400GbE AMs revisited

Pete Anslow, Ciena

IEEE P802.3bs Task Force, Logic Ad Hoc, 28 April 2016

Introduction

The set of alignment markers for 400GbE in D1.3 was analysed in anslow_01_0216_logic and found to have adequate performance for 4:1 bit interleaving for 100 Gb/s lanes.

The alignment marker structure as shown in Figure 119-4:

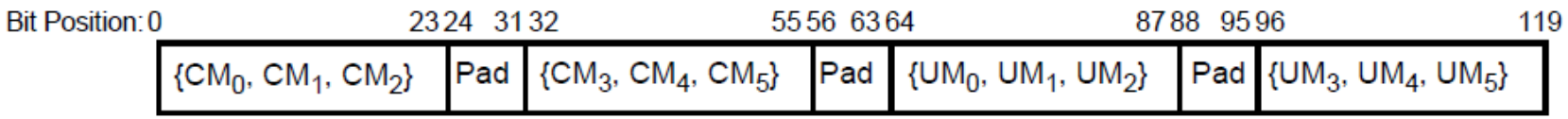


Figure 119–4—Alignment marker format

Where the Pad sections are PRBS9 makes the description and implementation of the AM blocks rather complex.

This contribution analyses the performance of this a revised format where the Pad sections shown in Figure 119-4 are replaced by extra UM bytes.

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.

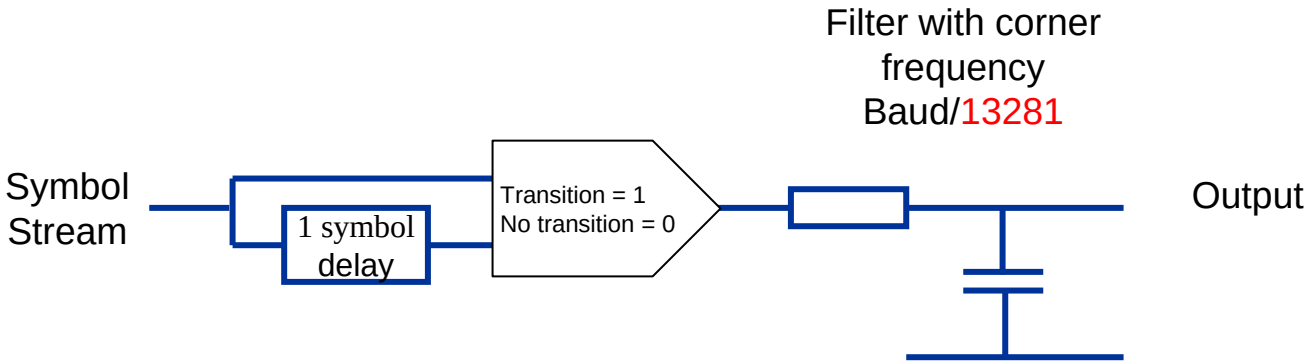
Clock content

The “clock content” parameter is defined here 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/13281.

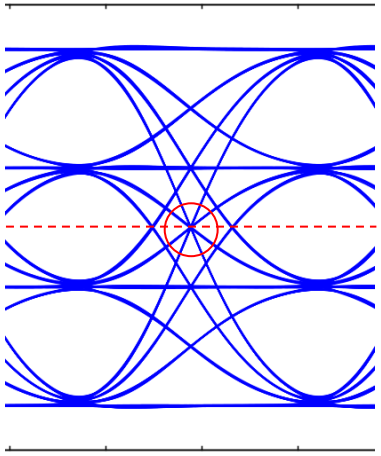
This analysis defines a transition as one of three possibilities (as per healey_3bs_01_1115):

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

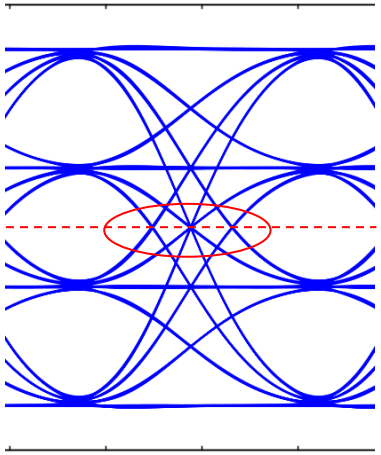


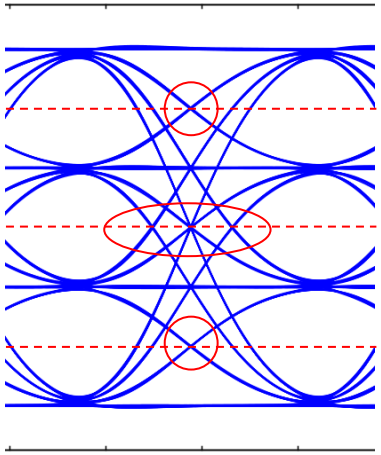
Clock content illustration

Symmetrical transitions through the signal average



Transitions through the signal average





All transitions

Revised alignment markers

This contribution proposes 120-bit alignment markers with a 48-bit common part of “0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9” (same as D1.3) and a 72-bit unique part:

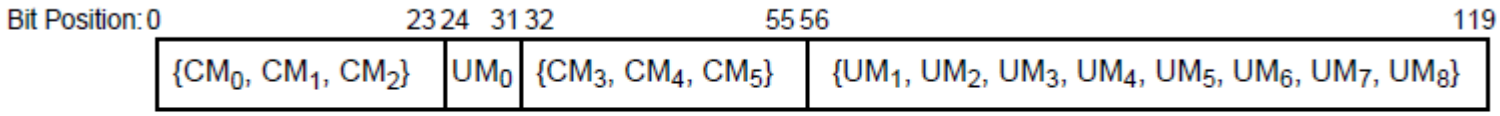


Figure 119–4—Alignment marker format

Where UM_5 to UM_8 are the inverse of UM_1 to UM_4.

