200GbE alignment markers

Pete Anslow, Ciena

IEEE P802.3bs Task Force, Whistler, Canada, May 2016

Introduction

A replacement set of alignment markers for 400GbE was analysed in anslow_03_0416_logic and found to have adequate performance for 4:1 bit interleaving for 100 Gb/s lanes.

The revised alignment marker structure is proposed to be as shown below (comment #48 against D1.3):

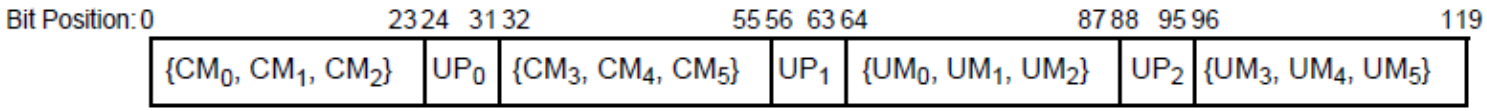


Figure 119–4—Alignment marker format

This contribution analyses the performance of a second set of alignment markers for 200 Gb/s Ethernet.

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.

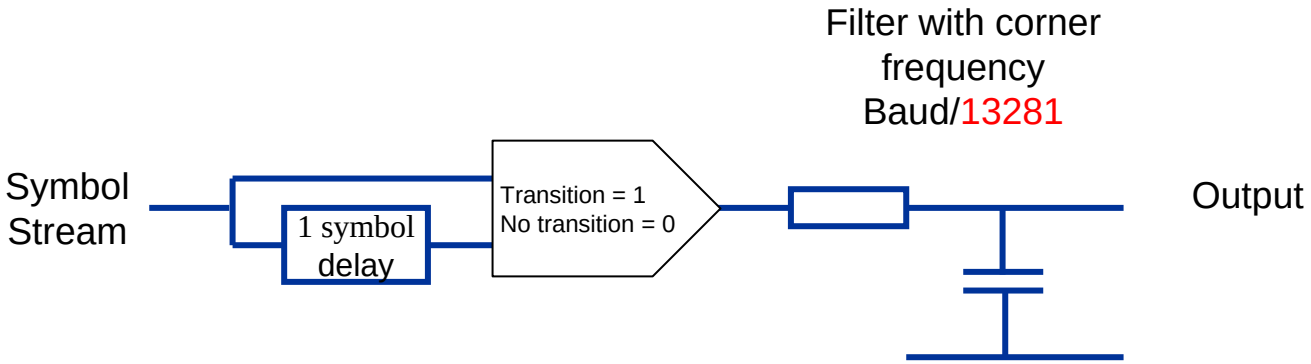
Clock content

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.

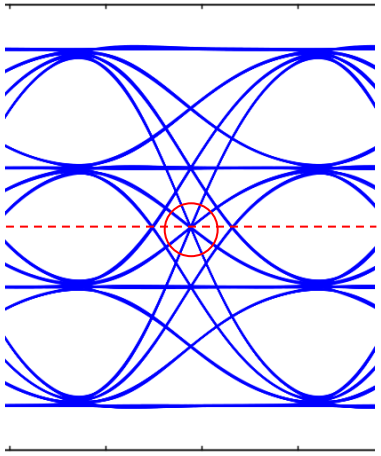
This analysis defines a transition as one of three possibilities (as per healey_3bs_01_1115):

