400GbE AMs revised proposal

Pete Anslow, Ciena

IEEE P802.3bs Task Force, Logic Ad Hoc, 8 January 2016

Introduction

A set of alignment markers for 400GbE was proposed in anslow_01_1215_logic. The worst case performance of the proposed codes was analysed for 4:1 bit interleaving for 100 Gb/s lanes and found to have a large “shoulder” on the clock content for “all transitions” (page 22).

In discussion on the Logic Ad Hoc call and also in association with wangz_01_1215_logic, it was proposed to reduce the length of the common part of the AMs to 48 bits rather than 64 bits.

Reducing the unique part of the AMs to also be 48 bits long gives AMs 96 bits long in total, which will fit in 6 x 257-bit blocks with only 6 bits set to the free-running PRBS9.

To further improve the transition density for 4:1 bit interleaving for 100 Gb/s lanes it is also proposed to make the 48-bit common part based on AM6 for 100GbE as it has more transitions than AM0.

Baseline wander

Previous NRZ contributions have used a “baseline wander” parameter

This was defined as:

Baseline wander is the instantaneous offset (in %) in the signal generated by AC coupling at the Baud rate / 10,000.

This analysis re-uses this definition unmodified, but it should be noted that for PAM4, the eye height is 1/3 that of NRZ so the effects of a given amount of baseline wander will be greater.

For NRZ contributions see:

P802.3ba anslow_01_0108
P802.3ba anslow_06_1108
P802.3bj anslow_01a_0112

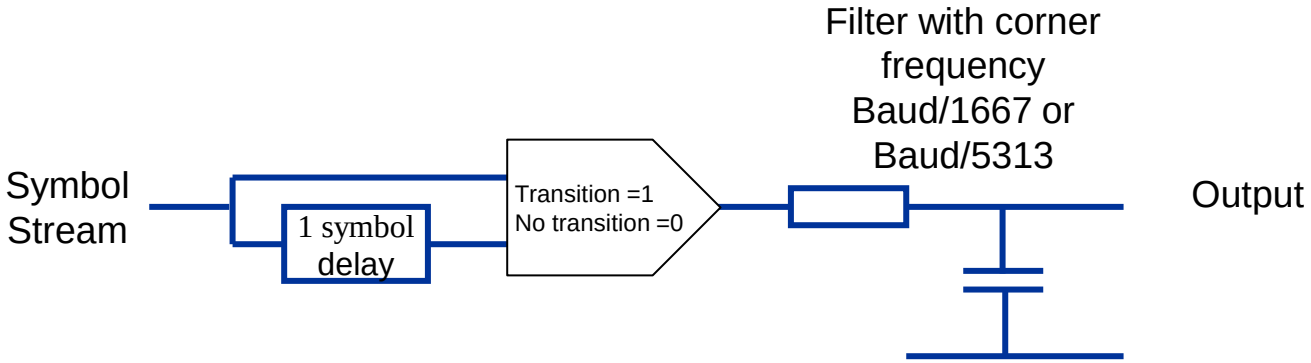
Clock content

Previous NRZ contributions have also used a “clock content” parameter defined as:

Create a function which is a 1 for a transition and a 0 for no transition and then filter the resulting sequence with a corner frequency of Baud/1667. (See additional slides for a corner frequency of Baud/5313)

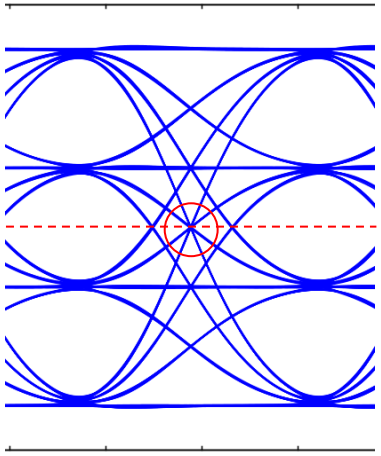
This analysis re-uses this definition unmodified but defines a transition as one of three possibilities (as per healey_3bs_01_1115):

- Symmetrical transitions through the signal average
- Transitions through the signal average
- All transitions

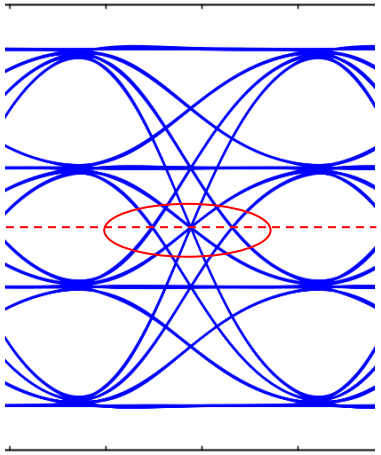


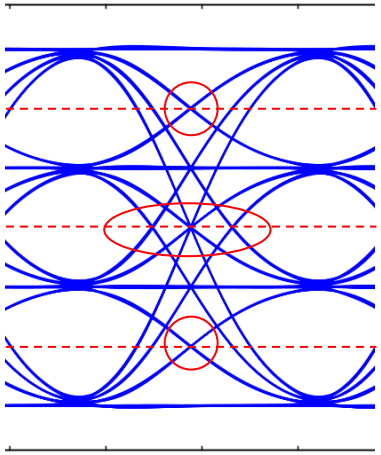
Clock content illustration

Symmetrical transitions through the signal average



Transitions through the signal average





All transitions

Revised alignment markers

This contribution (and comment #146 against D1.1) proposes 96-bit alignment markers with a 48-bit common part of “0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9” followed by a 48-bit unique part.

The 48-bit unique parts of the alignment markers are composed of the first three bytes of the unique parts in anslow_01_1215_logic followed by their inverse.

The proposed markers are shown on the next page.

Revised alignment marker proposal

Table 119–1—400GBASE-R Alignment marker encodings

PCS lane number	Encoding[a] {CM_0, CM_1, CM_2, CM_3, CM_4, CM_5, UM_0, UM_1, UM_2, UM_3, UM_4, UM_5}
0	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x9E, 0xEB, 0x27, 0x61, 0x14, 0xD8
1	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x50, 0x74, 0x88, 0xAF, 0x8B, 0x77
2	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0xB4, 0xB7, 0xEA, 0x4B, 0x48, 0x15
3	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0xE4, 0xFB, 0xF1, 0x1B, 0x04, 0x0E
4	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0xDC, 0x58, 0xEE, 0x23, 0xA7, 0x11
5	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0xBD, 0xA9, 0xBF, 0x42, 0x56, 0x40
6	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x97, 0x67, 0x77, 0x68, 0x98, 0x88
7	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x24, 0x35, 0xA5, 0xDB, 0xCA, 0x5A
8	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x57, 0x64, 0x51, 0xA8, 0x9B, 0xAE
9	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x28, 0xF9, 0x3E, 0xD7, 0x06, 0xC1
10	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0xCB, 0xD1, 0xAD, 0x34, 0x2E, 0x52
11	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x5E, 0x1E, 0x38, 0xA1, 0xE1, 0xC7
12	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x19, 0x98, 0xF9, 0xE6, 0x67, 0x06
13	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x84, 0xEC, 0x20, 0x7B, 0x13, 0xDF
14	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x13, 0xA4, 0xED, 0xEC, 0x5B, 0x12
15	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x3F, 0x8A, 0xBE, 0xC0, 0x75, 0x41

[a]Each octet is transmitted LSB to MSB.