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, UM_0, CM_3, CM_4, CM_5, UM_1, UM_2, UM_3, UM_4, UM_5, UM_6, UM_7, UM_8}
0	0x9A, 0x4A, 0x26, 0xB6, 0x65, 0xB5, 0xD9, 0xD9, 0x01, 0x71, 0xF3, 0x26, 0xFE, 0x8E, 0x0C
1	0x9A, 0x4A, 0x26, 0x04, 0x65, 0xB5, 0xD9, 0x67, 0x5A, 0xDE, 0x7E, 0x98, 0xA5, 0x21, 0x81
2	0x9A, 0x4A, 0x26, 0x46, 0x65, 0xB5, 0xD9, 0xFE, 0x3E, 0xF3, 0x56, 0x01, 0xC1, 0x0C, 0xA9
3	0x9A, 0x4A, 0x26, 0x5A, 0x65, 0xB5, 0xD9, 0x84, 0x86, 0x80, 0xD0, 0x7B, 0x79, 0x7F, 0x2F
4	0x9A, 0x4A, 0x26, 0xE1, 0x65, 0xB5, 0xD9, 0x19, 0x2A, 0x51, 0xF2, 0xE6, 0xD5, 0xAE, 0x0D
5	0x9A, 0x4A, 0x26, 0xF2, 0x65, 0xB5, 0xD9, 0x4E, 0x12, 0x4F, 0xD1, 0xB1, 0xED, 0xB0, 0x2E
6	0x9A, 0x4A, 0x26, 0x3D, 0x65, 0xB5, 0xD9, 0xEE, 0x42, 0x9C, 0xA1, 0x11, 0xBD, 0x63, 0x5E
7	0x9A, 0x4A, 0x26, 0x22, 0x65, 0xB5, 0xD9, 0x32, 0xD6, 0x76, 0x5B, 0xCD, 0x29, 0x89, 0xA4
8	0x9A, 0x4A, 0x26, 0x60, 0x65, 0xB5, 0xD9, 0x9F, 0xE1, 0x73, 0x75, 0x60, 0x1E, 0x8C, 0x8A
9	0x9A, 0x4A, 0x26, 0x6B, 0x65, 0xB5, 0xD9, 0xA2, 0x71, 0xC4, 0x3C, 0x5D, 0x8E, 0x3B, 0xC3
10	0x9A, 0x4A, 0x26, 0xFA, 0x65, 0xB5, 0xD9, 0x04, 0x95, 0xEB, 0xD8, 0xFB, 0x6A, 0x14, 0x27
11	0x9A, 0x4A, 0x26, 0x6C, 0x65, 0xB5, 0xD9, 0x71, 0x22, 0x66, 0x38, 0x8E, 0xDD, 0x99, 0xC7
12	0x9A, 0x4A, 0x26, 0x18, 0x65, 0xB5, 0xD9, 0x5B, 0xA2, 0xF6, 0x95, 0xA4, 0x5D, 0x09, 0x6A
13	0x9A, 0x4A, 0x26, 0x14, 0x65, 0xB5, 0xD9, 0xCC, 0x31, 0x97, 0xC3, 0x33, 0xCE, 0x68, 0x3C
14	0x9A, 0x4A, 0x26, 0xD0, 0x65, 0xB5, 0xD9, 0xB1, 0xCA, 0xFB, 0xA6, 0x4E, 0x35, 0x04, 0x59
15	0x9A, 0x4A, 0x26, 0xB4, 0x65, 0xB5, 0xD9, 0x56, 0xA6, 0xBA, 0x79, 0xA9, 0x59, 0x45, 0x86

[a]Each octet is transmitted LSB to MSB.

Simulations

Using these 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 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.