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

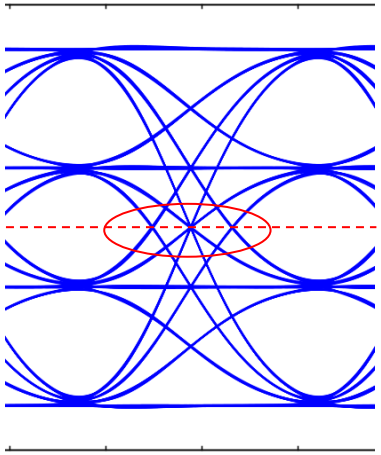


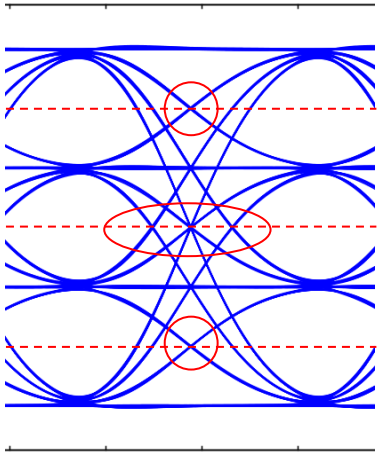
Clock content illustration

Symmetrical transitions through the signal average



Transitions through the signal average





All transitions

200Gb/s Ethernet alignment marker proposal

Table 119–1—200GBASE-R Alignment marker encodings

PCS lane number	Encoding[a] {CM_0, CM_1, CM_2, UP_0, CM_3, CM_4, CM_5, UP_1, UM_0, UM_1, UM_2, UP_2, UM_3, UM_4, UM_5}
0	0x9A, 0x4A, 0x26, 0x05, 0x65, 0xB5, 0xD9, 0xD6, 0xB3, 0xC0, 0x8C, 0x4B, 0x50, 0x79, 0x73
1	0x9A, 0x4A, 0x26, 0xCD, 0x65, 0xB5, 0xD9, 0x76, 0x4E, 0x38, 0x09, 0x7A, 0xE2, 0xA9, 0xF6
2	0x9A, 0x4A, 0x26, 0x64, 0x65, 0xB5, 0xD9, 0x08, 0x11, 0x47, 0x3D, 0xD5, 0x27, 0x57, 0xC2
3	0x9A, 0x4A, 0x26, 0xD3, 0x65, 0xB5, 0xD9, 0xA0, 0x18, 0x82, 0xDA, 0x81, 0xB2, 0xD6, 0x25
4	0x9A, 0x4A, 0x26, 0x75, 0x65, 0xB5, 0xD9, 0xA5, 0x40, 0xC2, 0x8F, 0xC0, 0xE3, 0xA7, 0x70
5	0x9A, 0x4A, 0x26, 0xB2, 0x65, 0xB5, 0xD9, 0x38, 0xF7, 0x74, 0x45, 0x74, 0x45, 0x01, 0xBA
6	0x9A, 0x4A, 0x26, 0x50, 0x65, 0xB5, 0xD9, 0xA3, 0xB1, 0x58, 0xFE, 0x78, 0xDA, 0x1B, 0x01
7	0x9A, 0x4A, 0x26, 0x7B, 0x65, 0xB5, 0xD9, 0xC4, 0x19, 0xA1, 0x39, 0xFD, 0xC6, 0xAA, 0xC6

[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 -40 to +40.

The worst case PCS lane combinations and delays were then used to generate the worst case PDFs for 200 GbE scrambled idle 100 Gb/s lanes.

