Report of Entropy estimates based on NIST SP 800-90B non-IID track

2024-Oct-20 17:19:16.351270

1 Identification information

1.1 Identification of acquisition data from entropy source

Table 1 Identification information of acquisition data from entropy source

URL of the acquisition data	`https://github.com/usnistgov/SP800-90B_EntropyAssessment/blob/master/bin/truerand_4bit.bin`
SHA-256 hash value of the acquisition data [hex]	`489bc841 bb364ba8 6da70b16 17138aef 76b25dd9 196ad669 eef40c14 41b6cb88`

- Name of the submitter of the acquisition data :
- Brief explanation of the acquisition data (or entropy source) :

1.2 Identification of analysis environment

Table 2 Identification information of analysis environment

Analysis tool	Name	Another entropy estimation tool with extensions
	Versioning information	1.0.56
	built as	64-bit application
	built by	Intel C++ Compiler (`__INTEL_LLVM_COMPILER`: 20240201)
	linked libraries	Boost C++ 1.86.0
Analysis environment	Hostname	[redacted]
	CPU information	AMD Ryzen [redacted]
	Physical memory size	[redacted] MiB
	OS name	`Microsoft Windows 11 Pro`
	OS version	`10.0.22621 N/A Build 22621`
	System type	64-bit
	Username	[redacted]

1.3 Identification of analysis conditions

Table 3 Identification information of analysis conditions

Number of samples	1000000
Bits per sample	4
Byte to bit conversion	Most Significant bit (MSb) first

1.4 Identification of analysis method

NIST SP 800-90B [1] 6.3 with corrections [2] is applied

2 Executive summary

2.1 Numerical results of min-entropy estimates based on non-IID track

Table 4 Numerical results

Estimator	H_{original} [a] [bit / 4 - bit]	Notes to H_{original}	$H_{\text{bitstring}}$ [b] [bit / 1 - bit]	Notes to $H_{\text{bitstring}}$
The Most Common Value Estimate	3.97119	see 3.1	0.99773	see 4.1
The Collision Estimate	—	—	0.928362	see 4.2
The Markov Estimate	—	—	0.99947	see 4.3
The Compression Estimate	—	—	0.900627	see 4.4
The t-Tuple Estimate	3.68775	see 3.2	0.929434	see 4.5
The Longest Repeated Substring (LRS) Estimate	3.93497	see 3.3	0.986687	see 4.6
Multi Most Common in Window Prediction Estimate	3.99229	see 3.4	0.99808	see 4.7
The Lag Prediction Estimate	3.97627	see 3.5	0.998649	see 4.8
The MultiMMC Prediction Estimate	3.98526	see 3.6	0.998205	see 4.9
The LZ78Y Prediction Estimate	3.98428	see 3.7	0.999355	see 4.10
The intial entropy source estimate [bit / 4 - bit] $H_I = \min(H_{\text{original}}, 4 \times H_{\text{bitstring}})$	3.60251			

[a] Entropy estimate of the sequential dataset [source: NIST SP 800-90B [1] 3.1.3]

[b] An additional entropy estimation (per bit) for the non-binary sequential dataset [see NIST SP 800-90B [1] 3.1.3]

