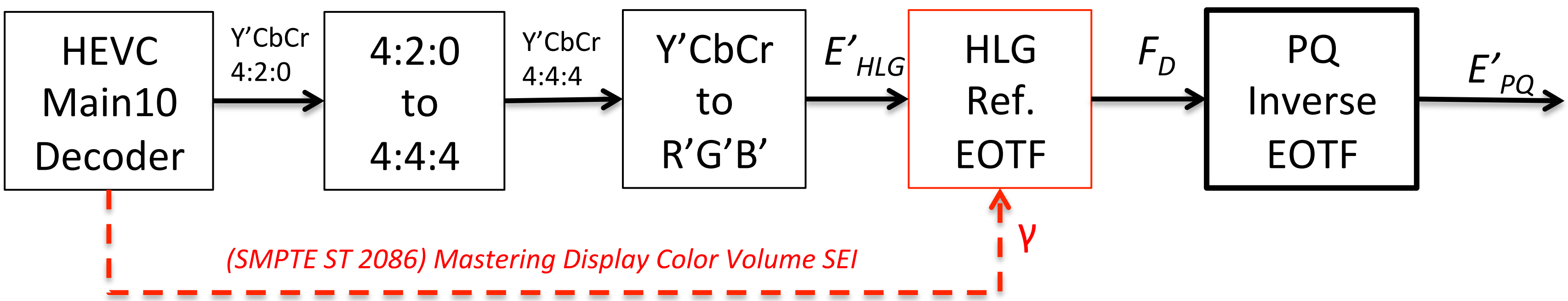


as per BT.2100 Annex 2

HLG proponents say no metadata is needed. This may be true for end-to-end HLG only systems, but for PQ <-> HLG interoperability, metadata (mastering display max luminance) or MaxCLL is needed…

Case 3a: use ST 2086 to guide HLG to PQ ?

To preserve intended rendered look, treat HLG as a display-referred transfer function…



Derive system gamma from SEI?
L_W = max_display_luminance
γ = BT2100_NOTE_5e(L_W)

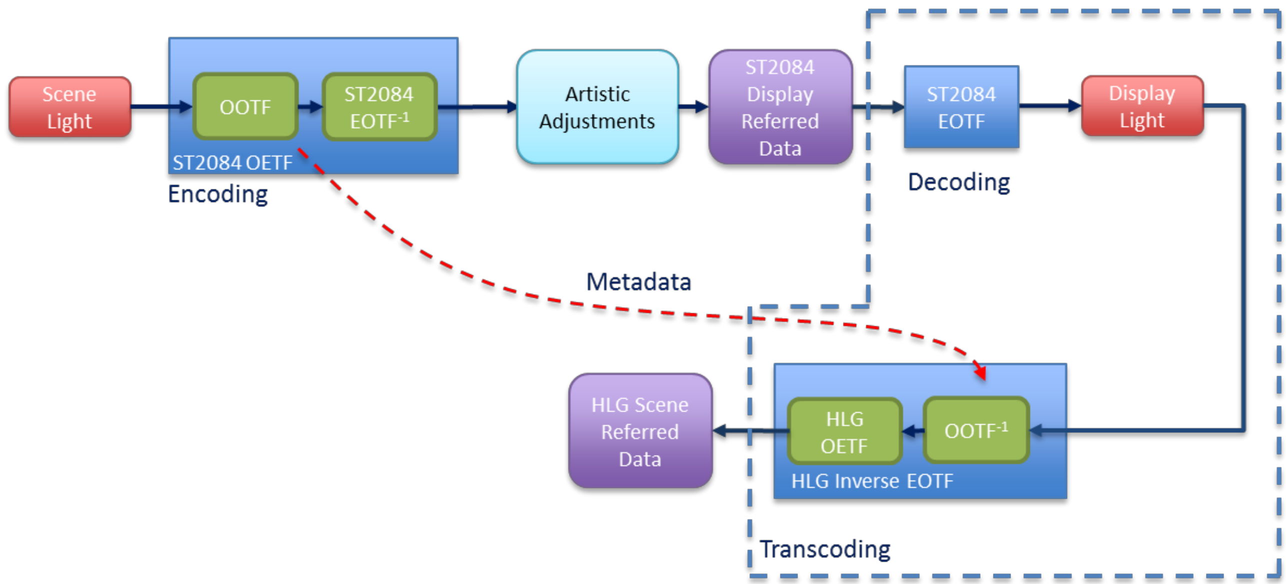
By reversibility criteria, this was implied in slide 25 of m37535_r2.pdf:
http://phenix.int-evry.fr/jct/doc_end_user/current_document.php?id=10317