Simulations

Using these new alignment codes, all possible combinations of PCS lanes for 4:1 bit interleaving for 100 Gb/s lanes were then analysed to find the worst cases for Baseline Wander (BW) and Clock Content (CC) after Gray coding to PAM4 symbols. These searches included lane delays of -20 to +20 as per the previous analysis for the 4:1 case. (The searches for Baud/5313 were -15 to +15.)

The worst case PCS lane combinations and delays were then used to generate the worst case PDFs for 400 GbE scrambled idle 100 Gb/s lanes.

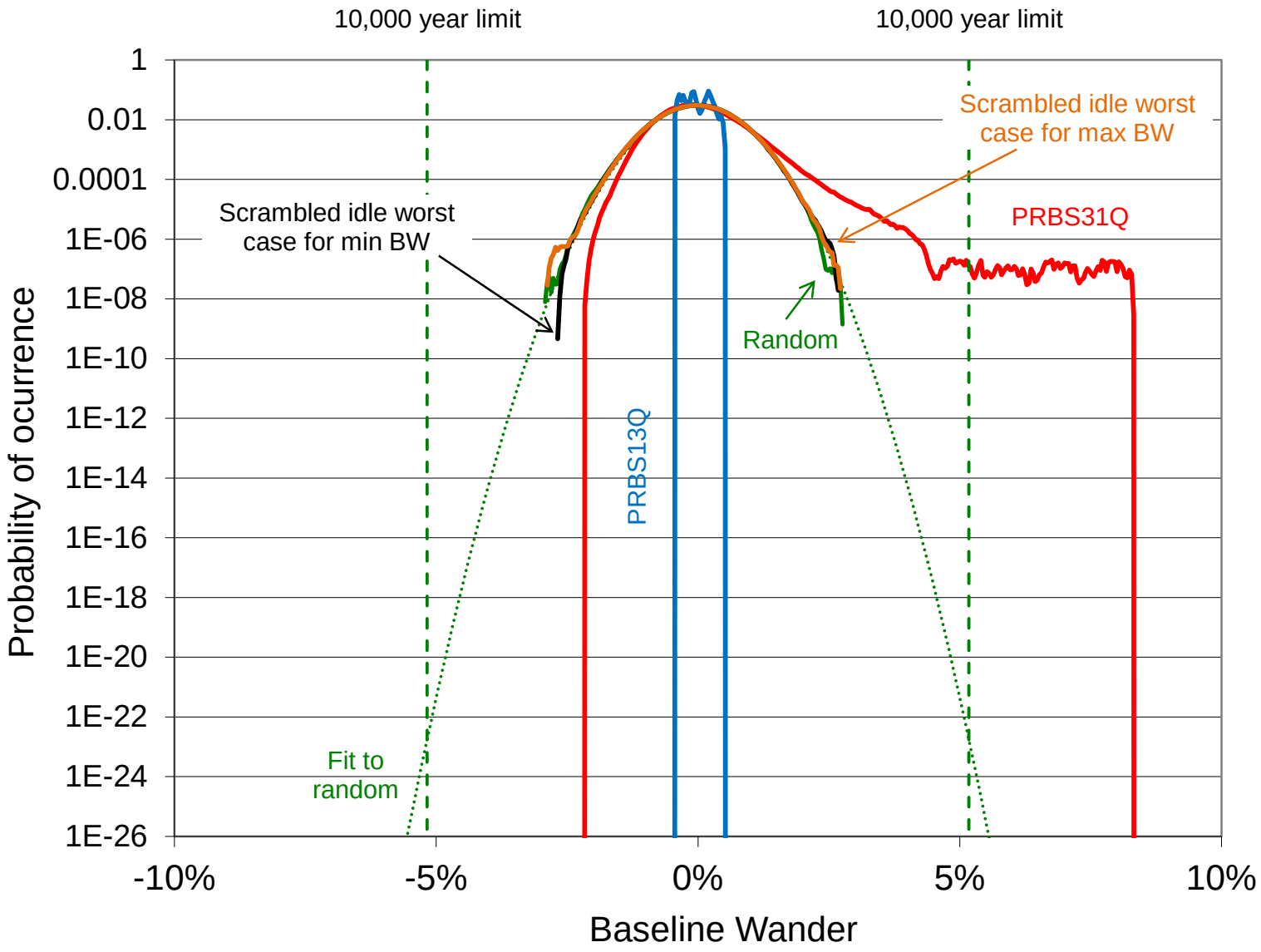
Scrambled idle construction

The scrambled idle symbol streams generated for this analysis were:

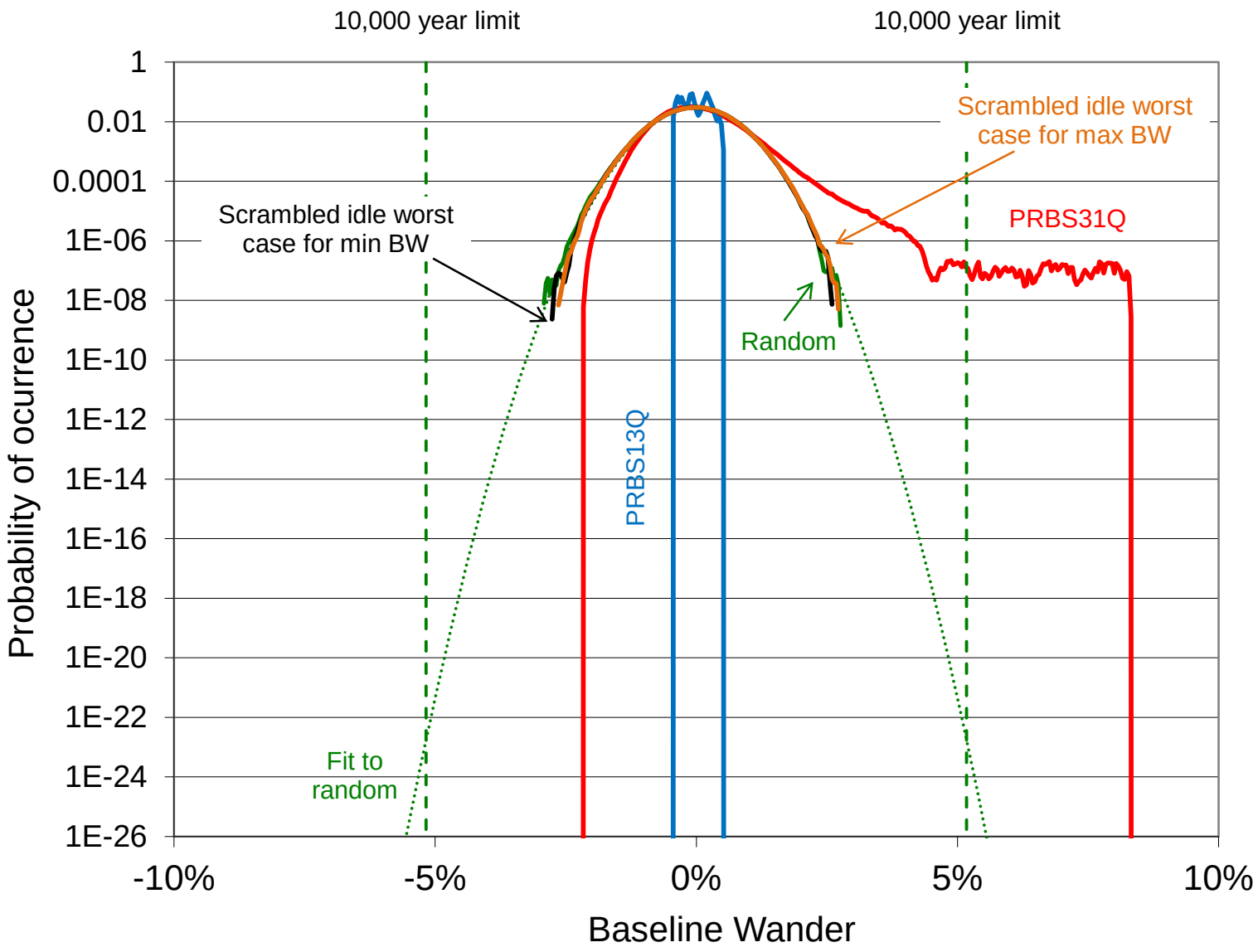
- Idle control characters
- 256B/257B transcoded
- Scrambled
- Distributed 10 bits at a time to two FEC codewords which start with alignment markers and 6 bits of PRBS9 one in every 8192 code words
- 300 bits of RS(544,514) FEC parity added
- Interleaved 10 bits at a time to form PCS lanes (option 8a)
- Bit interleaved with worst case PCS lane combinations and delays