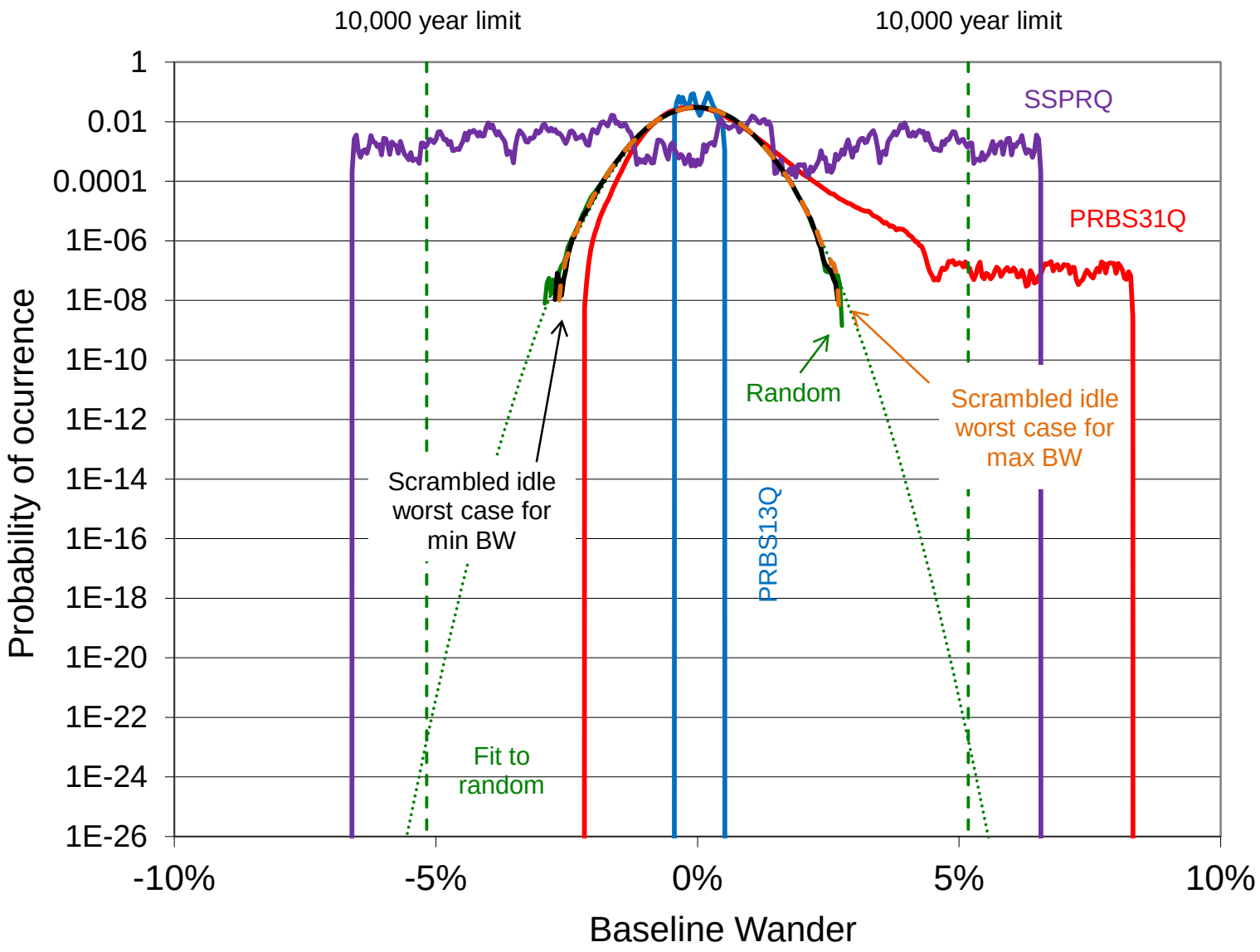
Scrambled idle construction

The scrambled idle symbol streams generated for this analysis were:

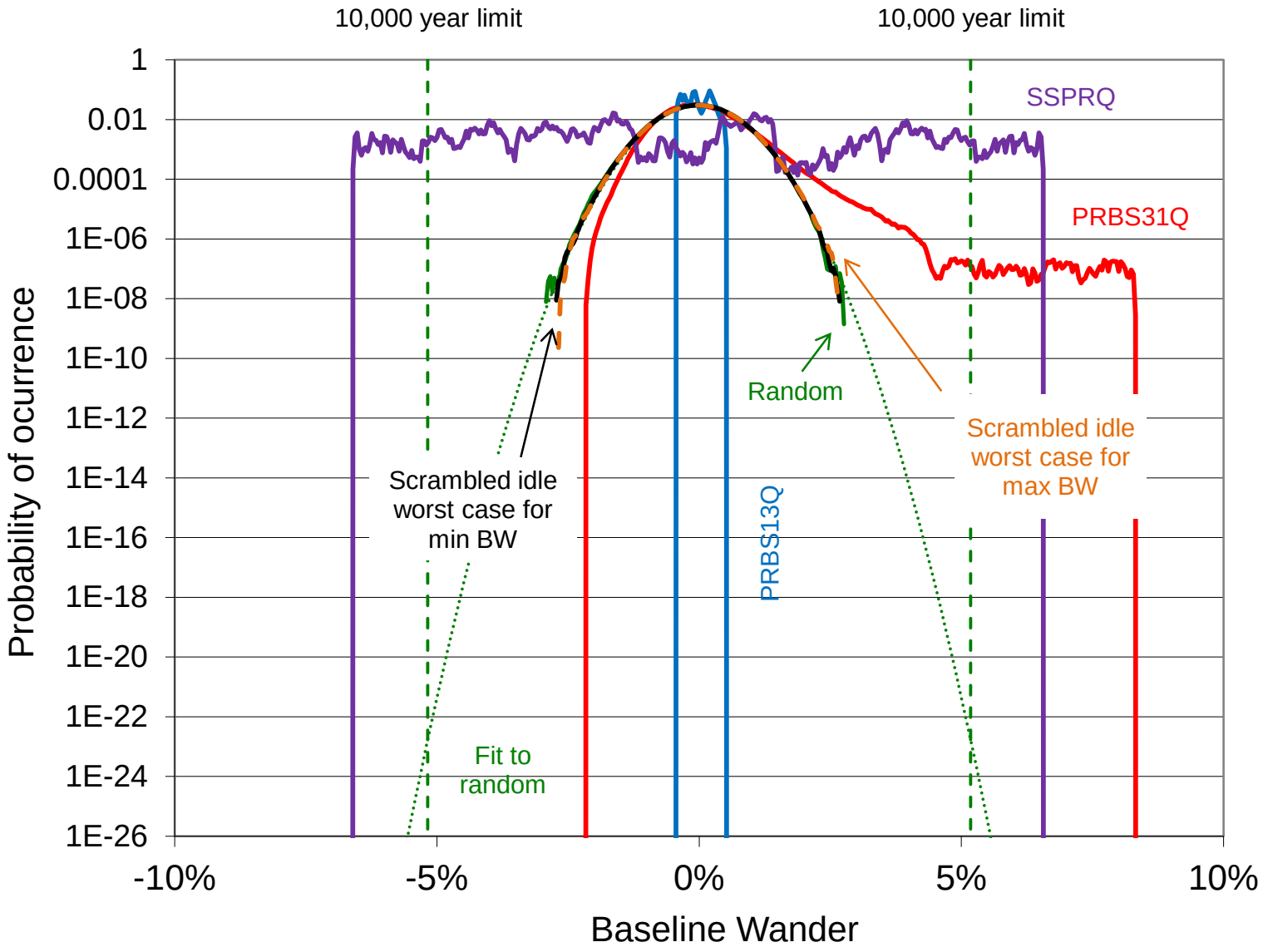
- Idle control characters
- 256B/257B transcoded
- Scrambled
- Distributed 10 bits at a time to two FEC codewords which start with alignment markers followed by 136 bits of PRBS9 one in every 4096 code words
- 300 bits of RS(544,514) FEC parity added
- Interleaved 10 bits at a time to form PCS lanes (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.