[see next slide]

Metadata (L_W) needed in HLG→PQ

slide 25 of m37535_r2.pdf:
http://phenix.int-evry.fr/jct/doc_end_user/current_document.php?id=10317

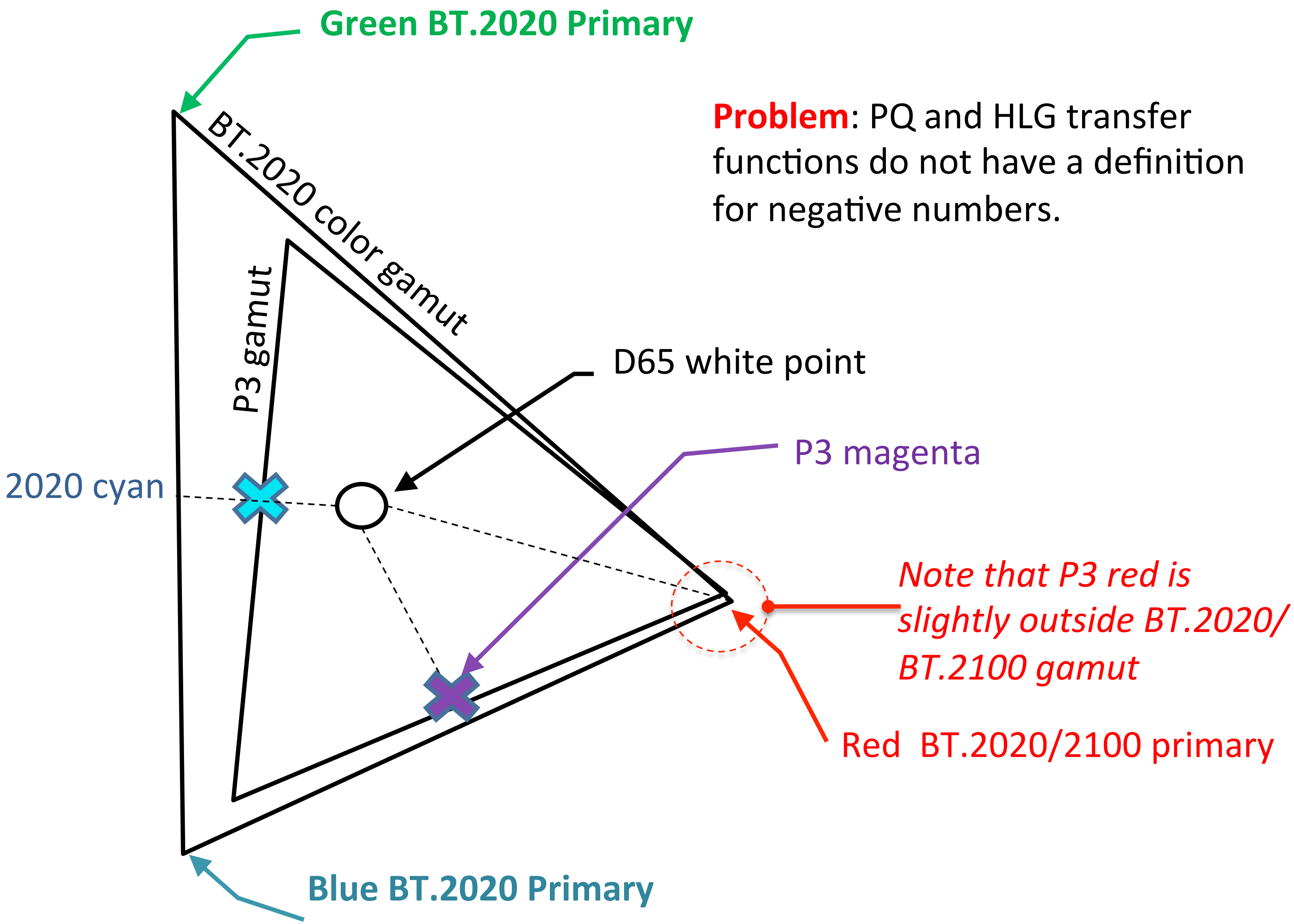
Transcoding from ST 2084 to HLG Therefore Straightforward, but BBC Believes Requires Implicit or Explicit Metadata