The results for baseline wander and clock content are in the following slides with the results for the previous proposal followed by those for the new proposal to aid comparison. For the clock content slides there are also plots for Baud/5313.

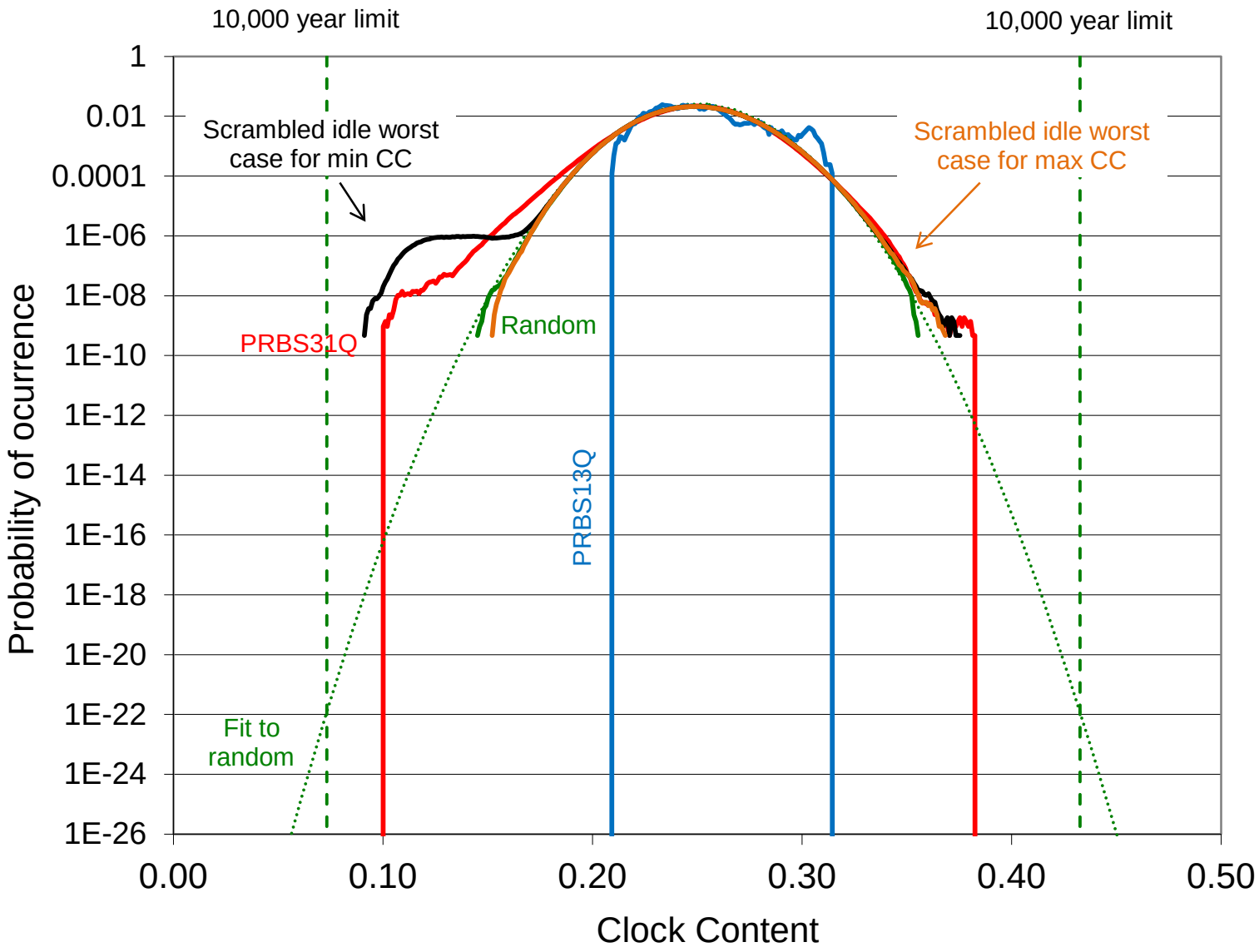
Baseline wander, 100G lanes, previous proposal