2.2 Visual comparison of min-entropy estimates from original samples

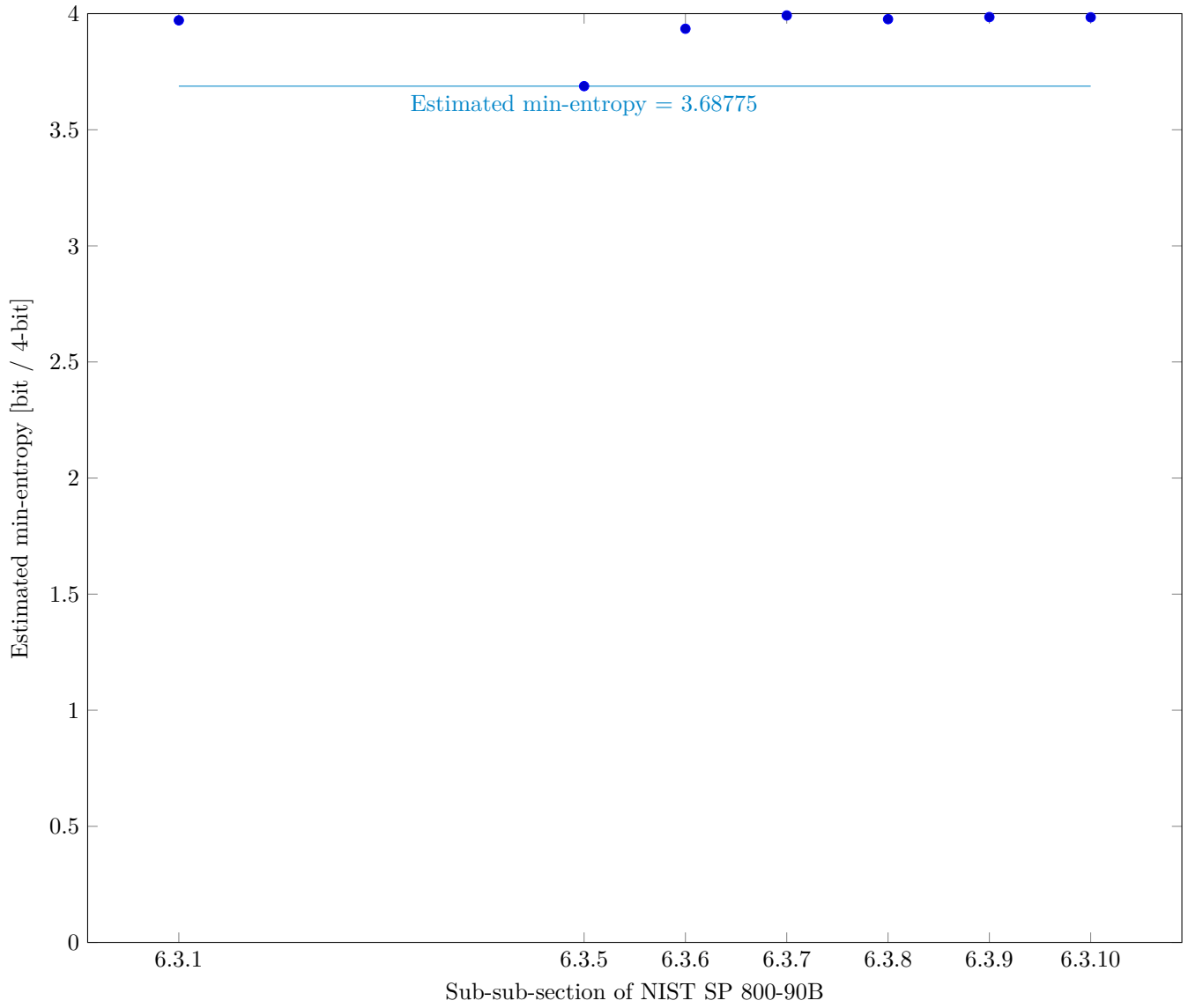

Fig. 1 Estimated Min-Entropy using §6.3 of NIST SP 800-90B

2.3 Visual comparison of min-entropy estimates by interpreting each sample as bitstring

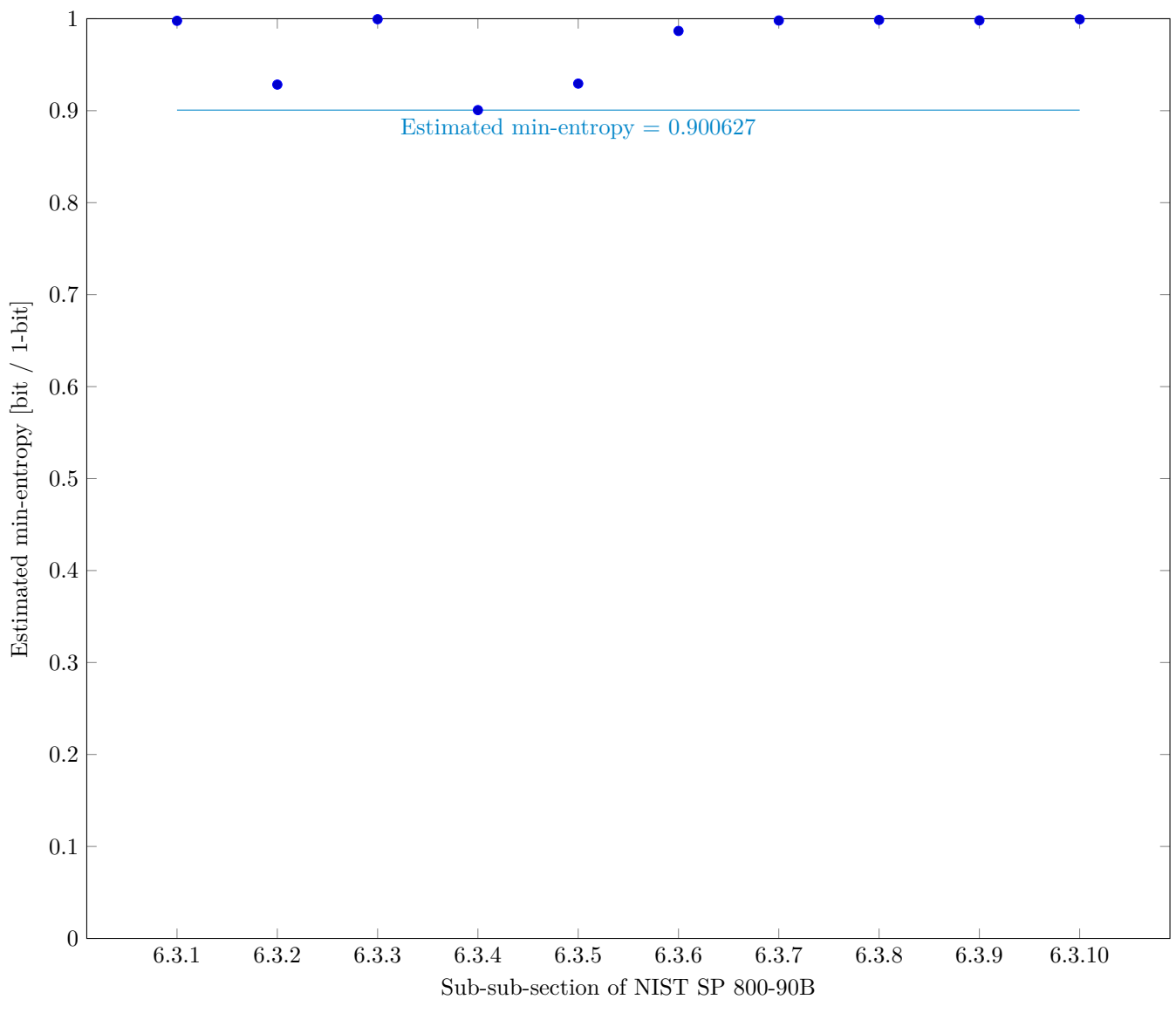

Fig. 2 Estimated Min-Entropy using §6.3 of NIST SP 800-90B

3 Detailed results of analysis from original samples

3.1 The Most Common Value Estimate (NIST SP 800-90B Section 6.3.1)

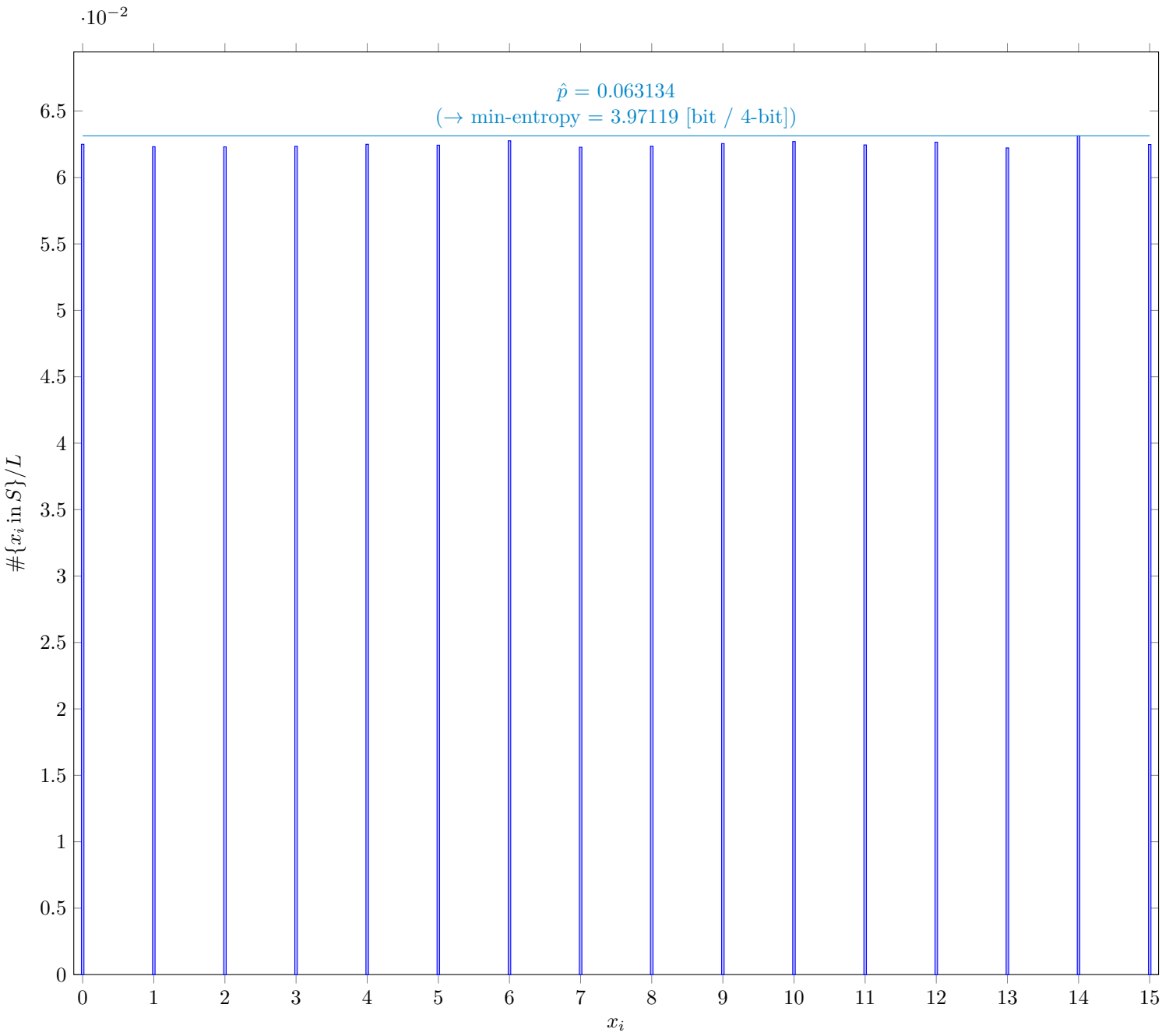

Fig. 3 Distribution of x_i

3.1.1 Supplemental information for traceability

Table 5 Supplemental information for traceability (NIST SP 800-90B Section 6.3.1)