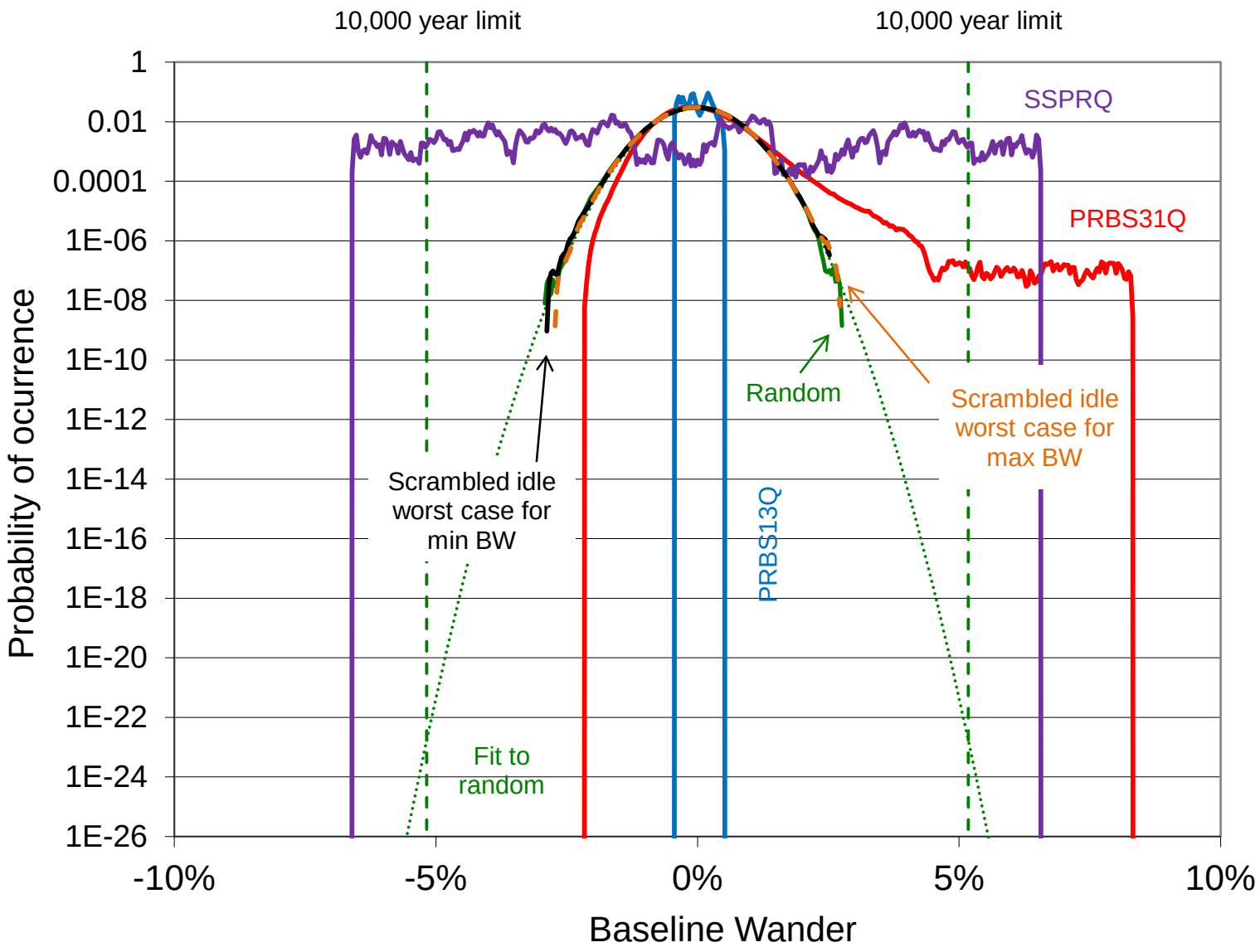
Scrambled idle construction

The scrambled idle symbol streams generated for this analysis were:

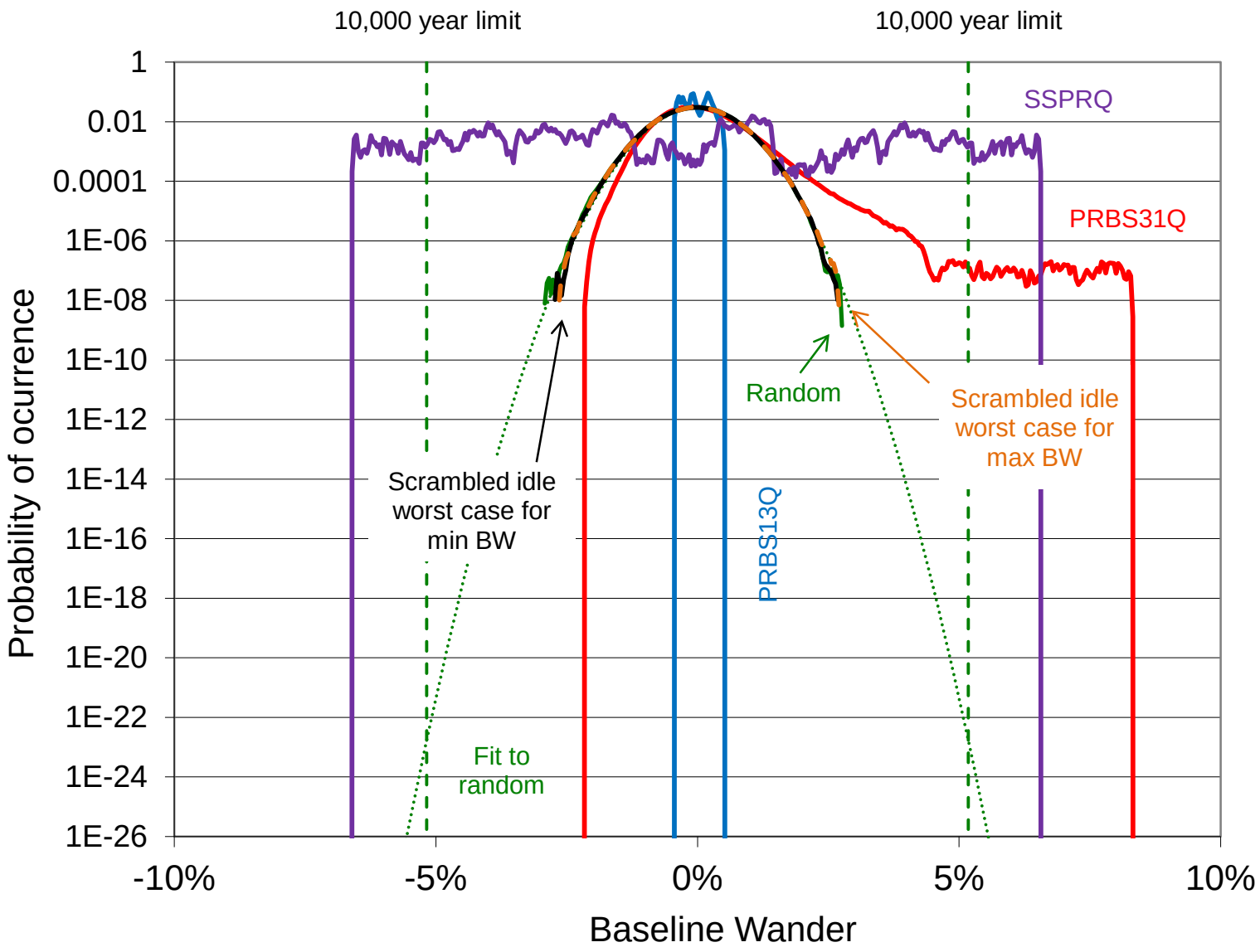
- Idle control characters
- 256B/257B transcoded
- Scrambled
- Distributed 10 bits at a time to two FEC codewords which start with alignment markers followed by 136 bits of PRBS9 one in every 8192 code words
- 300 bits of RS(544,514) FEC parity added
- Interleaved 10 bits at a time to form PCS lanes (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.