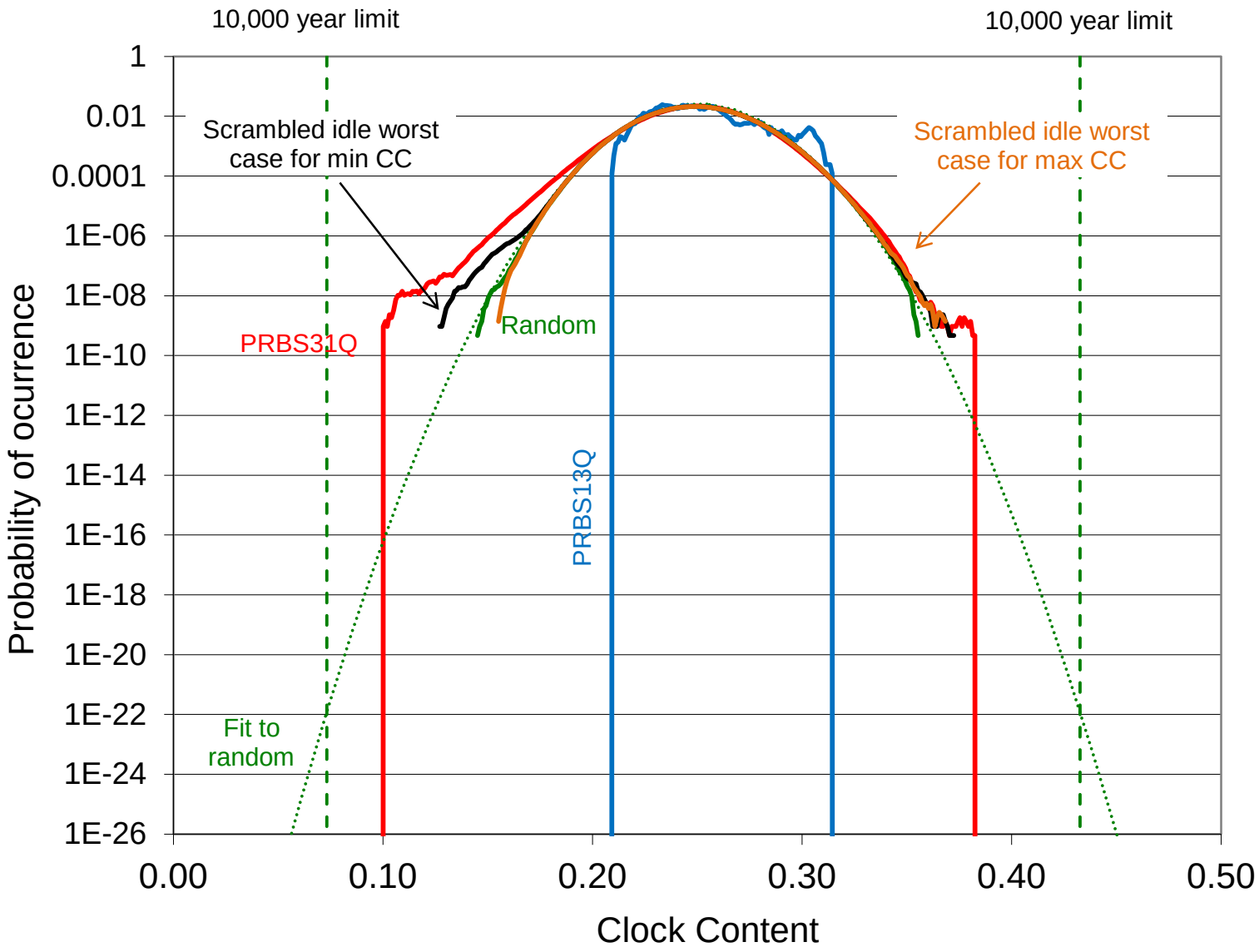
Baseline wander, 100G lanes, new proposal



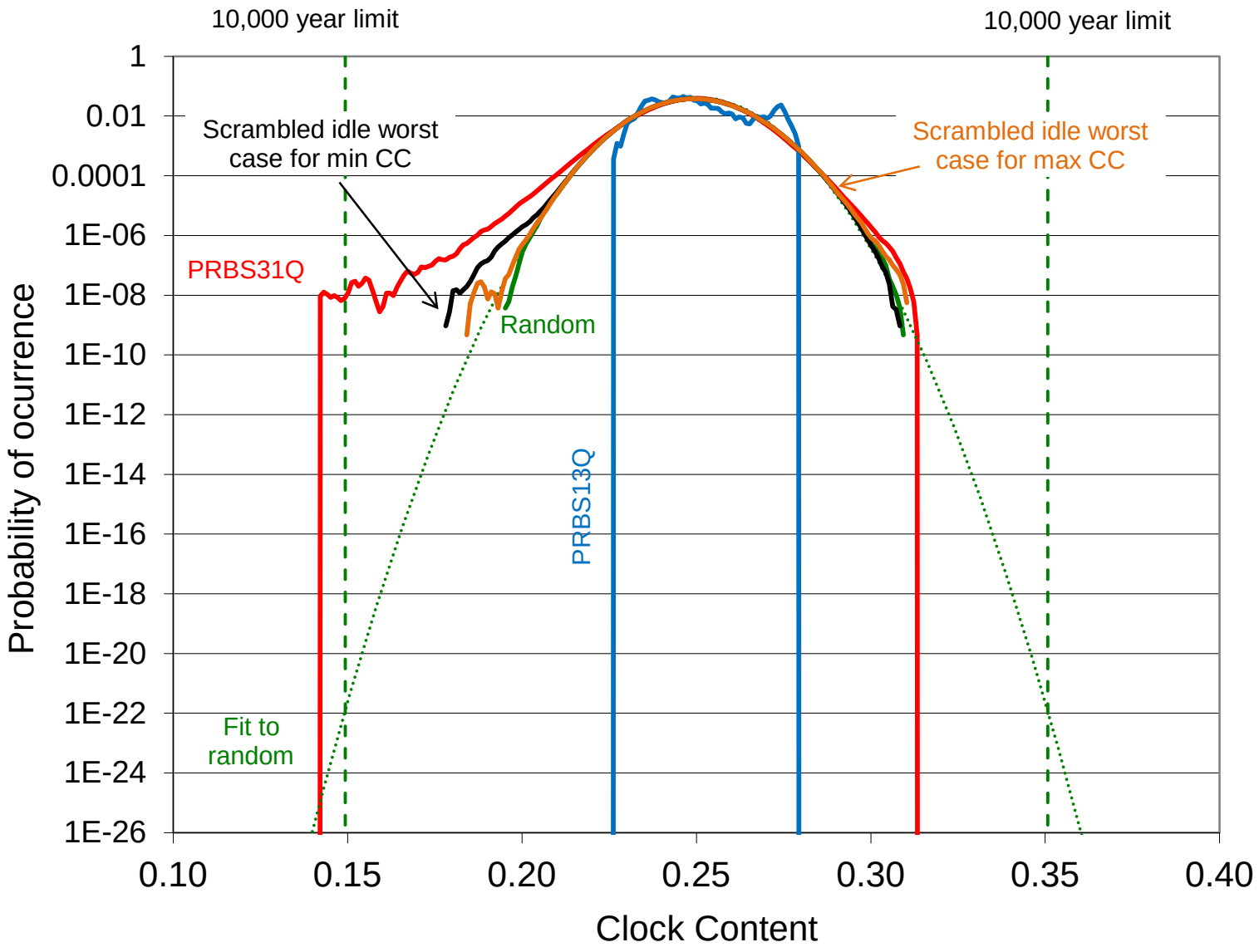
Clock, sym trans thro ave, 100G lanes, previous