Symbol	Value
mode	63134
$\hat{p}$	0.063134
p_u	0.0637605

3.2 The t-tuple Estimate (NIST SP 800-90B Section 6.3.5)

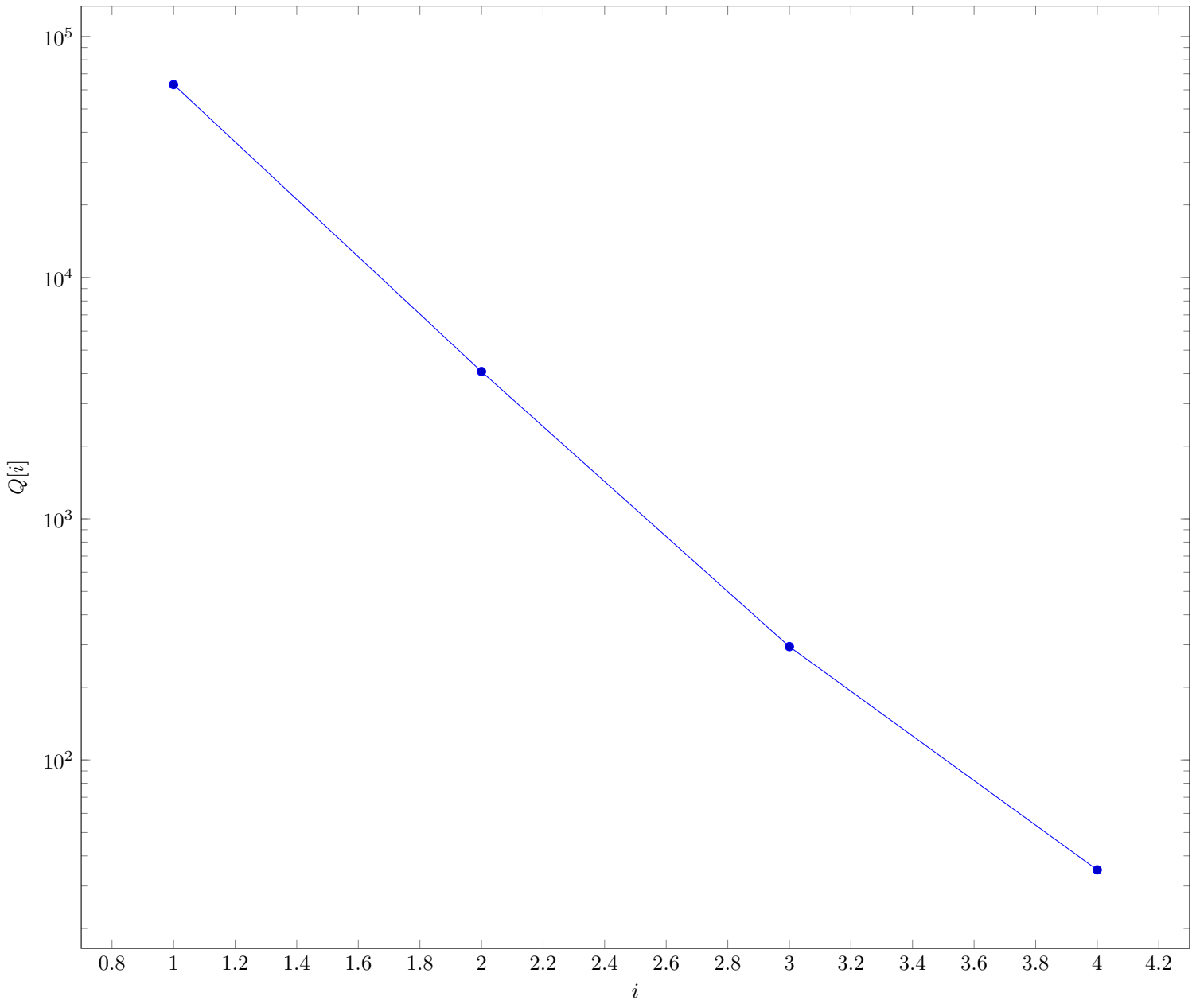

Fig. 4 Intermediate value $Q[i]$ in §6.3.5 of NIST SP 800-90B

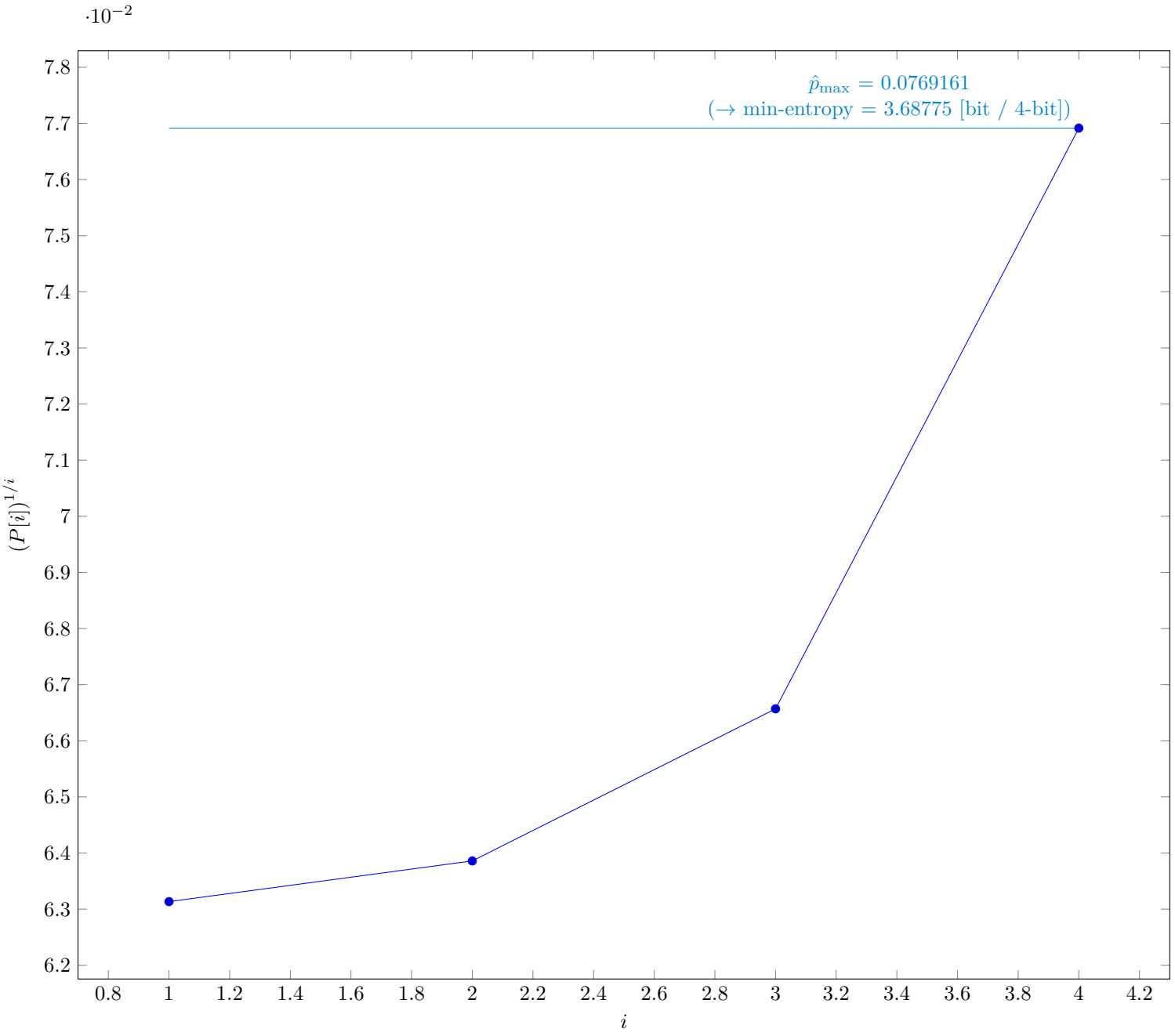

Fig. 5 $P[i]^{1/i}$ in §6.3.5 of NIST SP 800-90B

3.2.1 Supplemental information for traceability

Table 6 Supplemental information for traceability (NIST SP 800-90B Section 6.3.5)

Symbol	Value
t	4
$\hat{p}_{\max}$	0.0769161
p_u	0.0776025

3.3 The LRS Estimate (NIST SP 800-90B Section 6.3.6)

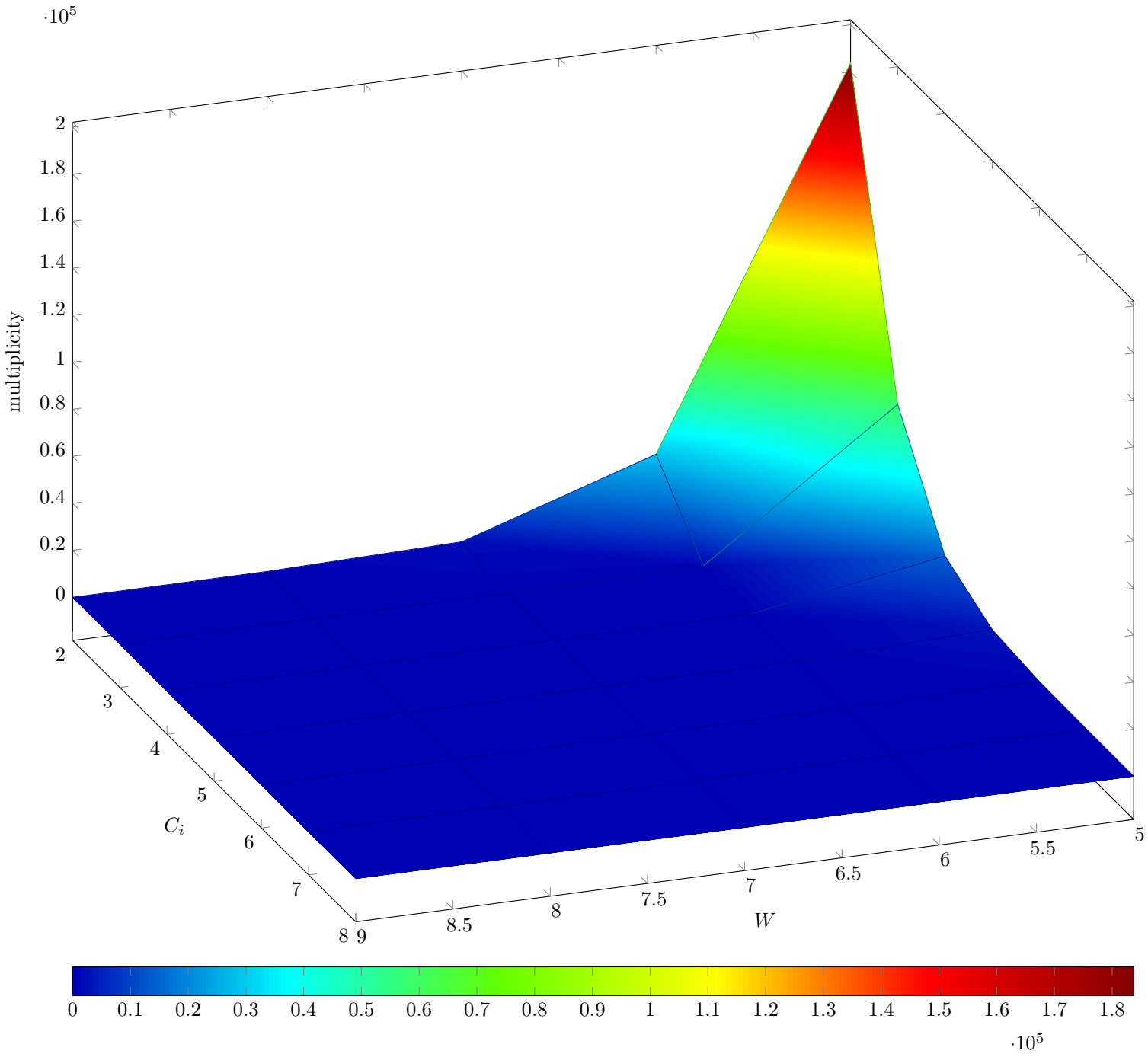

Fig. 6 Estimated W-tuple collision probability in Step 3 of §6.3.6 of NIST SP 800-90B