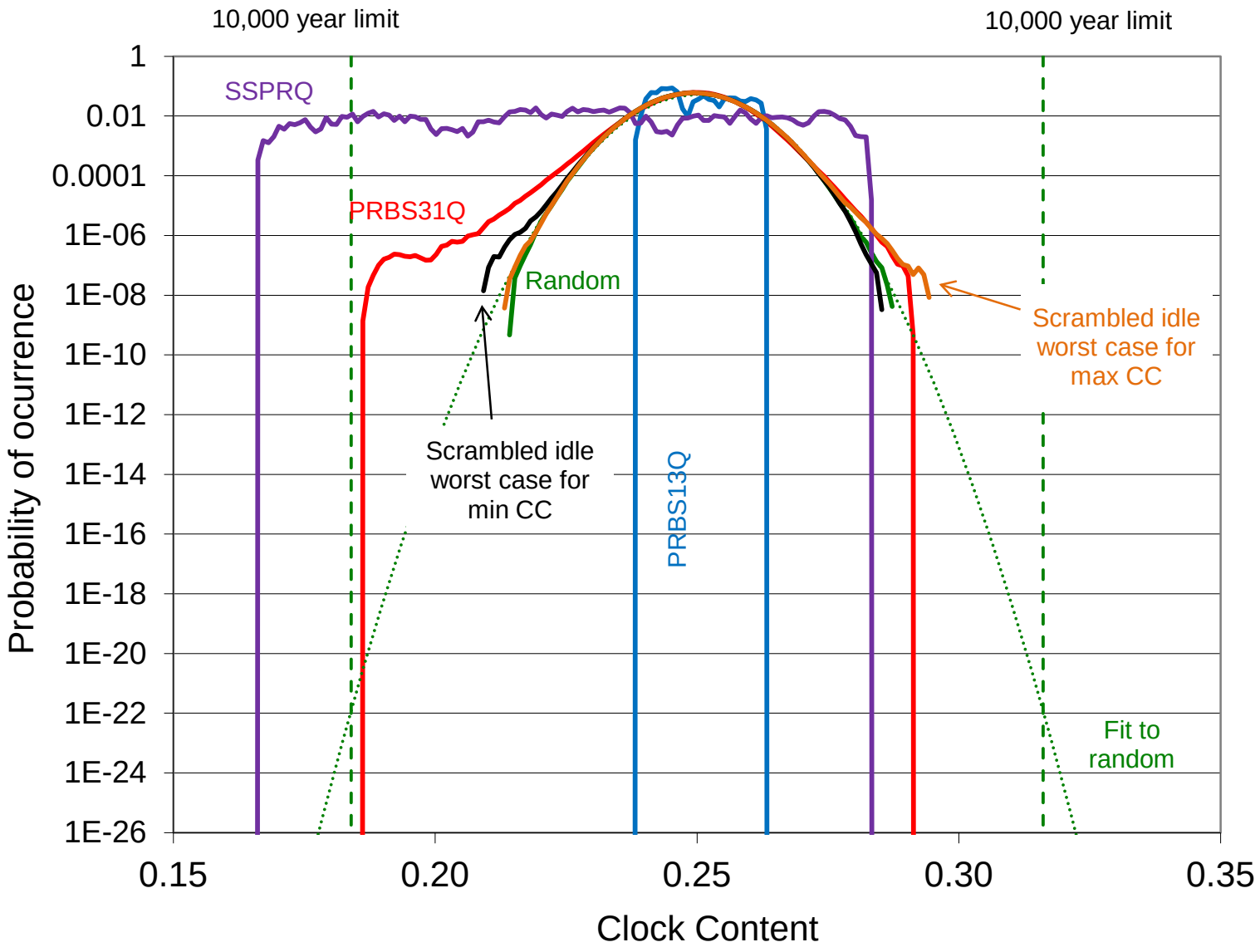
Baseline wander, 100G lanes, 400G markers