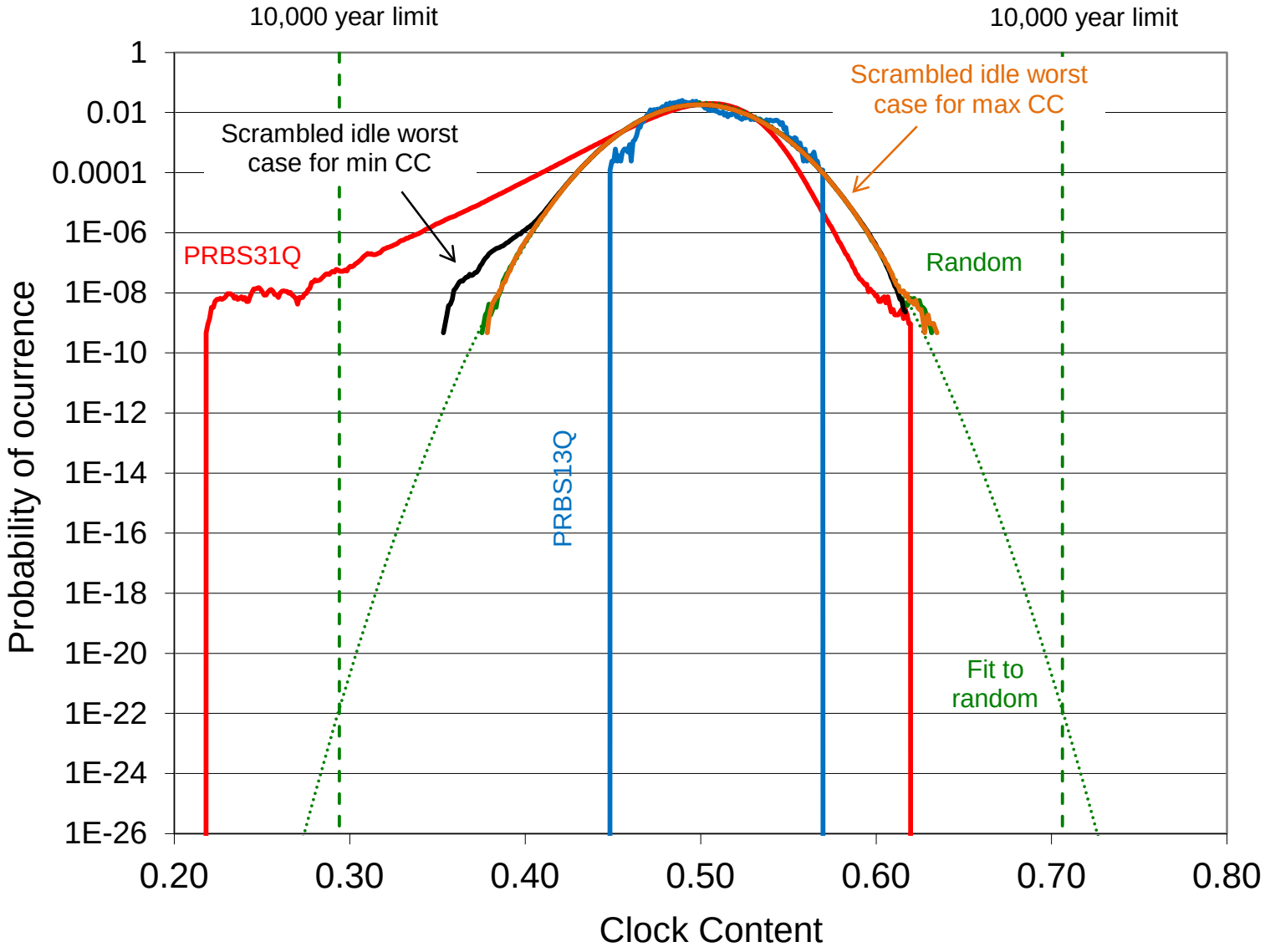
Clock, sym trans thro ave, 100G lanes, new



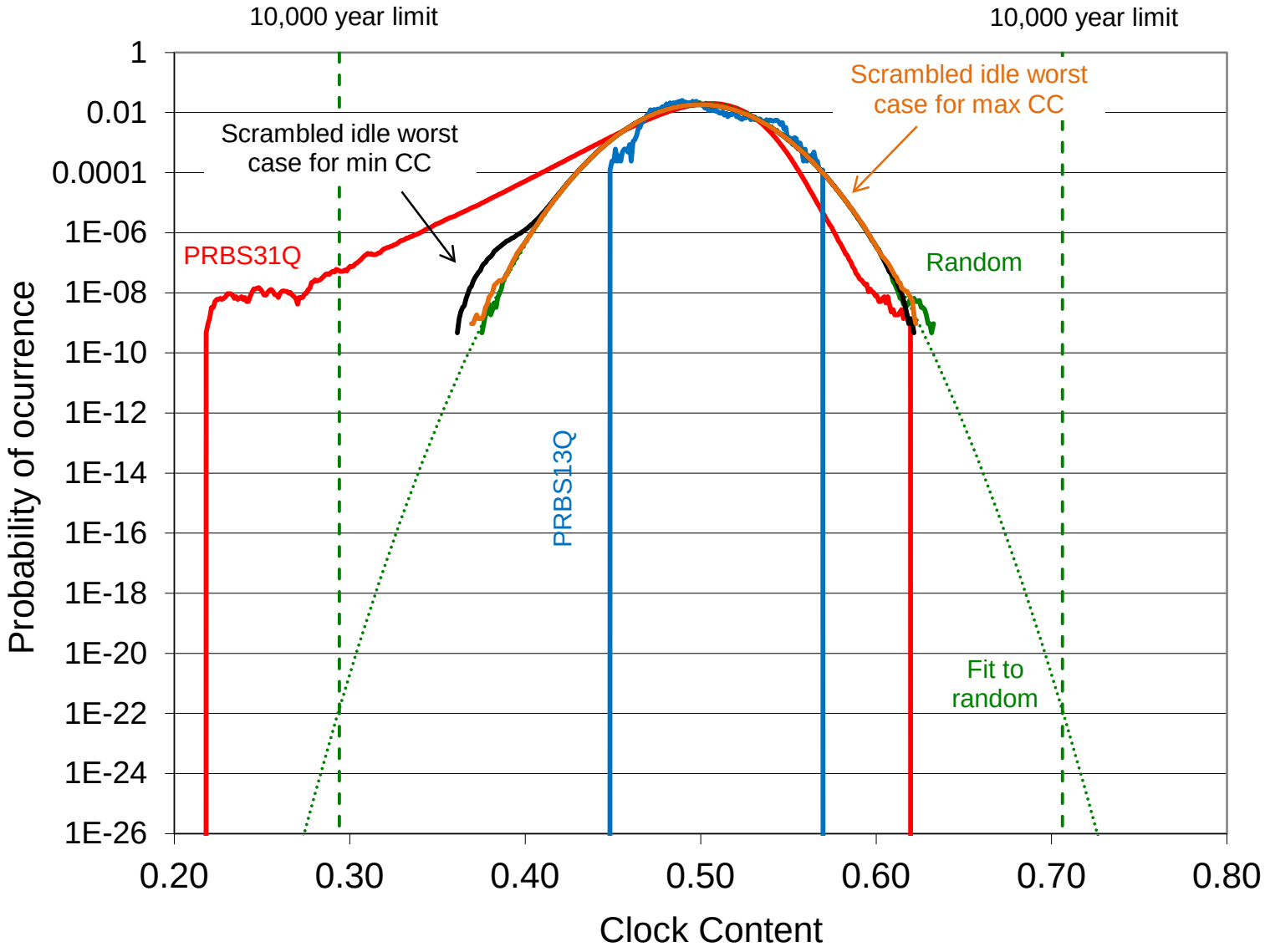
Clock, sym trans thro ave, 100G, new, Bd/5313