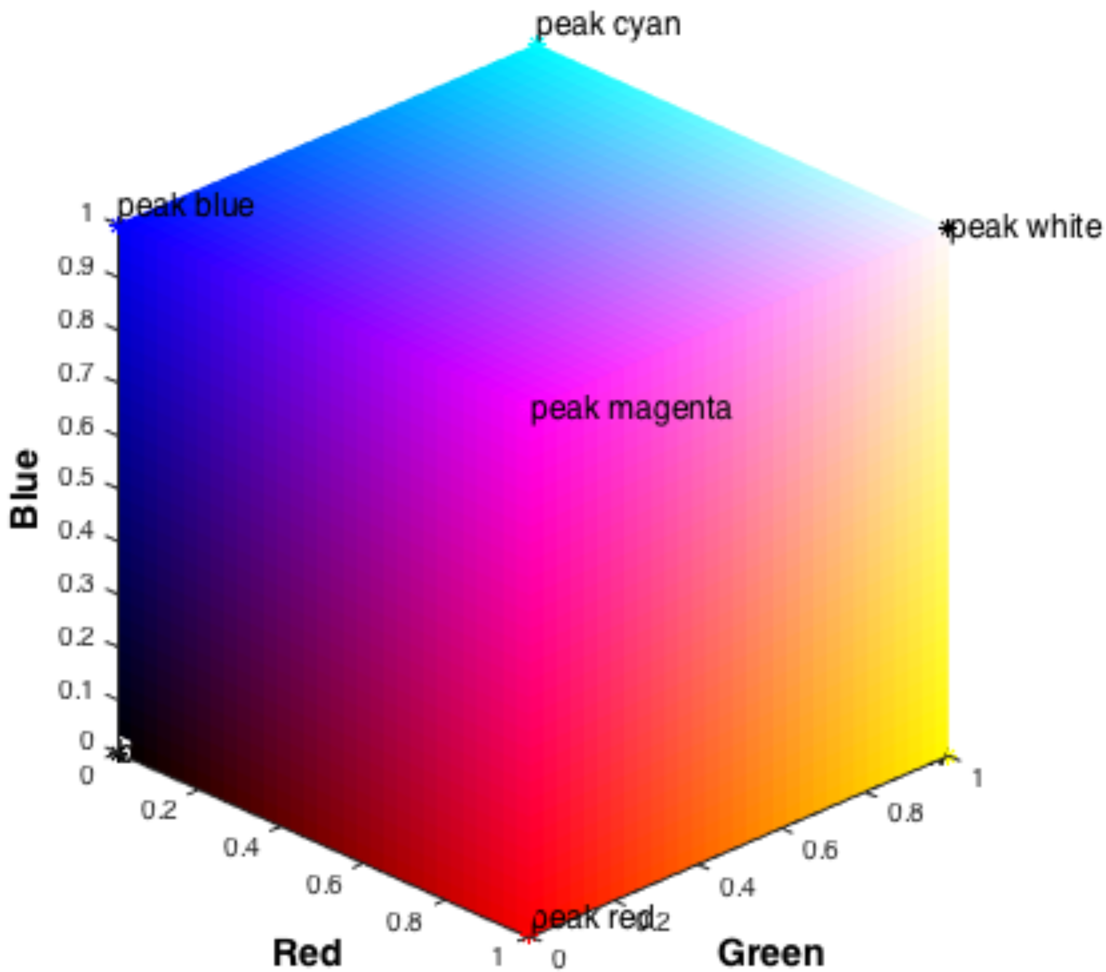
Preserving P3 red primary in BT.2020/2100

allow -2 % (linear) Blue to pass through signal margin?