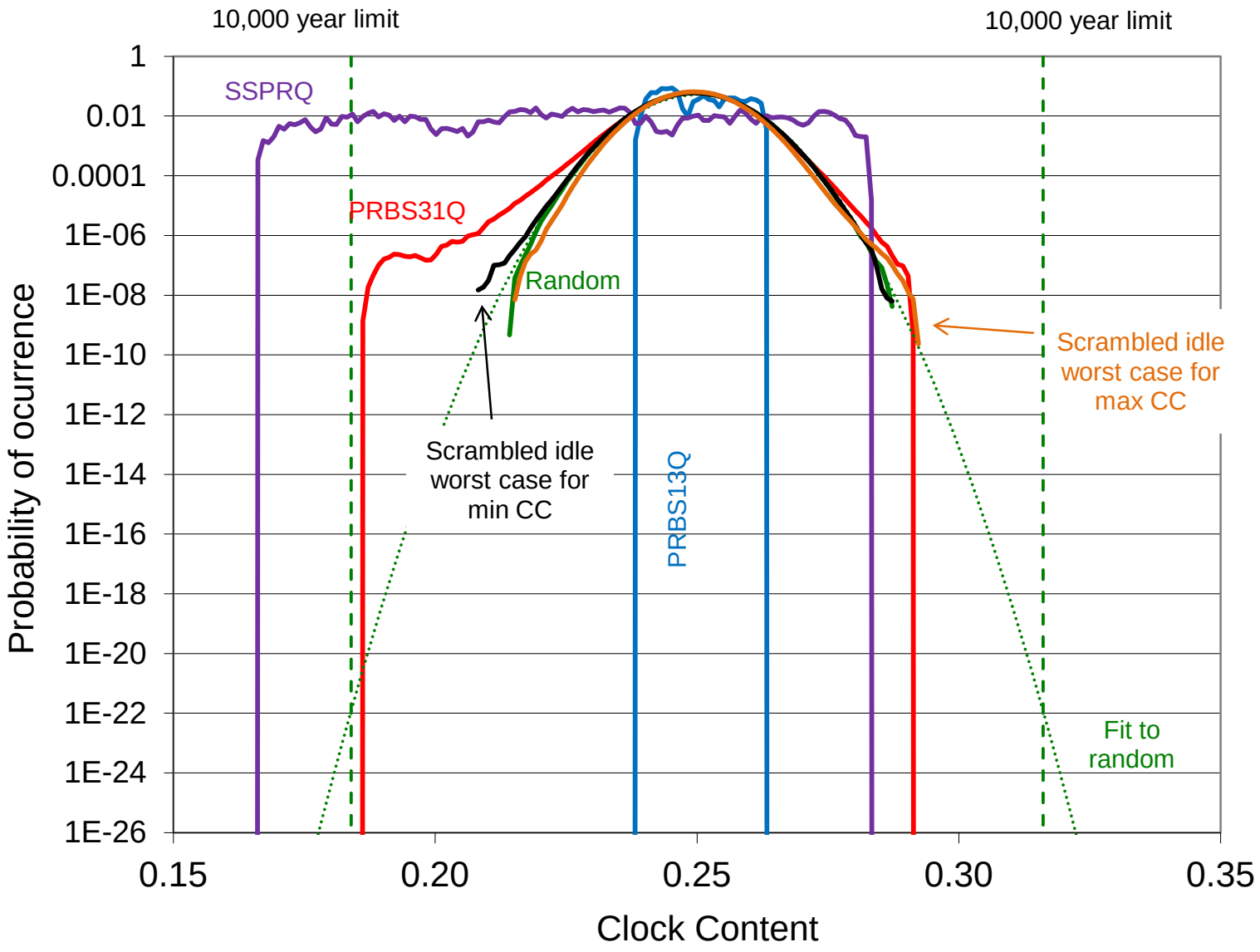
Baseline wander, 100G lanes, 200G markers