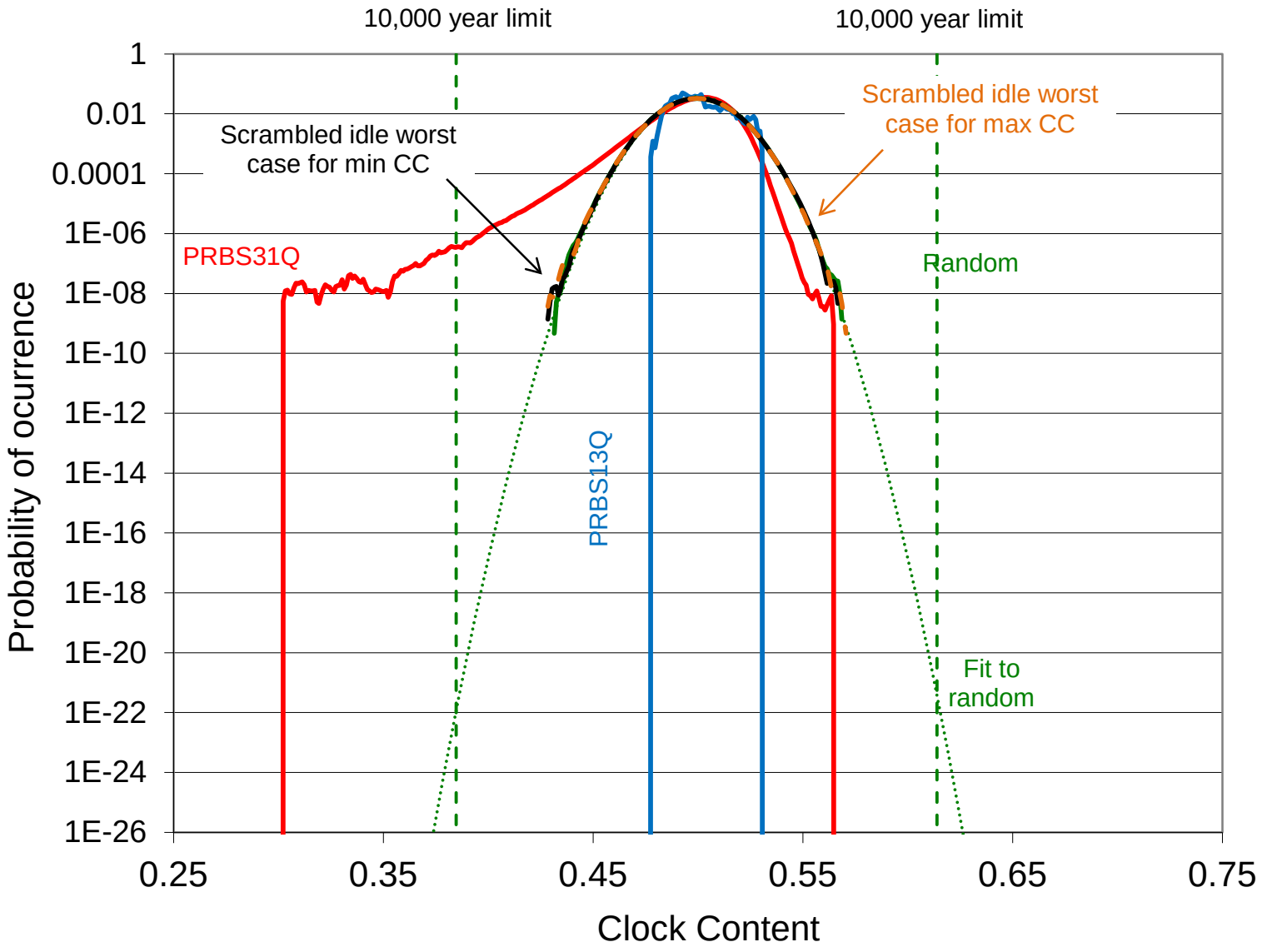
Clock, trans through ave, 100G lanes, previous



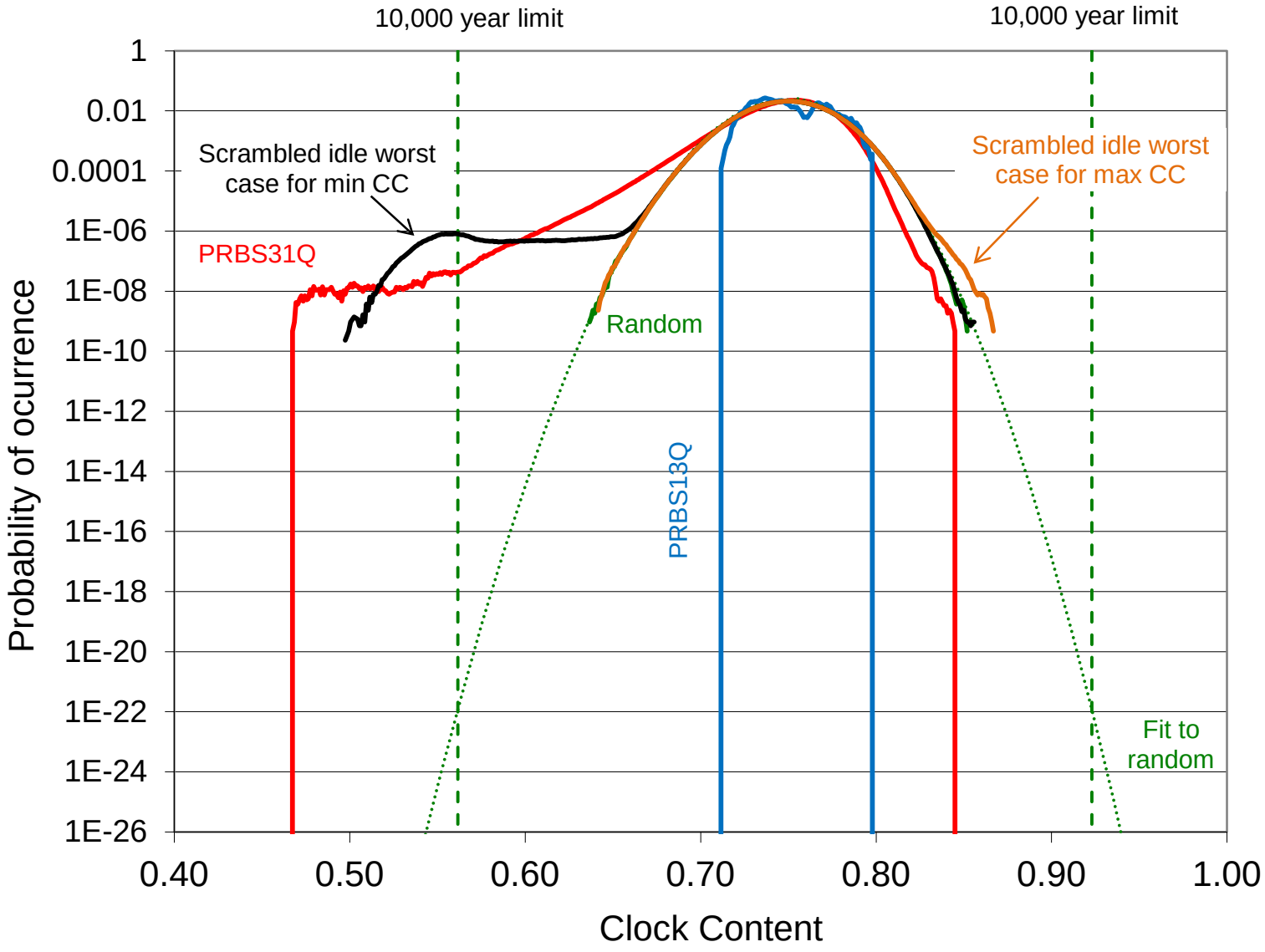
Clock, trans through ave, 100G lanes, new