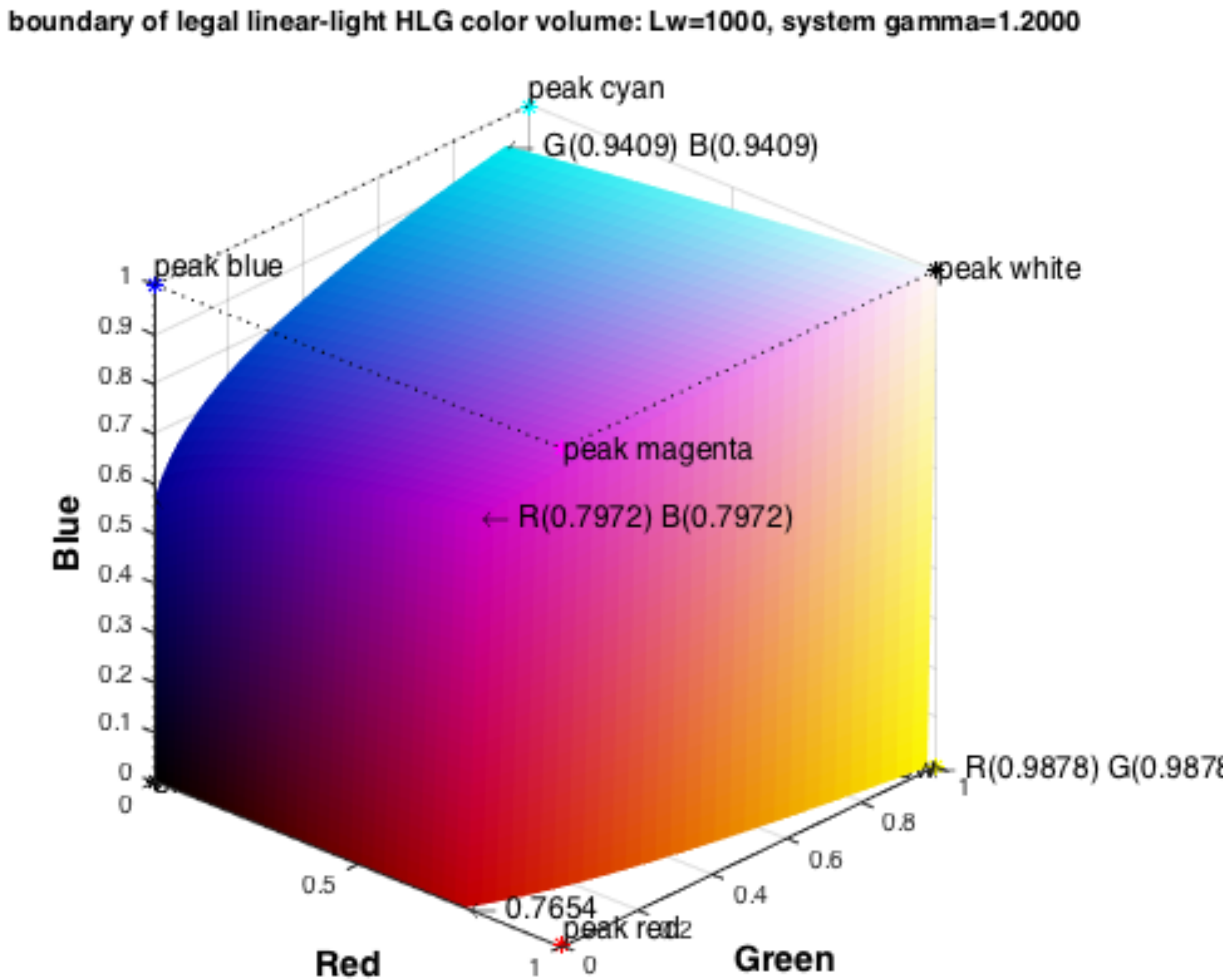
Problem: PQ and HLG transfer functions do not have a definition for negative numbers.