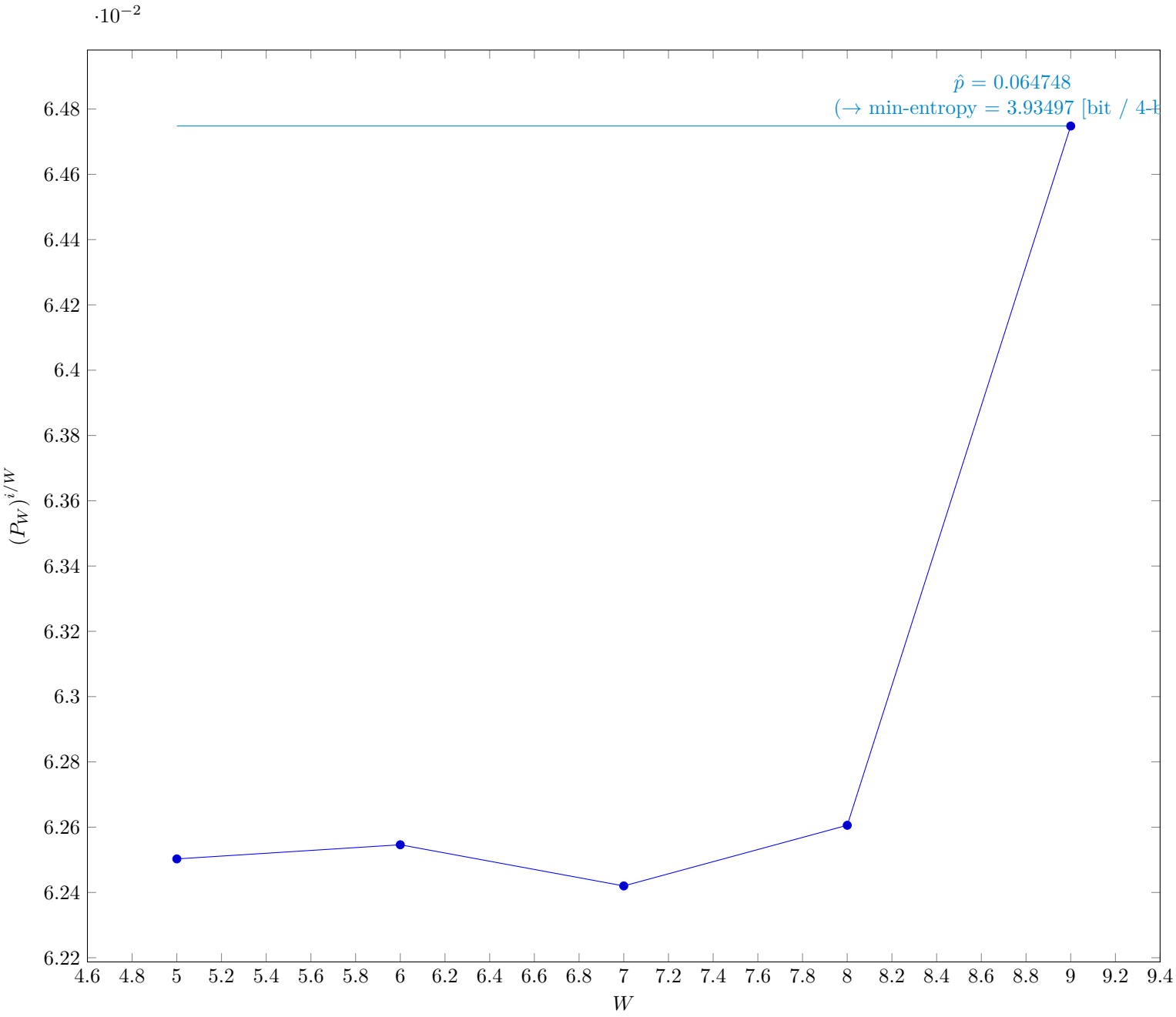

Fig. 7 Estimated average collision probability per string symbol in Step 3 of §6.3.6 of NIST SP 800-90B

3.3.1 Supplemental information for traceability

Table 7 Supplemental information for traceability (NIST SP 800-90B Section 6.3.6)

Symbol	Value
u	5
v	9
$\hat{p}$	0.064748
p_u	0.0653819

3.4 Multi Most Common in Window Prediction Estimate (NIST SP 800-90B Section 6.3.7)

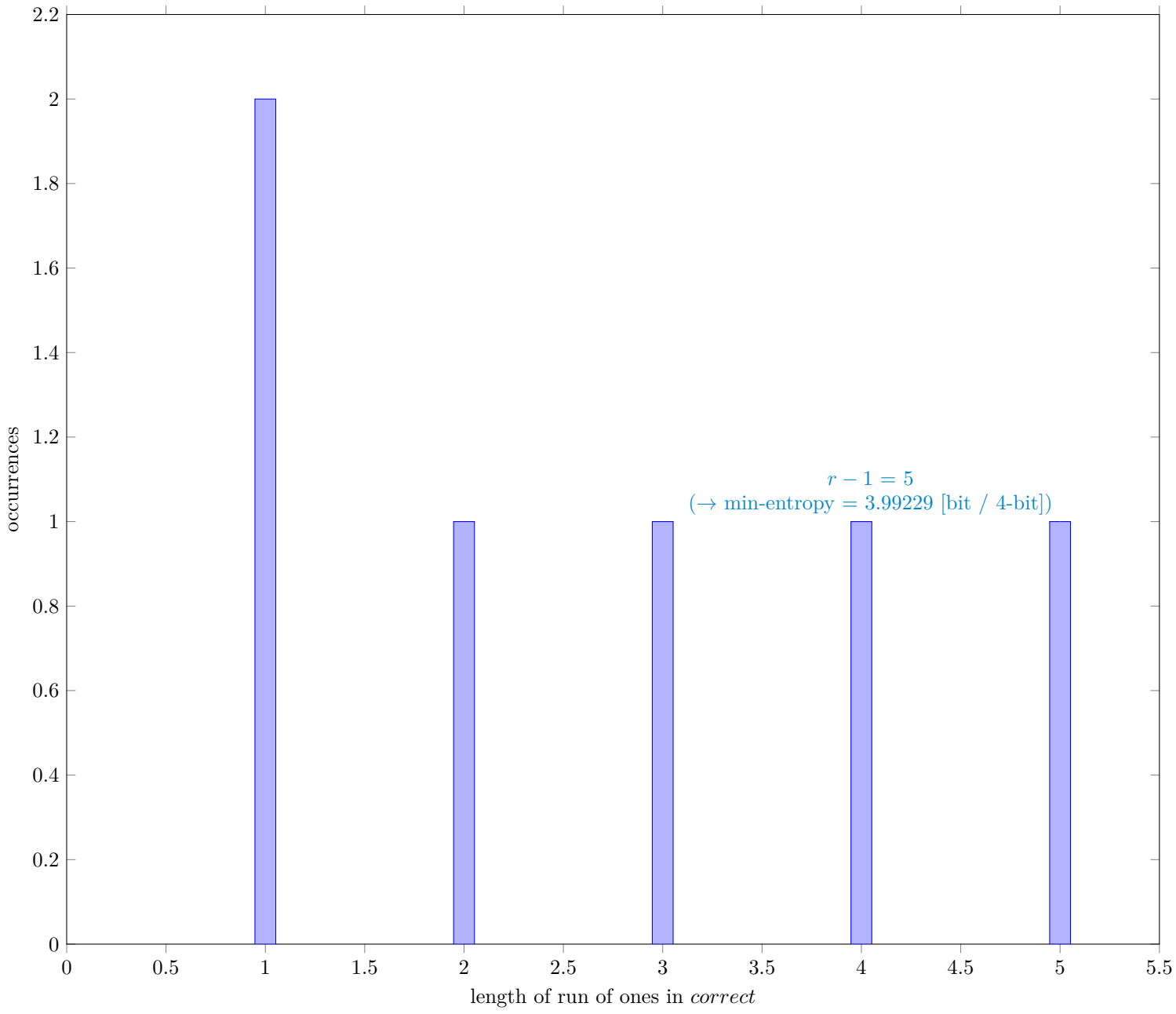

Fig. 8 Distribution of *correct*

3.4.1 Supplemental information for traceability

Table 8 Supplemental information for traceability (NIST SP 800-90B Section 6.3.7)