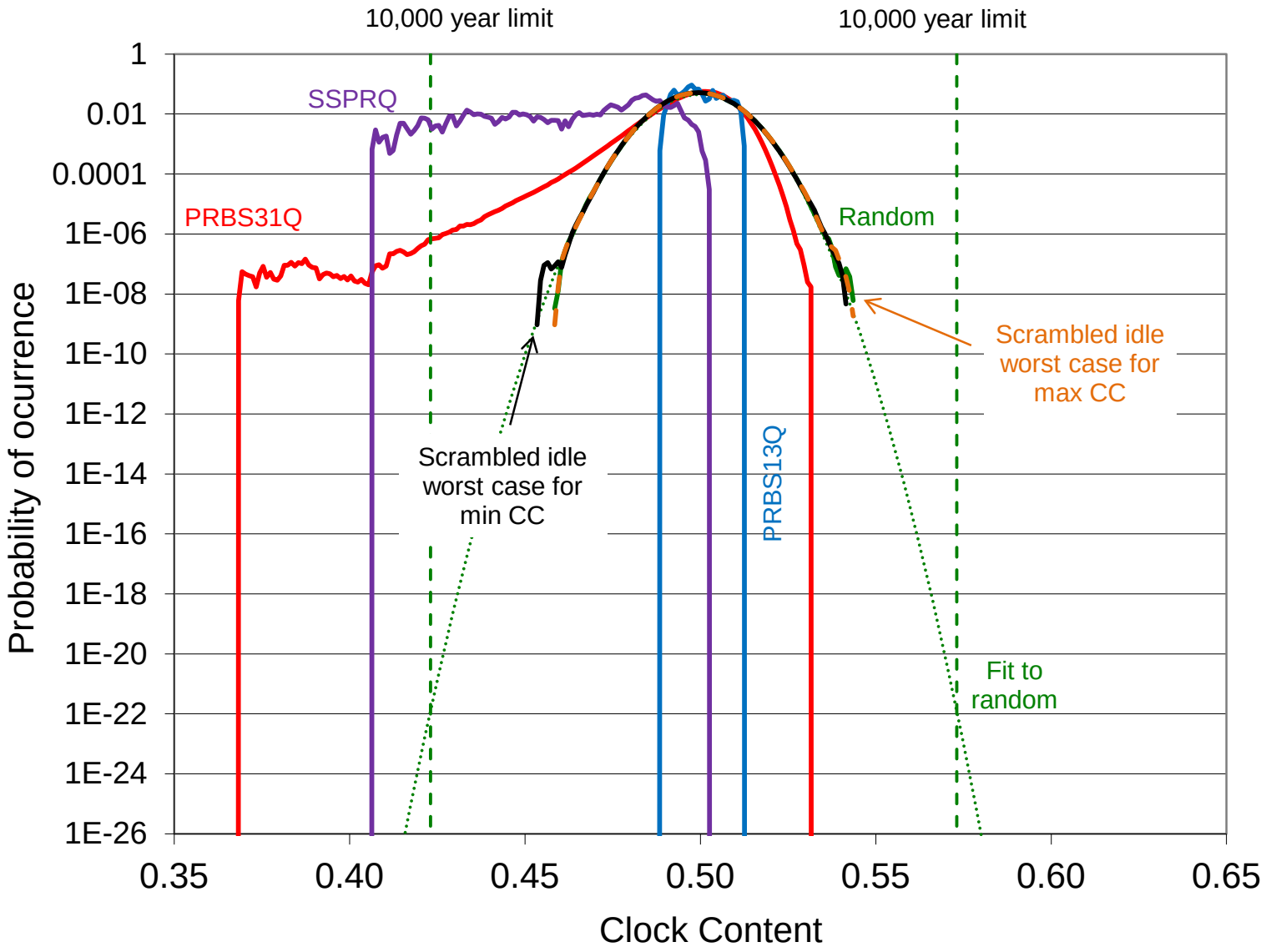
Clock, sym trans through ave, 100G lanes, 400G



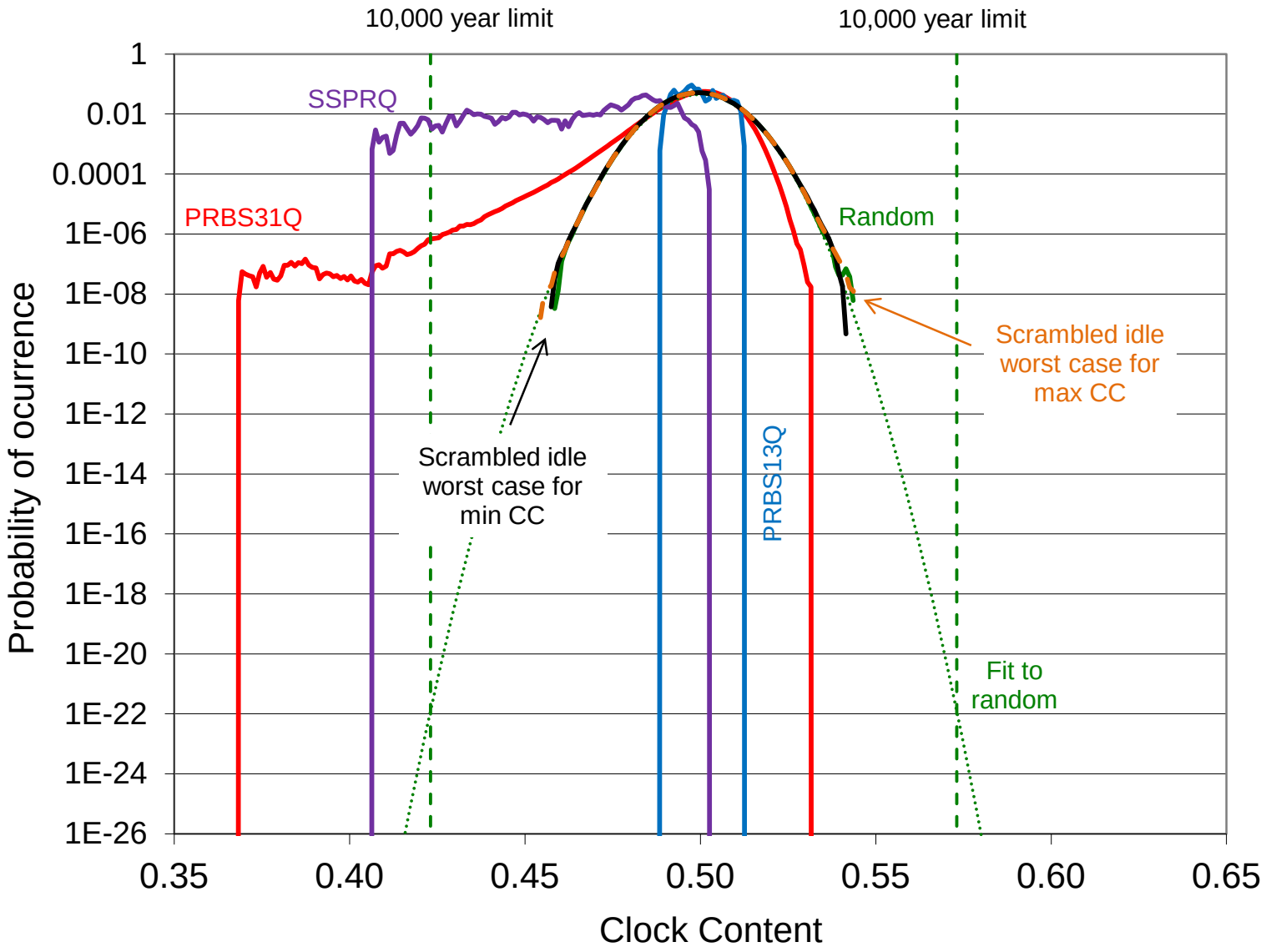
Clock, sym trans through ave, 100G lanes, 200G