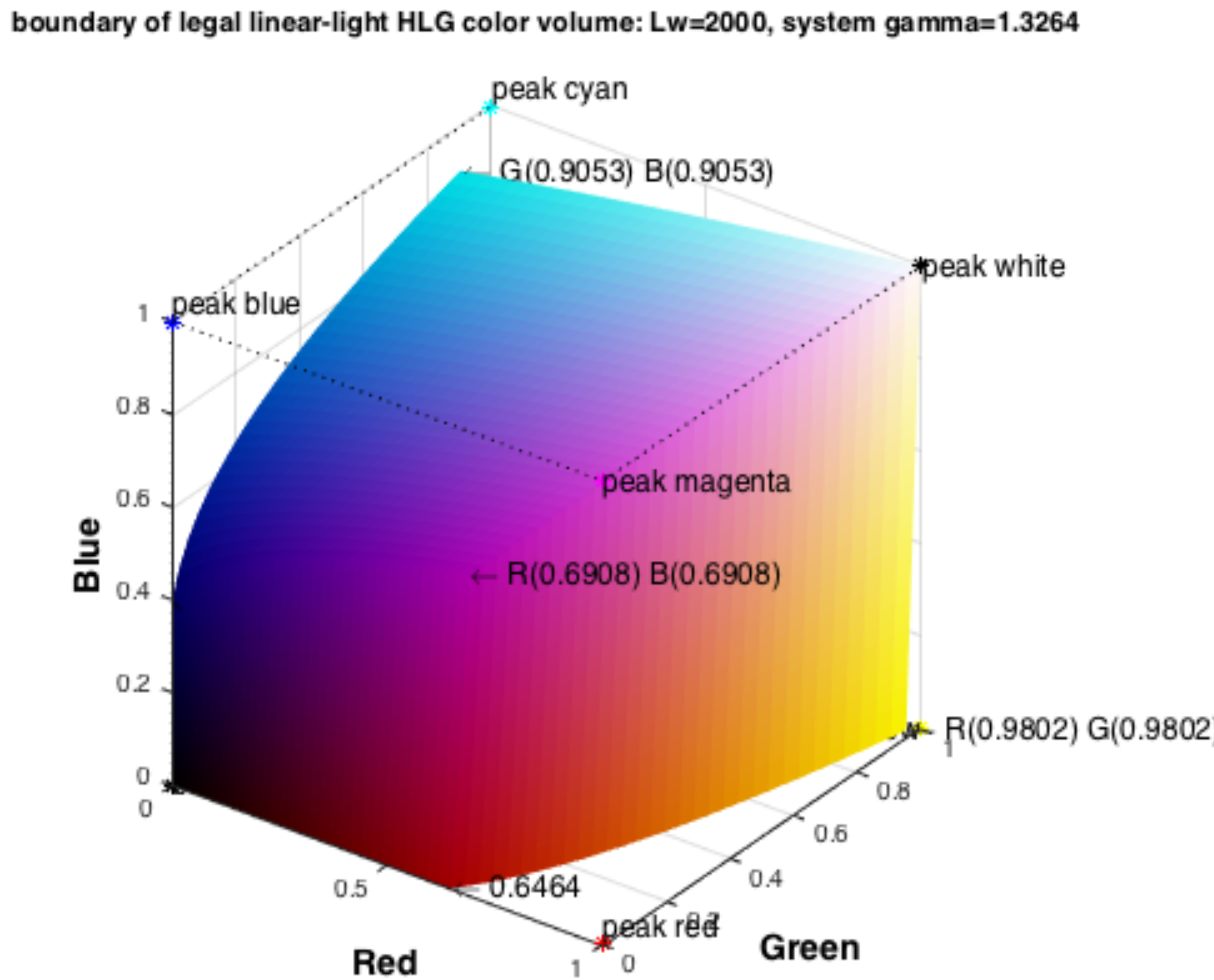
The Full RGB(-PQ) cube



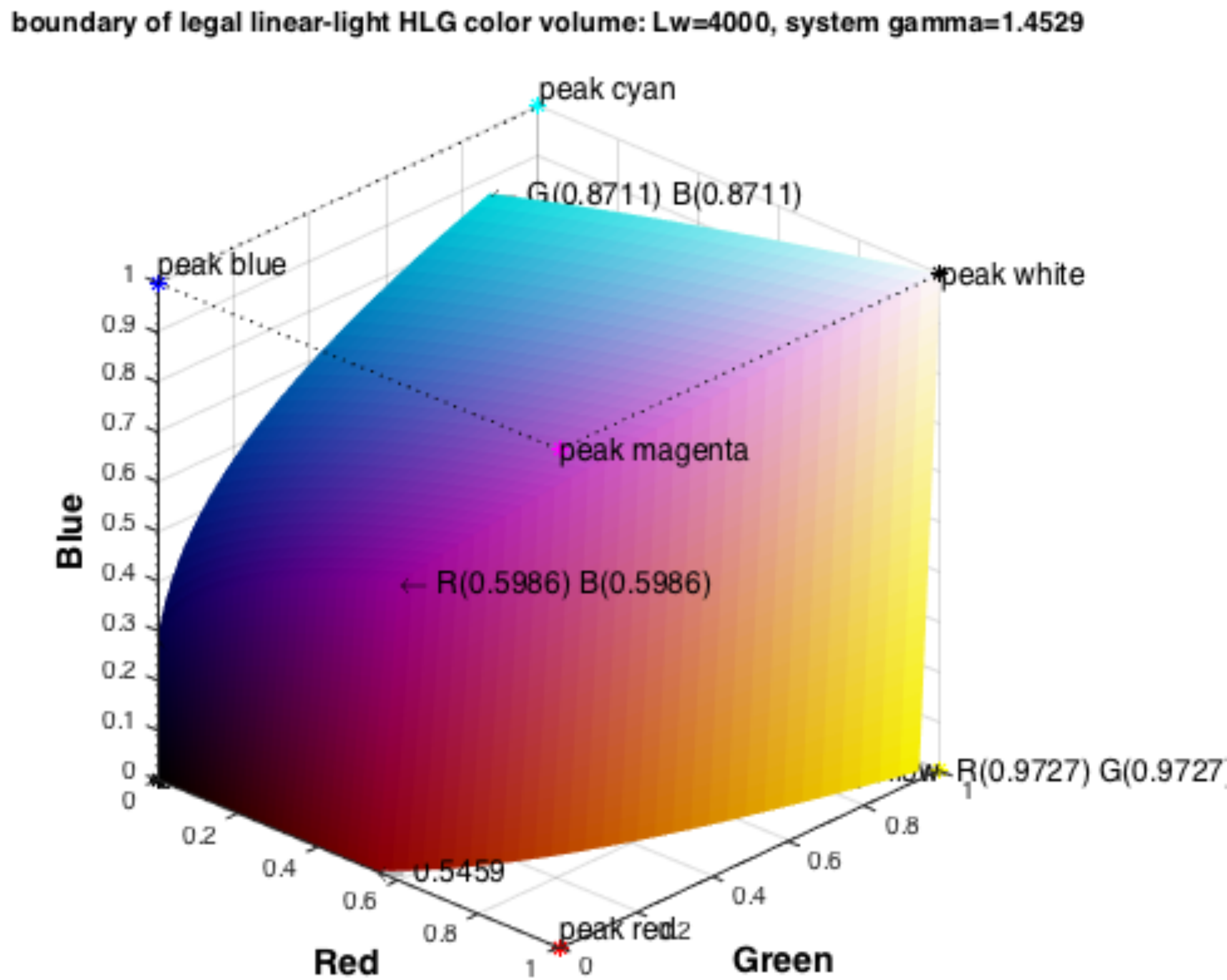
1000 nits tapered cube that avoids E’ [0.0, 1.0] clamping



2000 nits tapered cube that avoids E’ [0.0, 1.0] clamping



4000 nits tapered cube that avoids E’ [0.0, 1.0] clamping



Usual Y’CbCr color volume