Symbol	Value
N	999937
C	62209
P_{global}	0.0622129
P'_{global}	0.0628351
r	6
P_{local}	0.0468281

3.5 Lag Prediction Estimate (NIST SP 800-90B Section 6.3.8)

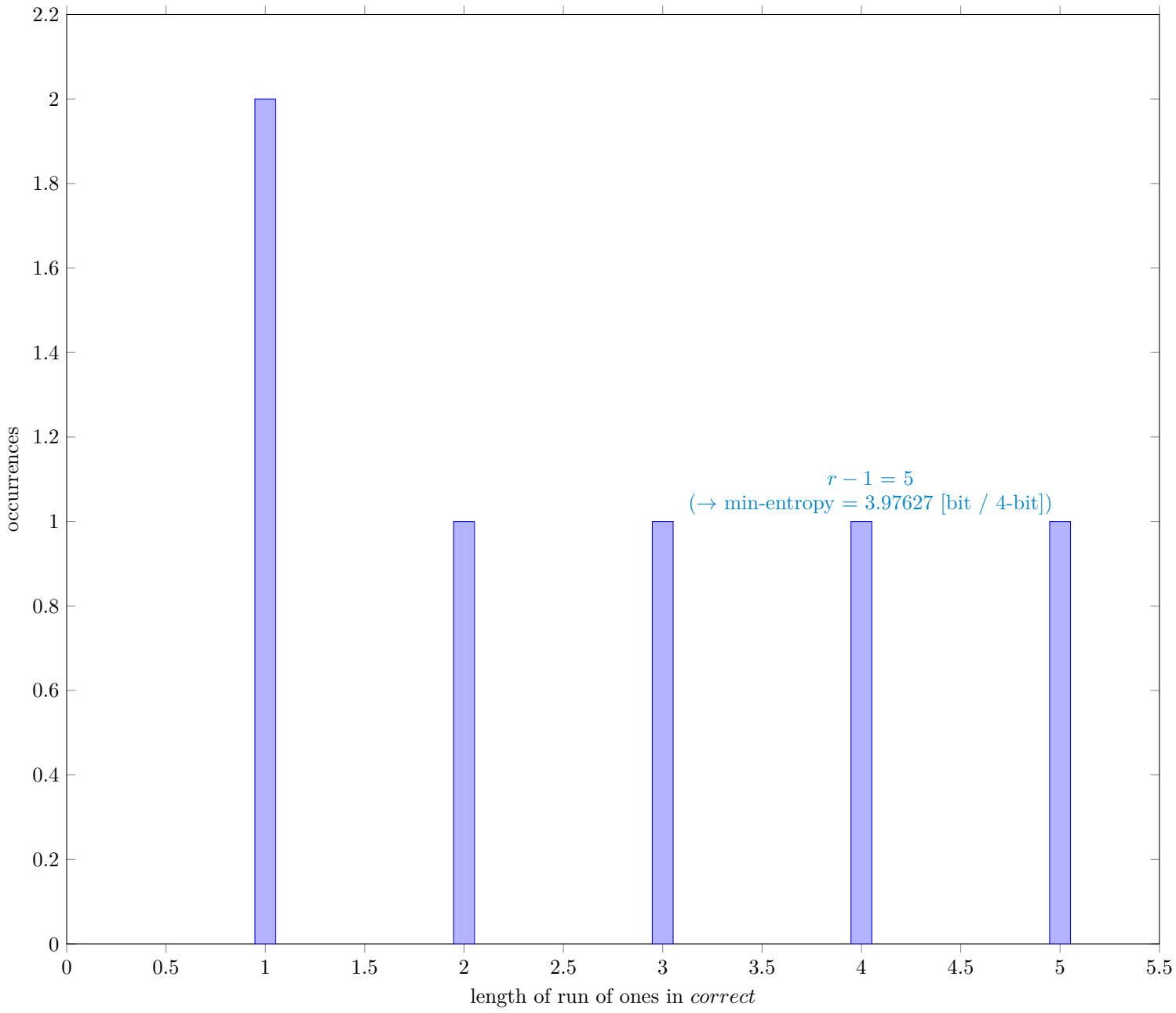

Fig. 9 Distribution of *correct*

3.5.1 Supplemental information for traceability

Table 9 Supplemental information for traceability (NIST SP 800-90B Section 6.3.8)

Symbol	Value
N	999999
C	62911
P_{global}	0.0629111
P'_{global}	0.0635365
r	6
P_{local}	0.0468276

3.6 The MultiMMC Prediction Estimate (NIST SP 800-90B Section 6.3.9)

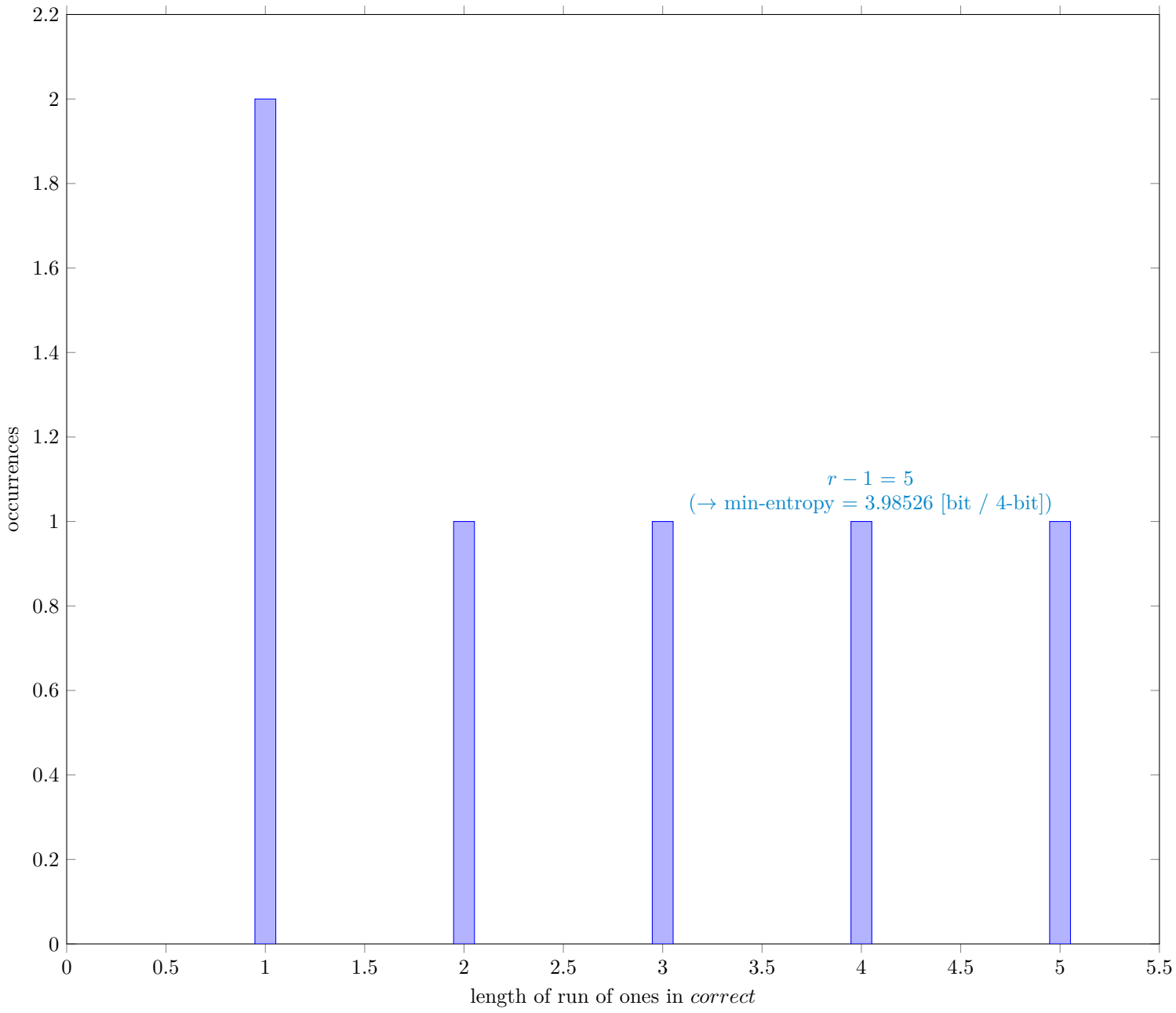

Fig. 10 Distribution of *correct*

3.6.1 Supplemental information for traceability

Table 10 Supplemental information for traceability (NIST SP 800-90B Section 6.3.9)

Symbol	Value
N	999998
C	62518
P_{global}	0.0625181
P'_{global}	0.0631417
r	6
P_{local}	0.0468276