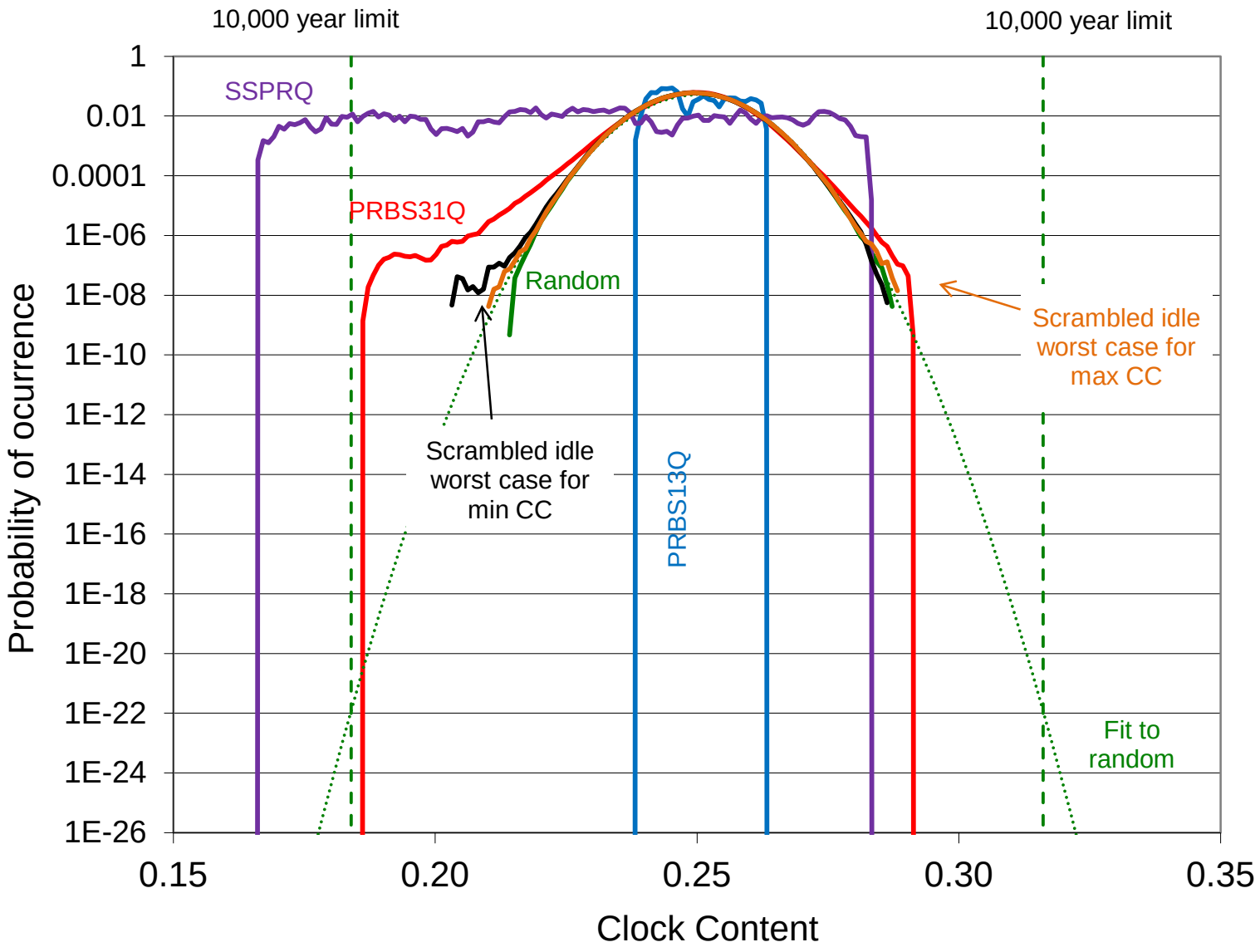
Baseline wander, 100G lanes, D1.3 markers