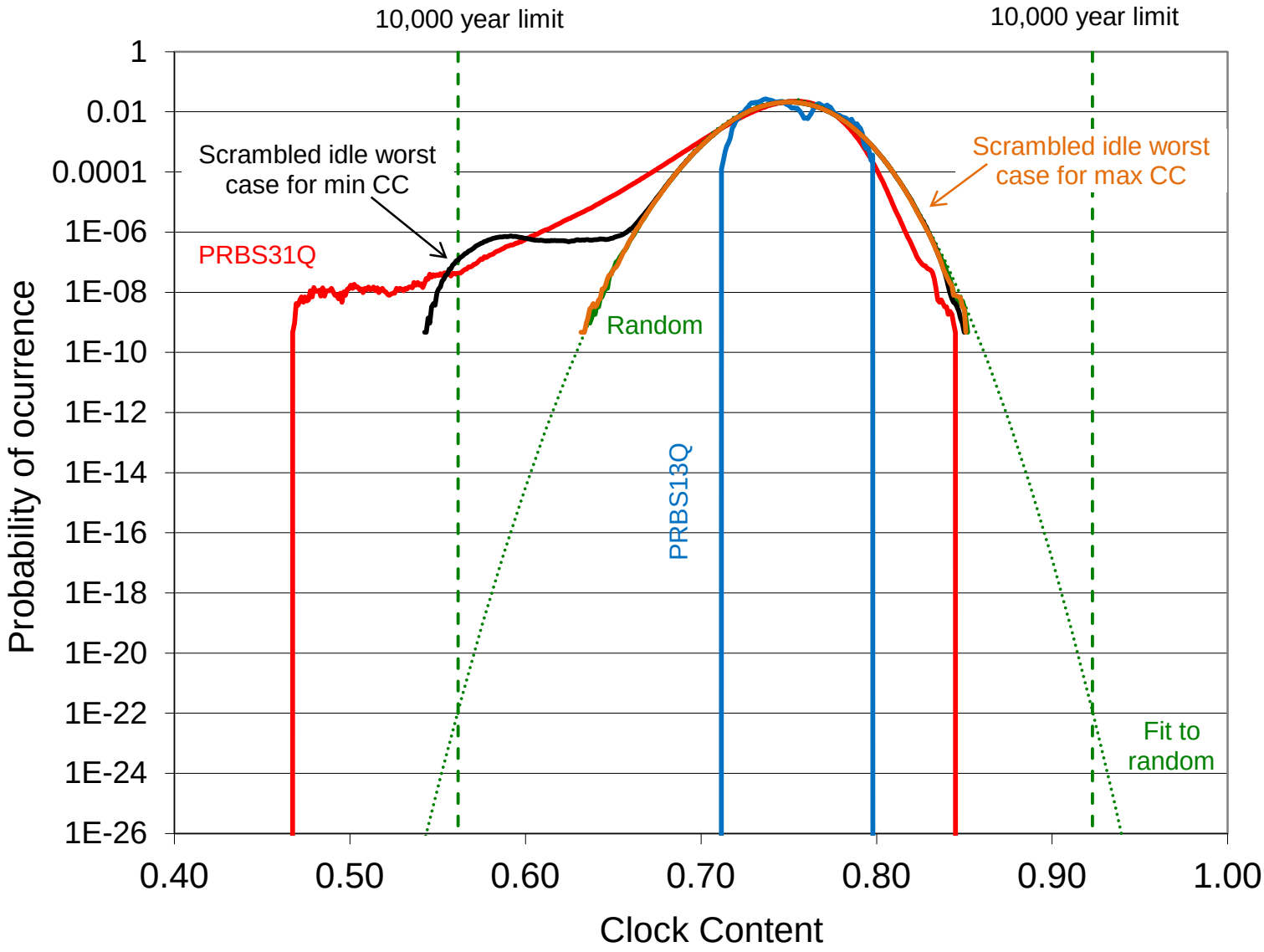
Clock, trans through ave, 100G, new, Bd/5313