3.7 The LZ78Y Prediction Estimate (NIST SP 800-90B Section 6.3.10)

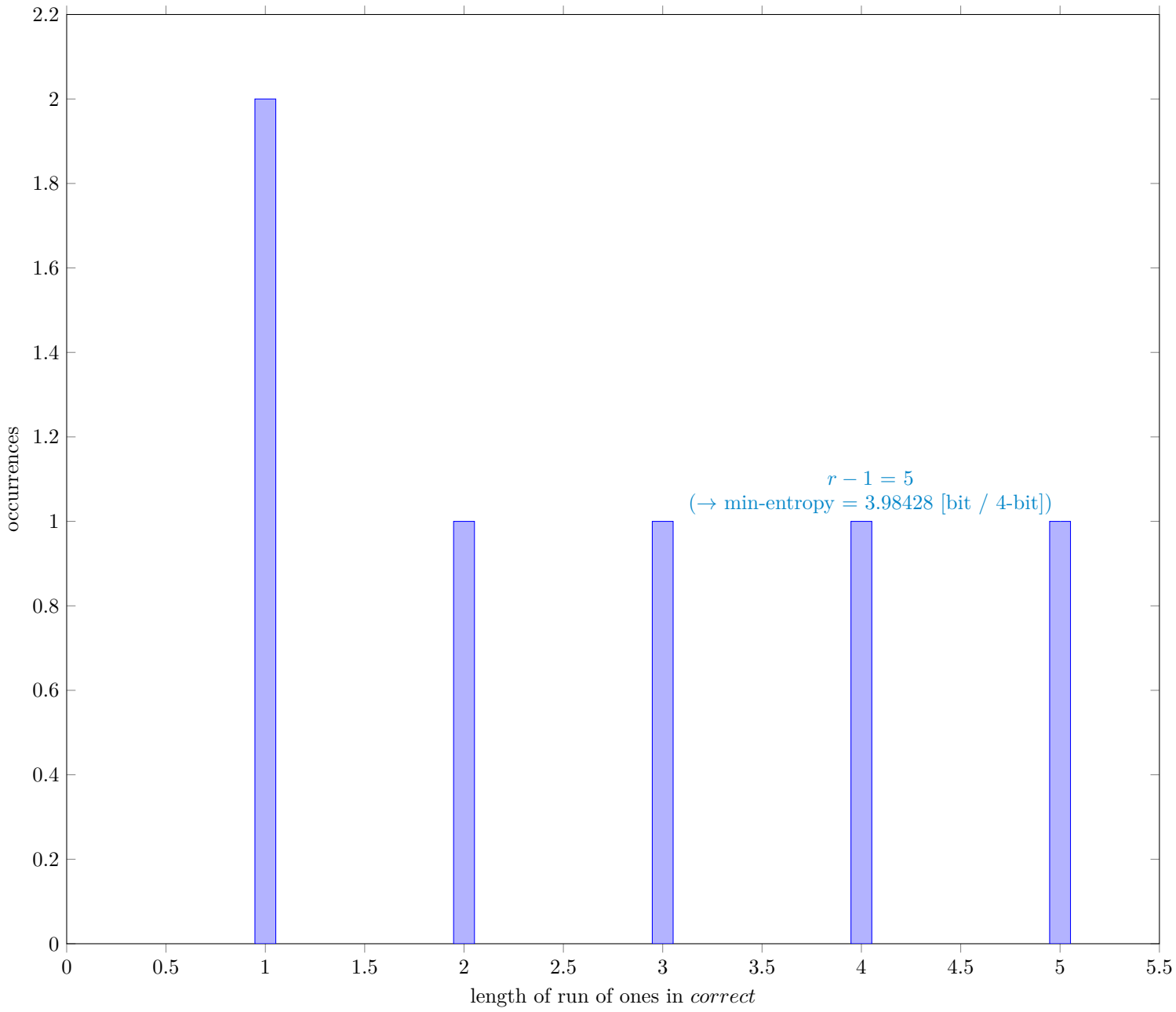

Fig. 11 Distribution of *correct*

3.7.1 Supplemental information for traceability

Table 11 Supplemental information for traceability (NIST SP 800-90B Section 6.3.10)

Symbol	Value
N	999983
C	62560
P_{global}	0.0625611
P'_{global}	0.0631849
r	6
P_{local}	0.0468277

4 Detailed results of analysis by interpreting each sample as bitstrings

4.1 The Most Common Value Estimate (NIST SP 800-90B Section 6.3.1)

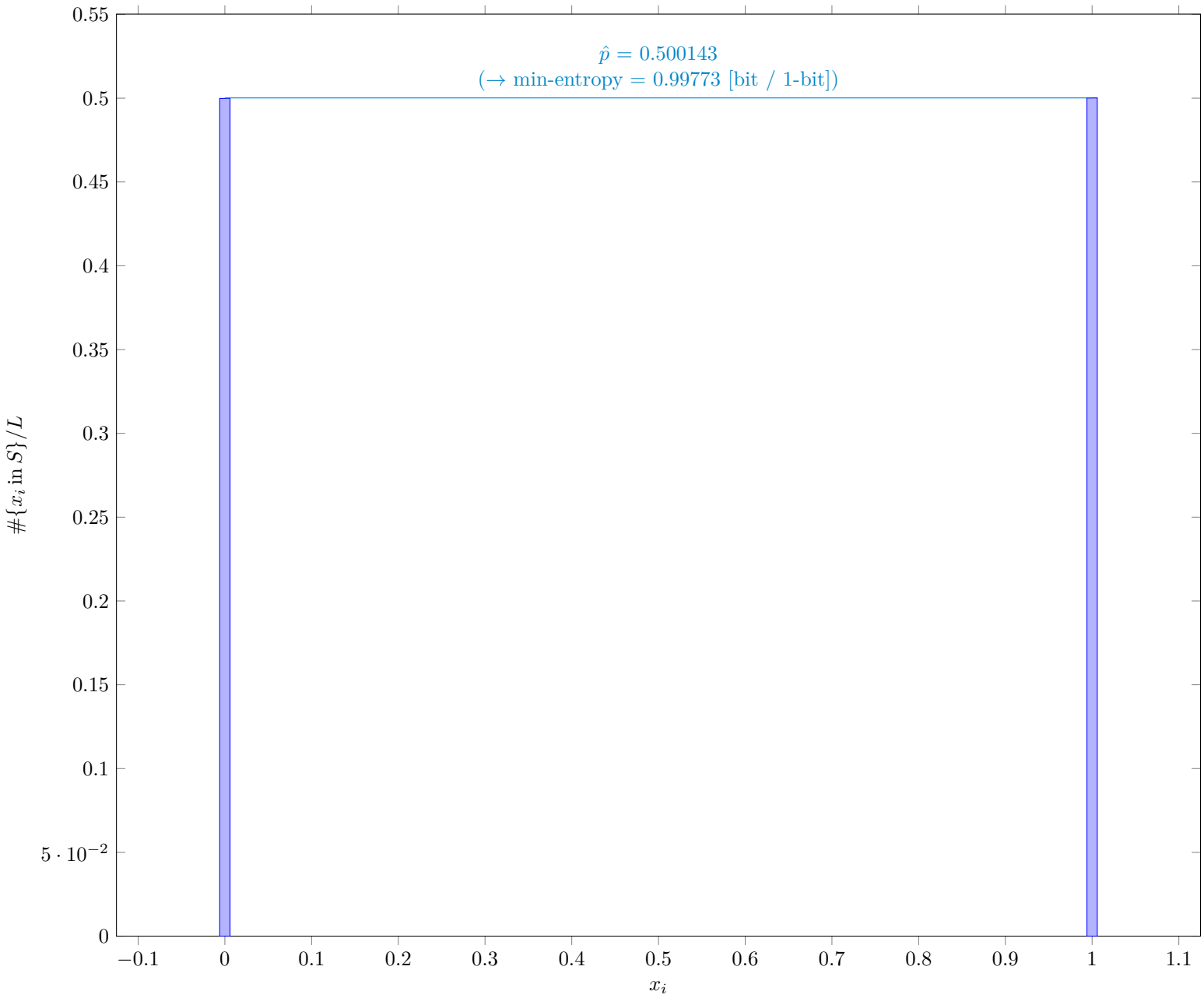

Fig. 12 Distribution of x_i

4.1.1 Supplemental information for traceability

Table 12 Supplemental information for traceability (NIST SP 800-90B Section 6.3.1)