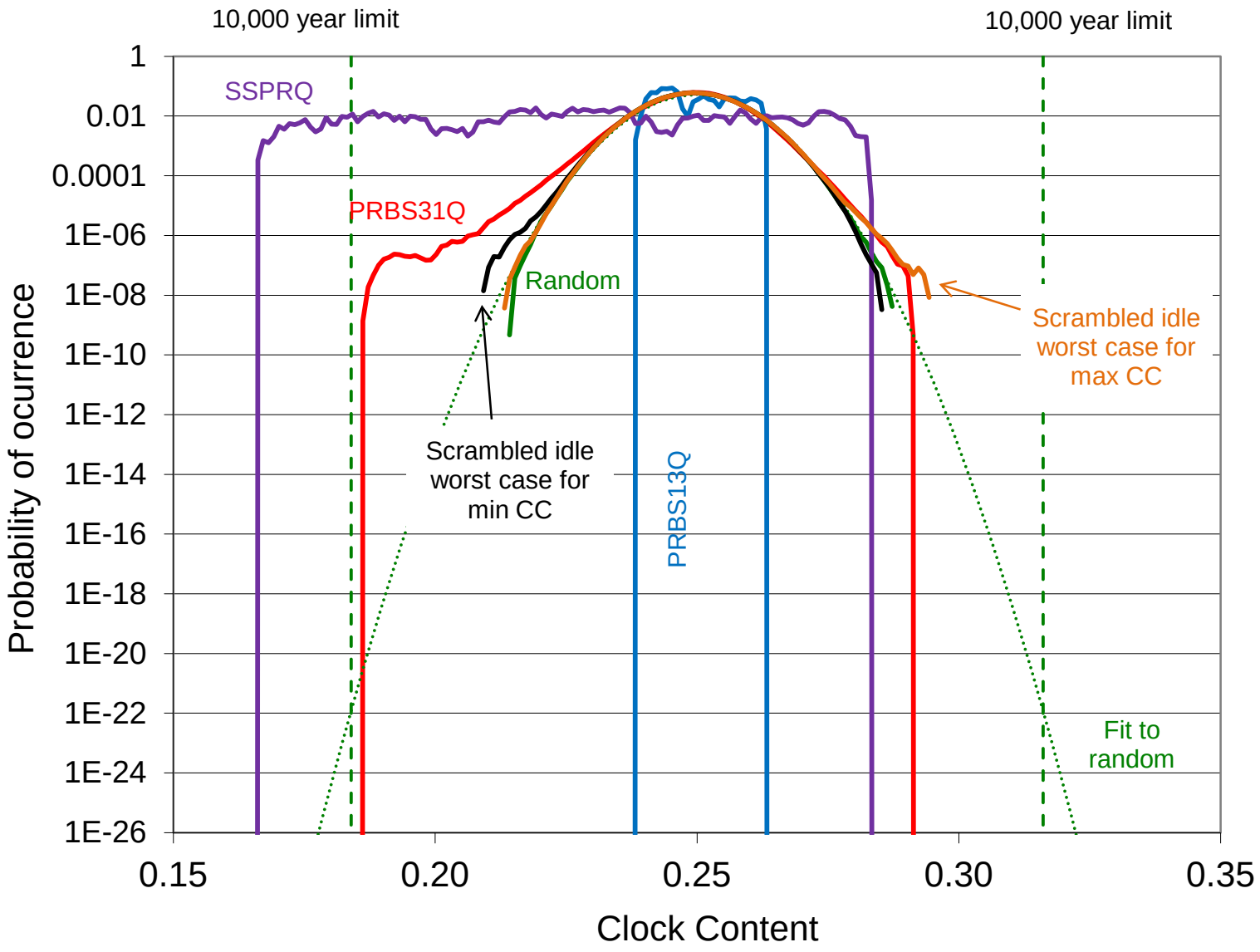
Baseline wander, 100G lanes, new markers