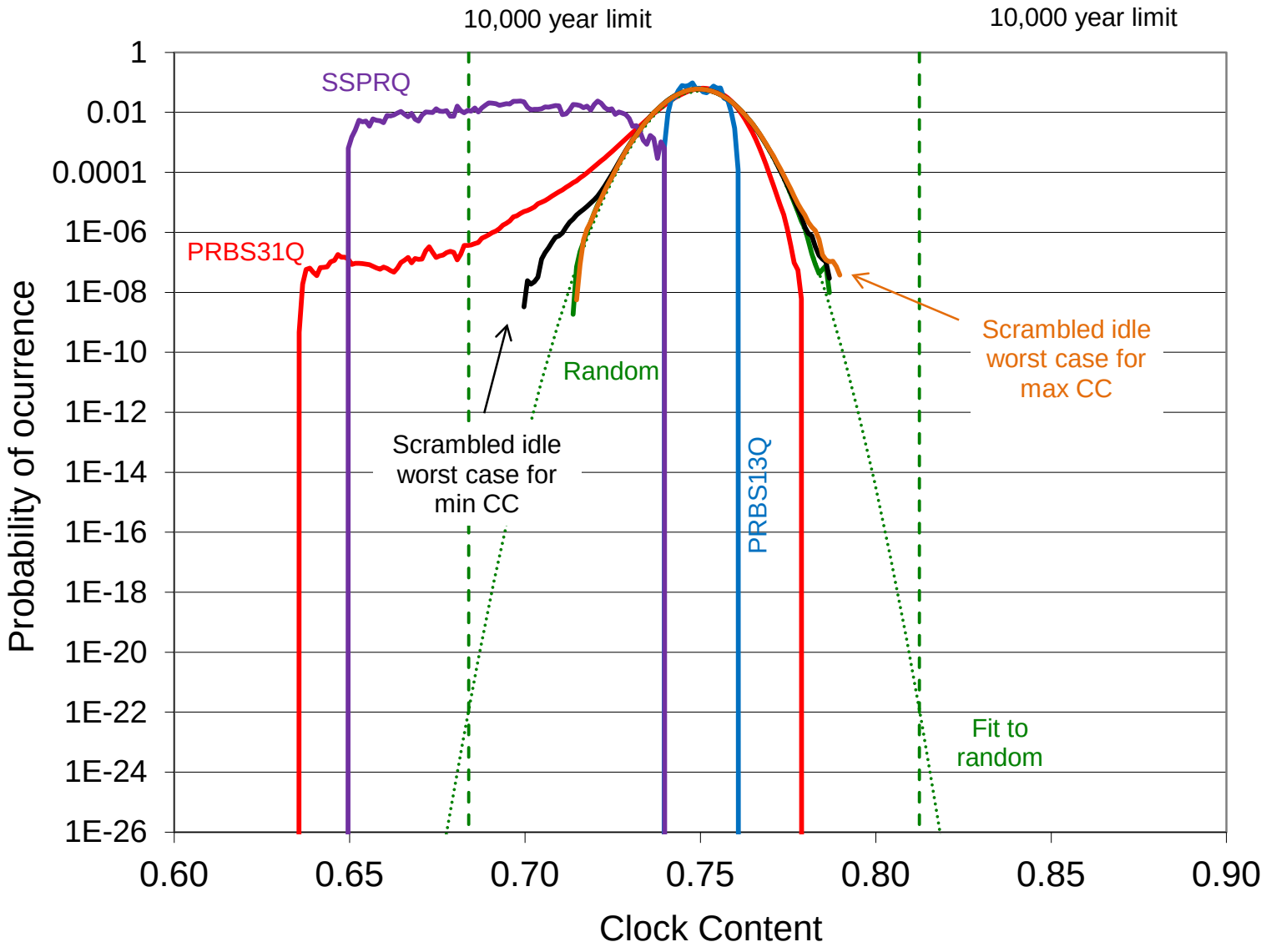
Clock, trans through ave, 100G lanes, 400G