Symbol	Value
mode	2000573
$\hat{p}$	0.500143
p_u	0.500787

4.2 The Collision Estimate (NIST SP 800-90B Section 6.3.2)

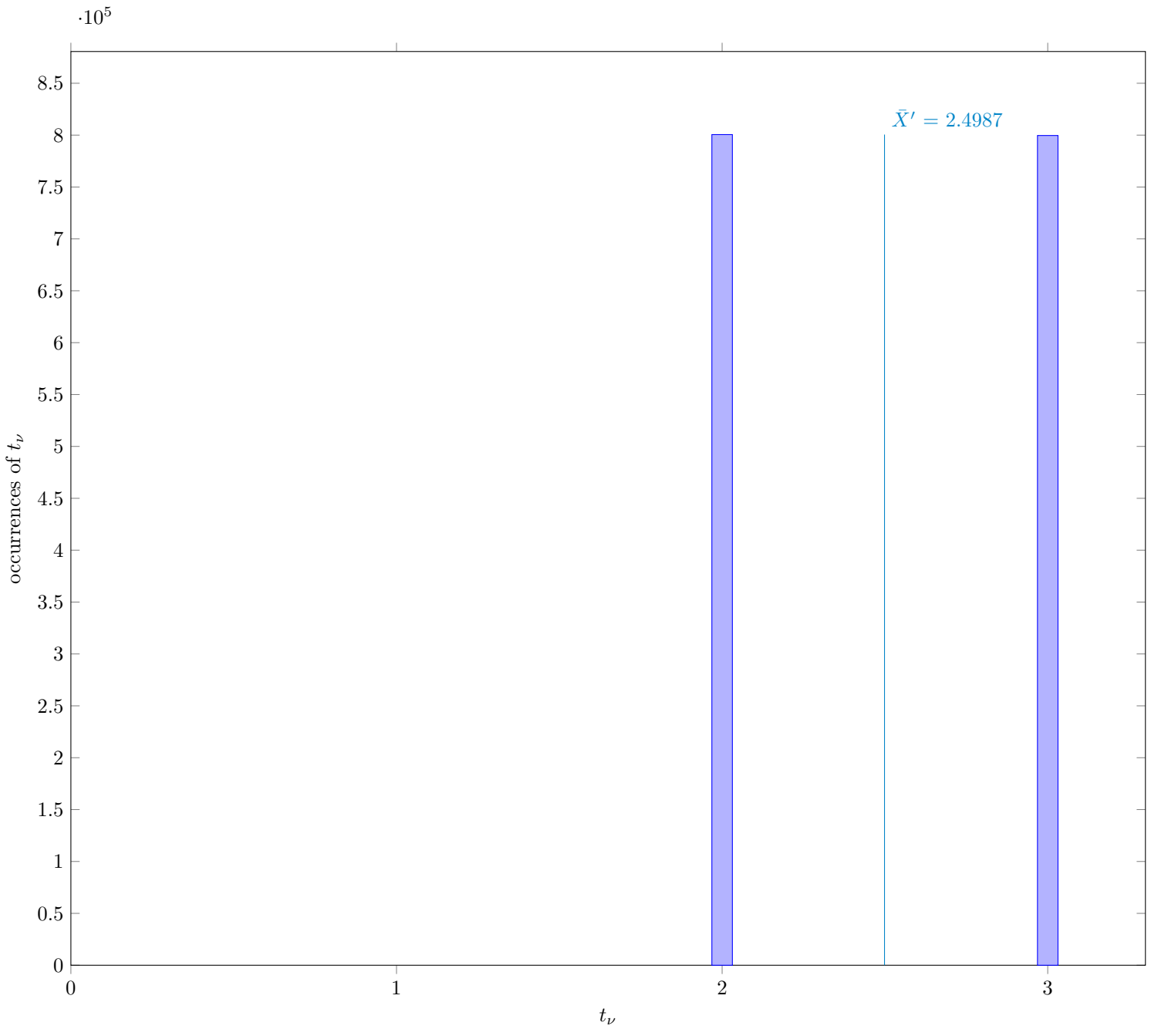

Fig. 13 Distribution of intermediate value t_ν

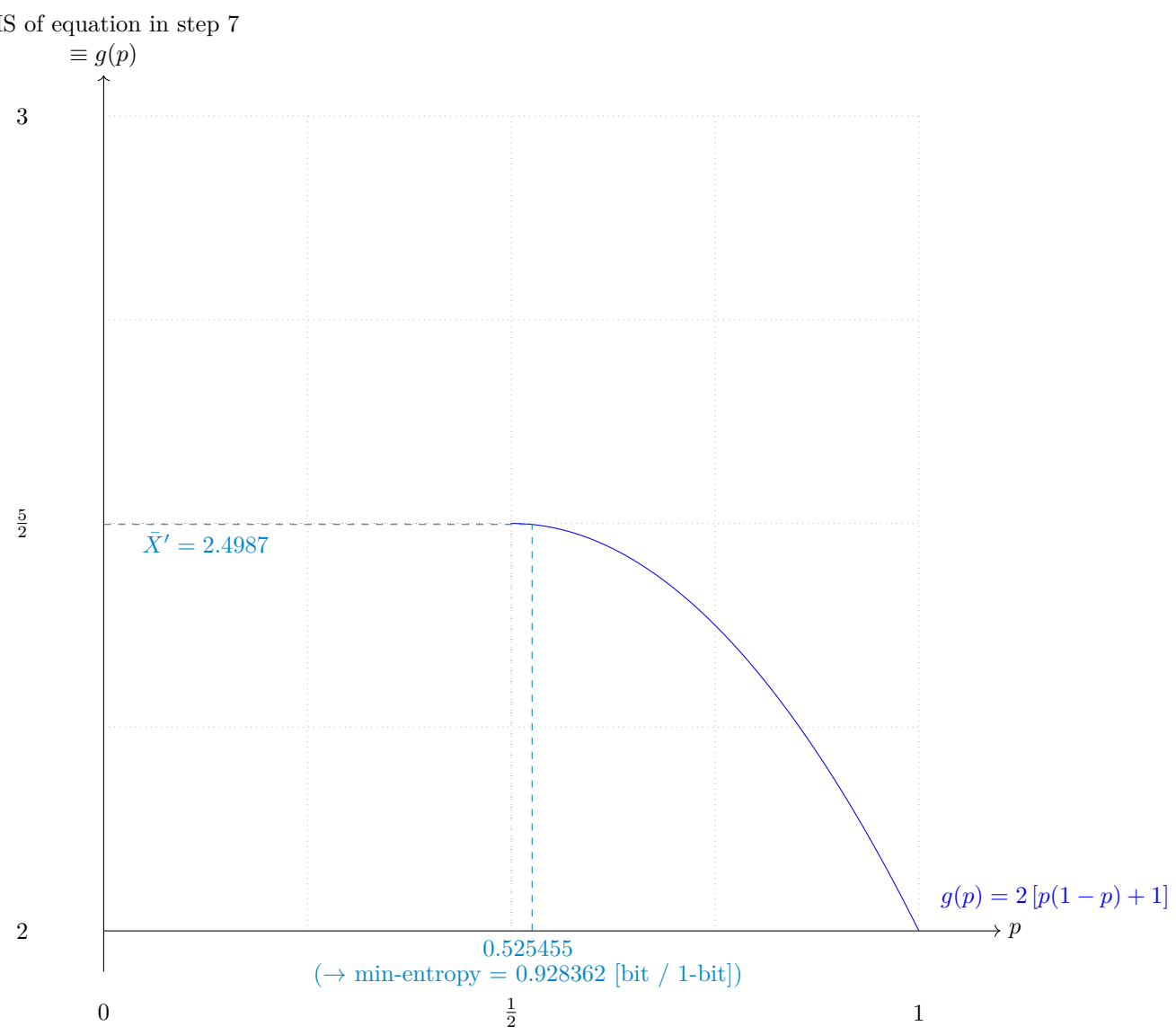

Fig. 14 Solution to the equation in step 7

4.2.1 Supplemental information for traceability

Table 13 Supplemental information for traceability (NIST SP 800-90B Section 6.3.2)