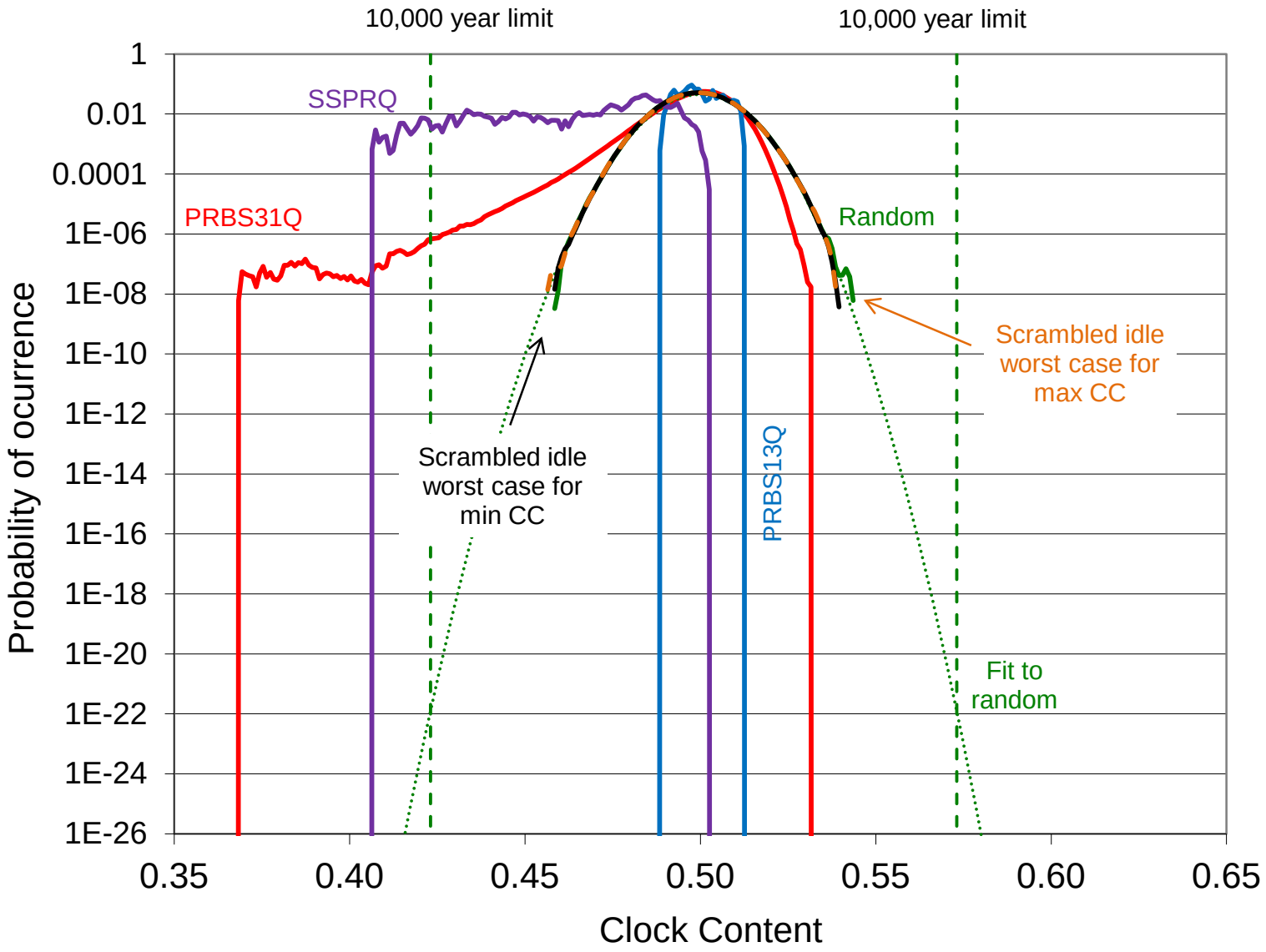
Clock, sym trans through ave, 100G lanes, D1.3



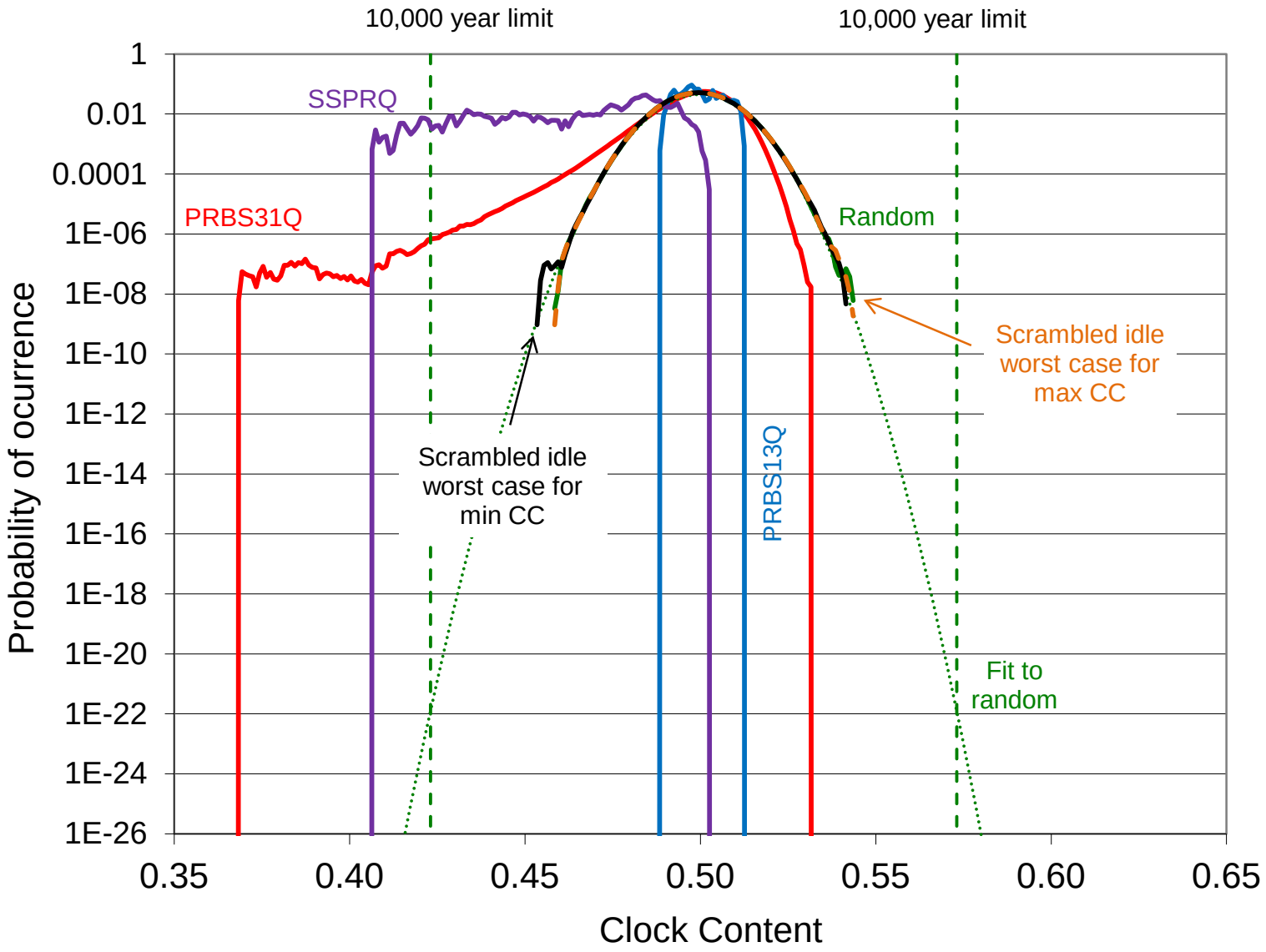
Clock, sym trans through ave, 100G lanes, new