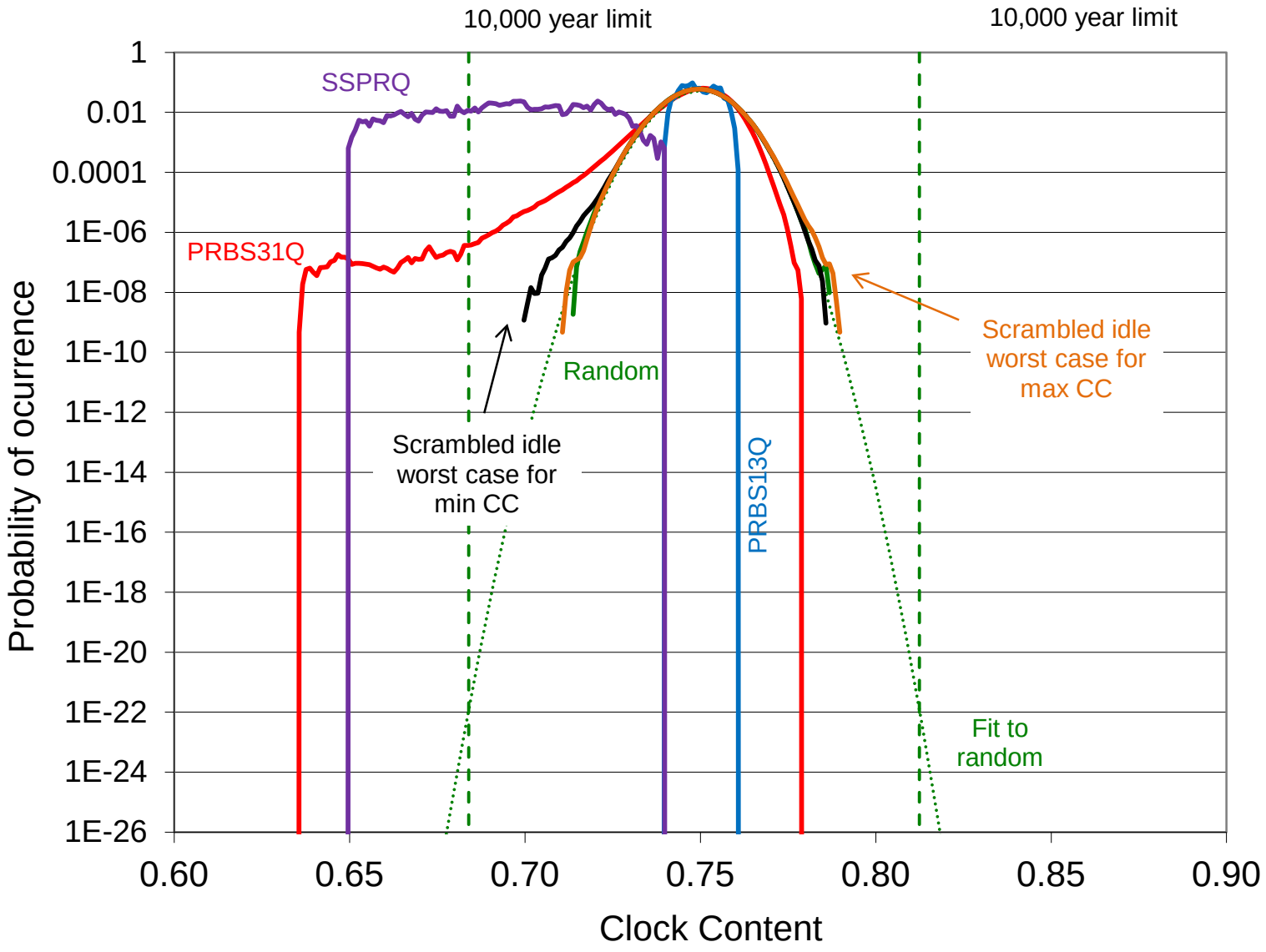
Clock, trans through ave, 100G lanes, 200G



Clock, all transitions, 100G lanes, 400G



Clock, all transitions, 100G lanes, 200G



Conclusion

The baseline wander and clock content for the proposed 200G alignment markers don’t show any worse “shoulders” than the corresponding 400G plots do.

It is therefore proposed to use these alignment markers for 200 Gb/s Ethernet.

Backup

Worst case lane combinations 200GbE markers

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	2	4	3	7	0	0	-1	-1
wander_min	6	0	7	1	0	32	-3	-4
clock25_max	0	1	2	5	0	1	1	1
clock25_min	2	5	6	4	0	-2	-3	-1
clock50_max	5	0	7	1	0	39	-6	-17
clock50_min	2	0	7	1	0	2	0	2
clock75_max	1	5	7	3	0	1	1	-1
clock75_min	0	2	6	4	0	0	0	0

Thanks!