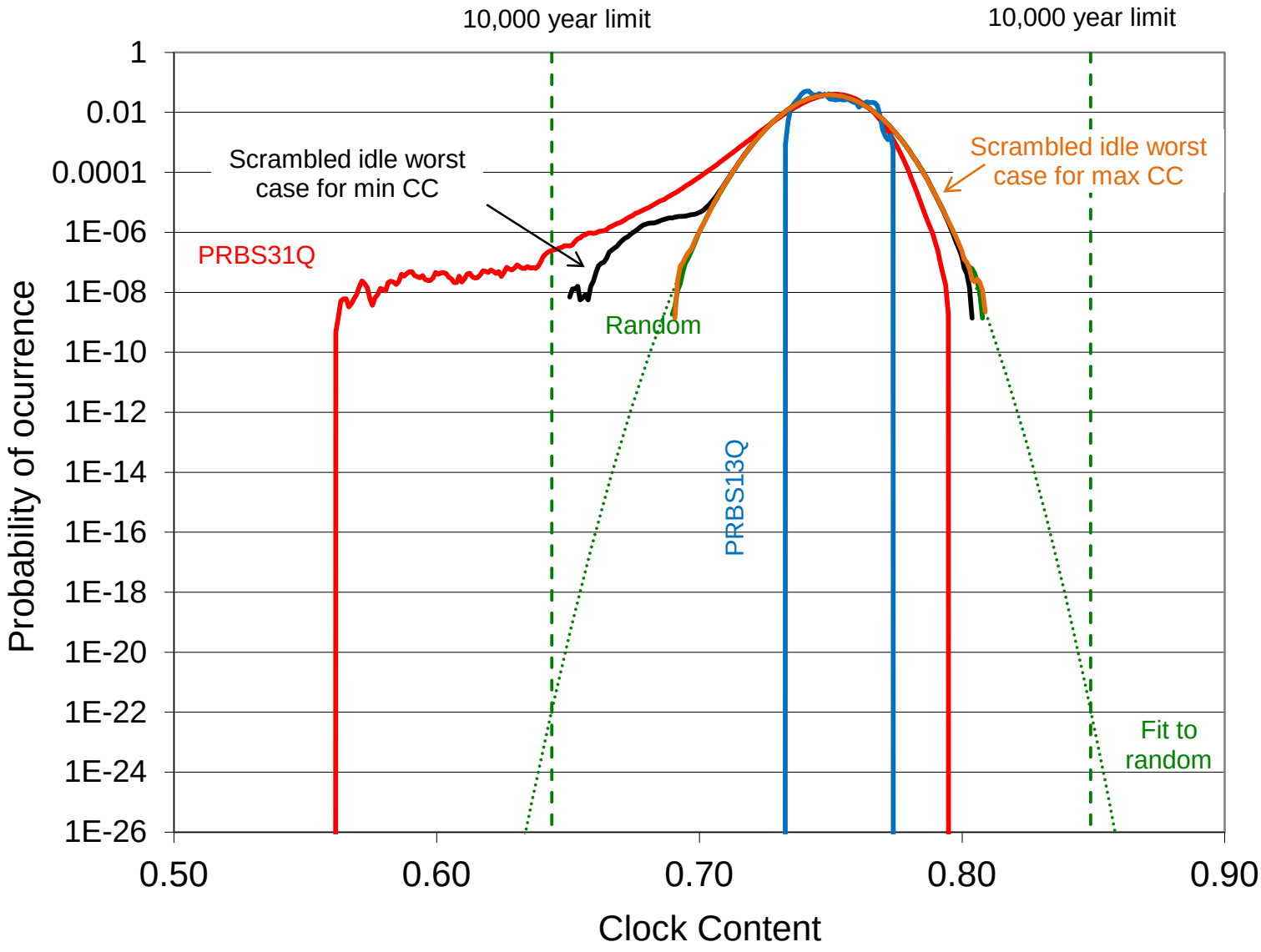
Clock, all transitions, 100G lanes, previous



Clock, all transitions, 100G lanes, new

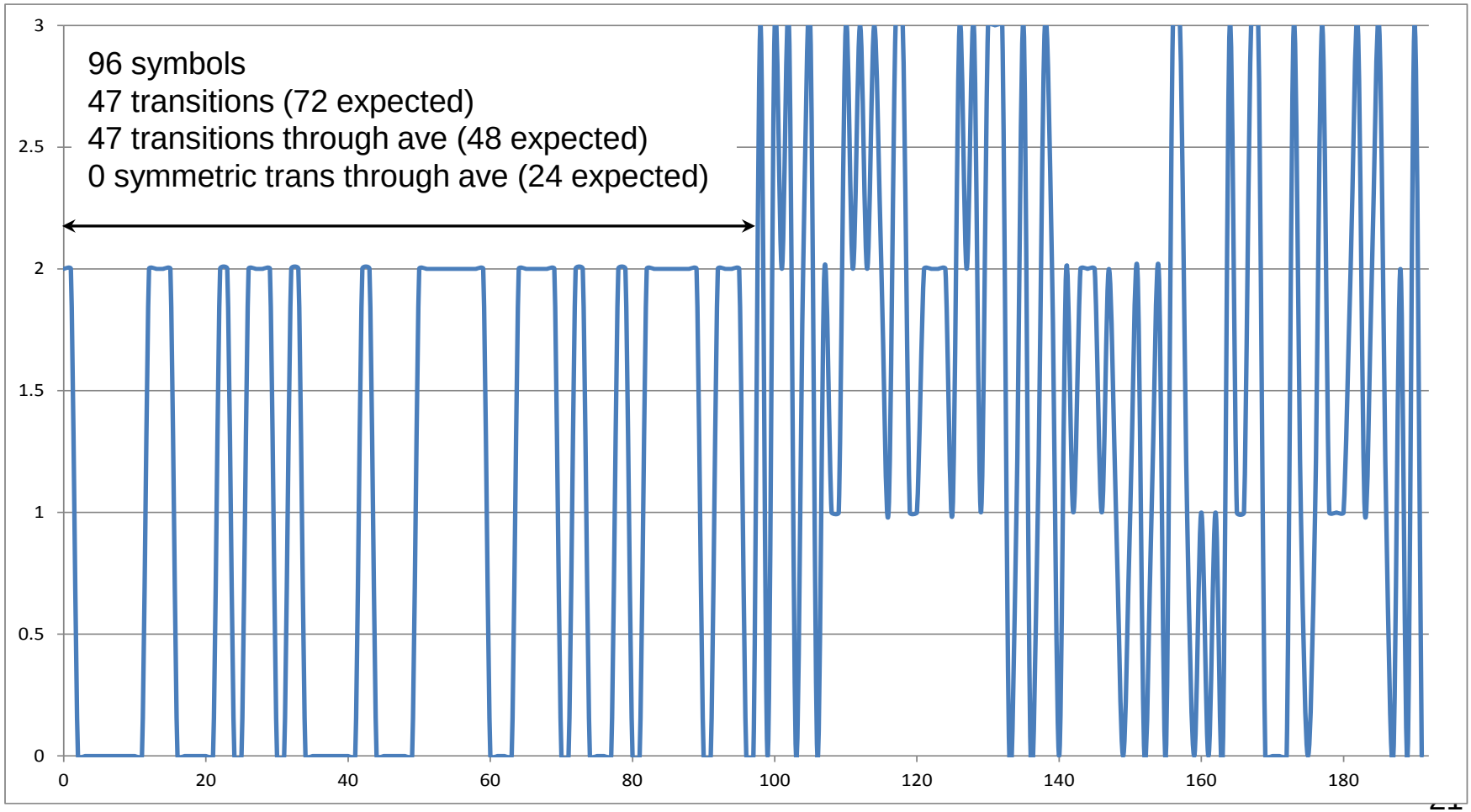


Clock, all transitions, 100G lanes, new, Bd/5313



Effect of common part of lane marker, 4:1

Worst case clock content for all transitions, 100G lanes. Common part of alignment marker gives 96 symbols as below for zero skew:



Conclusion

The baseline wander and clock content for the revised alignment marker proposal still show a “shoulder”, but they are within the baseline wander and clock content for the PRBS31Q test pattern.

It is therefore proposed to use this alignment marker scheme for 400 Gb/s Ethernet in accordance with comment #146 against P802.3bs D1.1

Backup

Worst case lane combinations revised proposal

4:1 bit interleaving for 100 Gb/s lanes

	First lane	Second lane	Third lane	Fourth lane	First lane delay	Second lane delay	Third lane delay	Fourth lane delay
wander_max	0	6	5	15	0	0	-5	-5
wander_min	0	1	2	3	0	1	-7	17
clock25_max	0	1	2	3	0	-19	1	-19
clock25_min	2	7	4	14	0	0	15	15
clock50_max	0	1	2	3	0	19	-18	-11
clock50_min	12	0	14	1	0	15	-1	14
clock75_max	0	1	2	3	0	-2	-2	-1
clock75_min	6	14	10	12	0	0	0	0

Thanks!