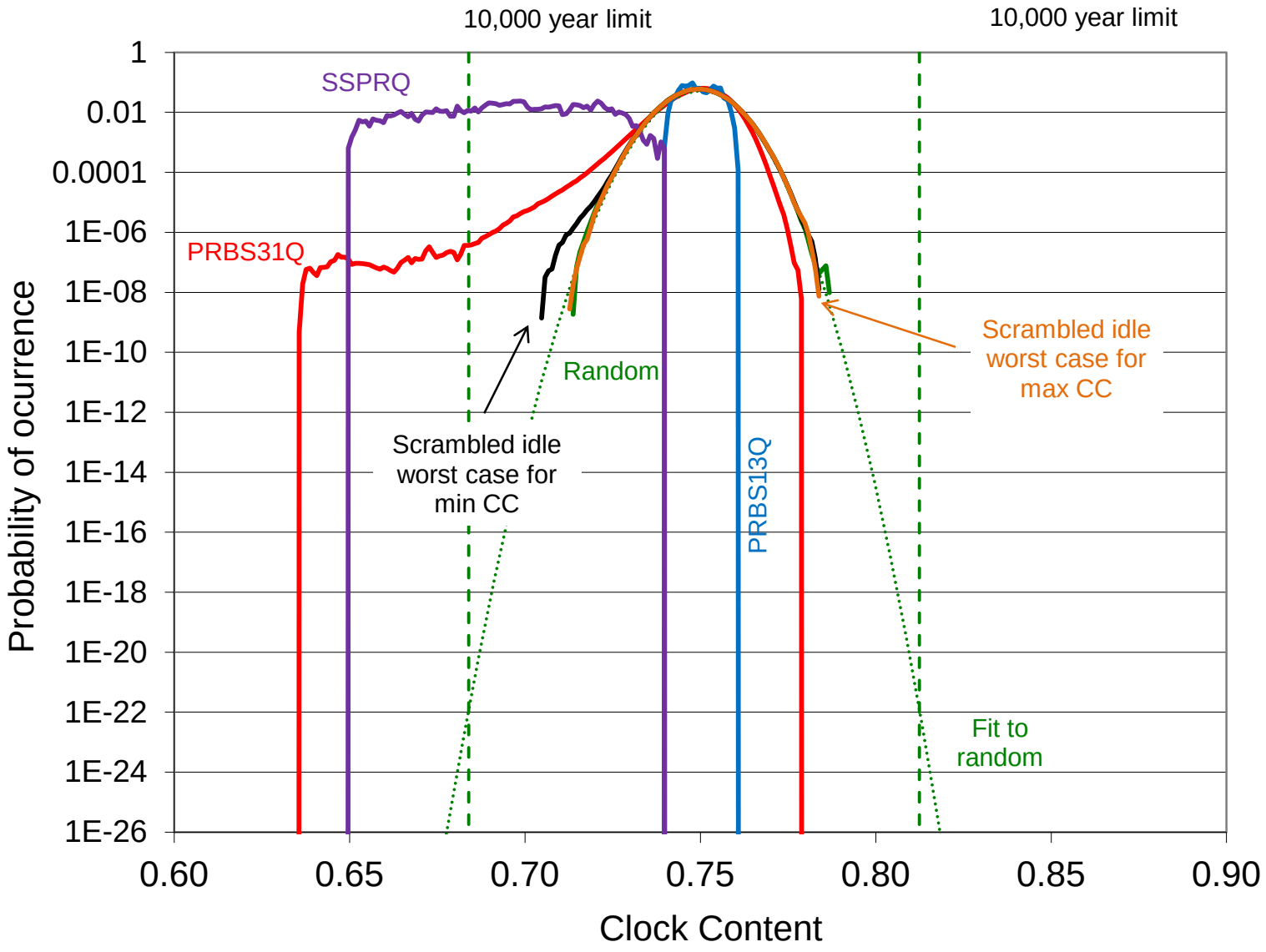
Clock, trans through ave, 100G lanes, D1.3



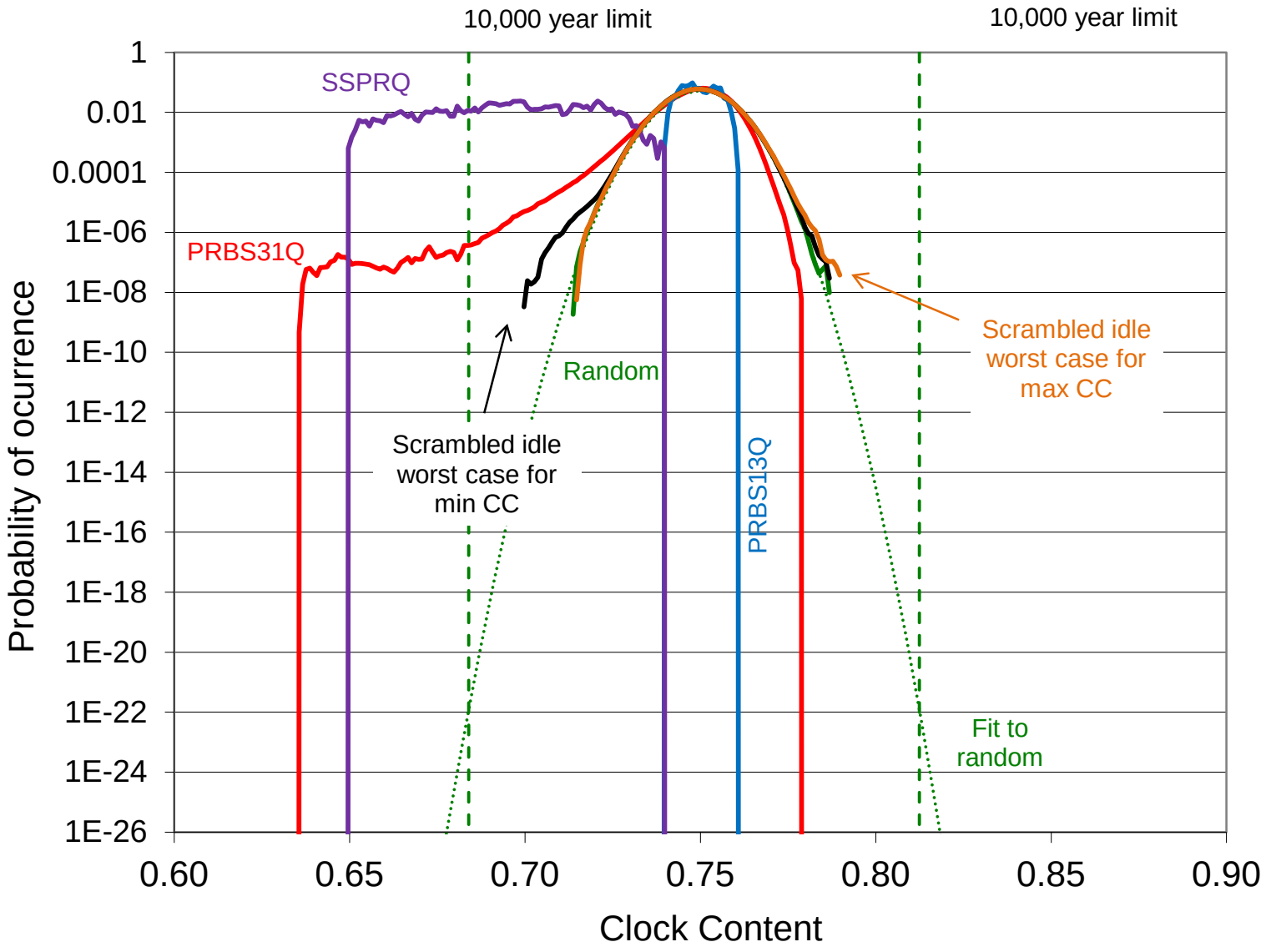
Clock, trans through ave, 100G lanes, new



Clock, all transitions, 100G lanes, D1.3



Clock, all transitions, 100G lanes, new



Conclusion

The baseline wander and clock content for the revised alignment marker proposal show a slightly worse “shoulder” on some clock content plots as expected, but they remain within the baseline wander and clock content for the PRBS31Q test pattern.

It is therefore proposed to use this revised alignment marker scheme for 400 Gb/s Ethernet.

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	8	0	3	10	0	0	20	20
wander_min	2	3	14	11	0	1	-15	17
clock25_max	2	5	11	10	0	0	1	0
clock25_min	9	8	10	3	0	0	-5	-5
clock50_max	6	0	9	1	0	11	-2	-4
clock50_min	5	0	10	1	0	2	0	2
clock75_max	2	9	5	6	0	-2	-2	-1
clock75_min	5	11	13	9	0	0	-1	-1

Thanks!