Symbol	Value
p	0.525455
$\bar{X}$	2.49972
$\bar{X}'$	2.4987
$\hat{\sigma}$	0.5

4.3 The Markov Estimate (NIST SP 800-90B Section 6.3.3)

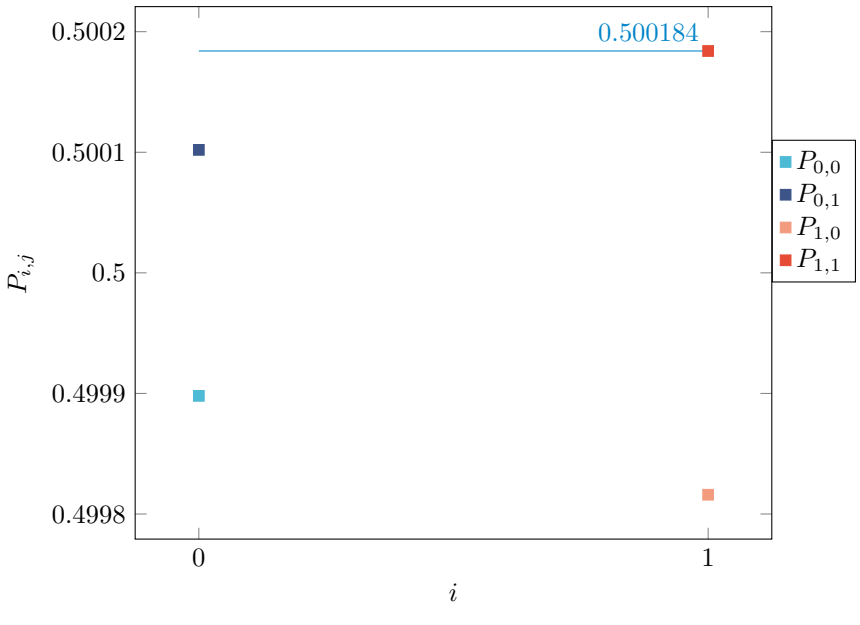

Fig. 15 Transition probability $P_{i,j}$ of §6.3.3 of NIST SP 800-90B

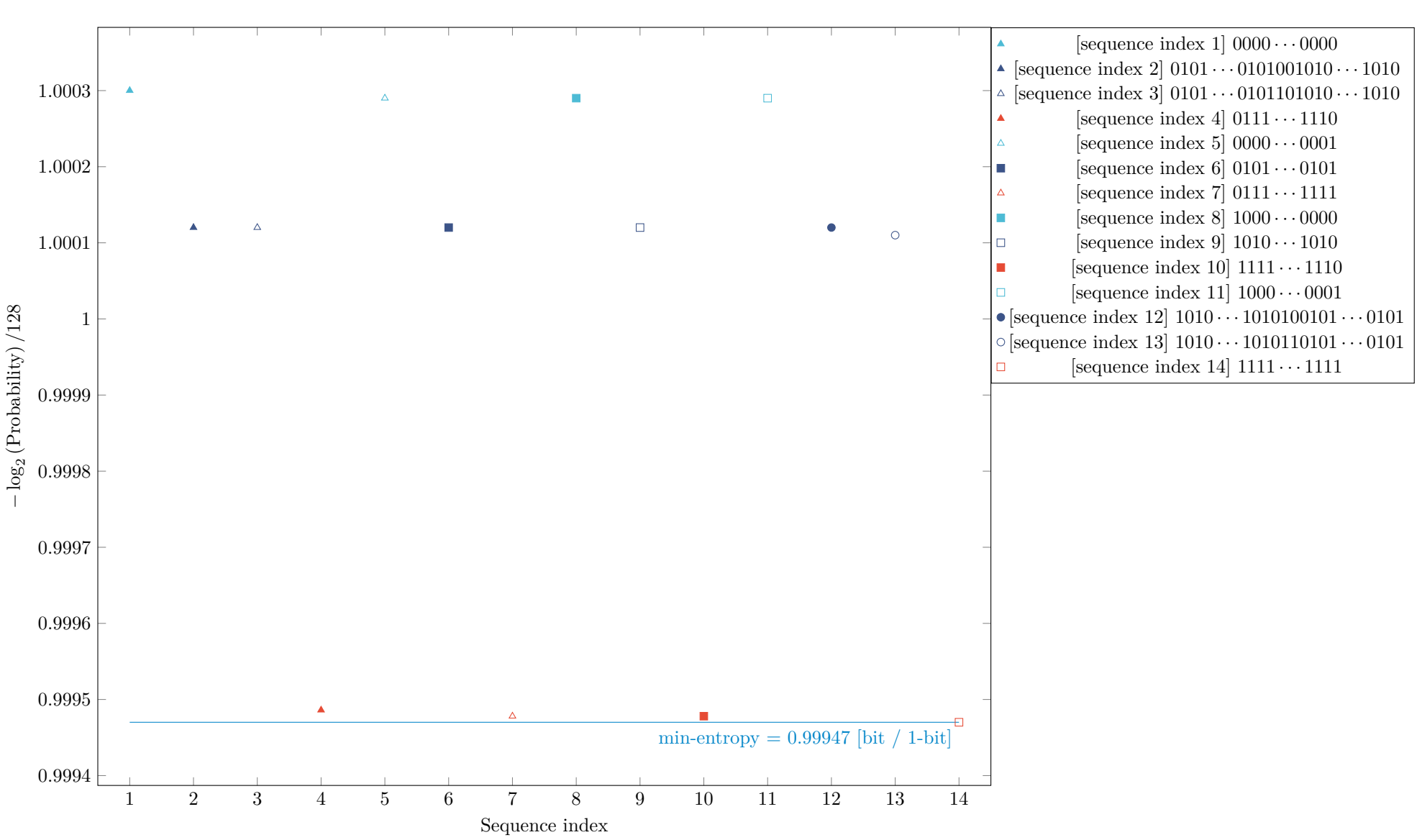

Fig. 16 Estimated Min-Entropy using §6.3.3 of NIST SP 800-90B

4.4 The Compression Estimate (NIST SP 800-90B Section 6.3.4)

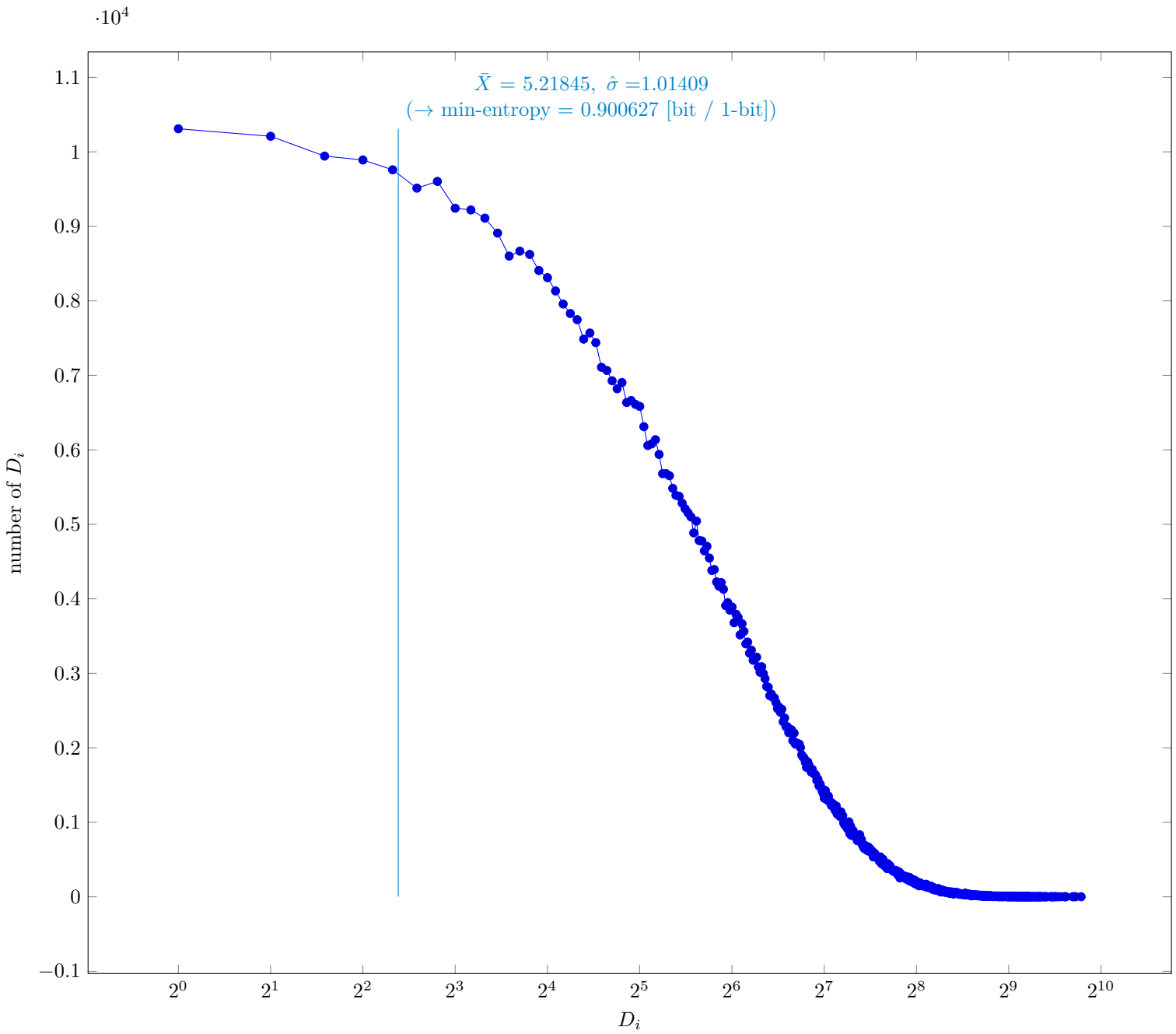

Fig. 17 Distribution of intermediate value D_i

4.4.1 Supplemental information for traceability

Table 14 Supplemental information for traceability (NIST SP 800-90B Section 6.3.4)

Symbol	Value
p	0.0236214
$\bar{X}$	5.21845
$\hat{\sigma}$	1.01409
$\bar{X}'$	5.21525

4.5 The t-tuple Estimate (NIST SP 800-90B Section 6.3.5)

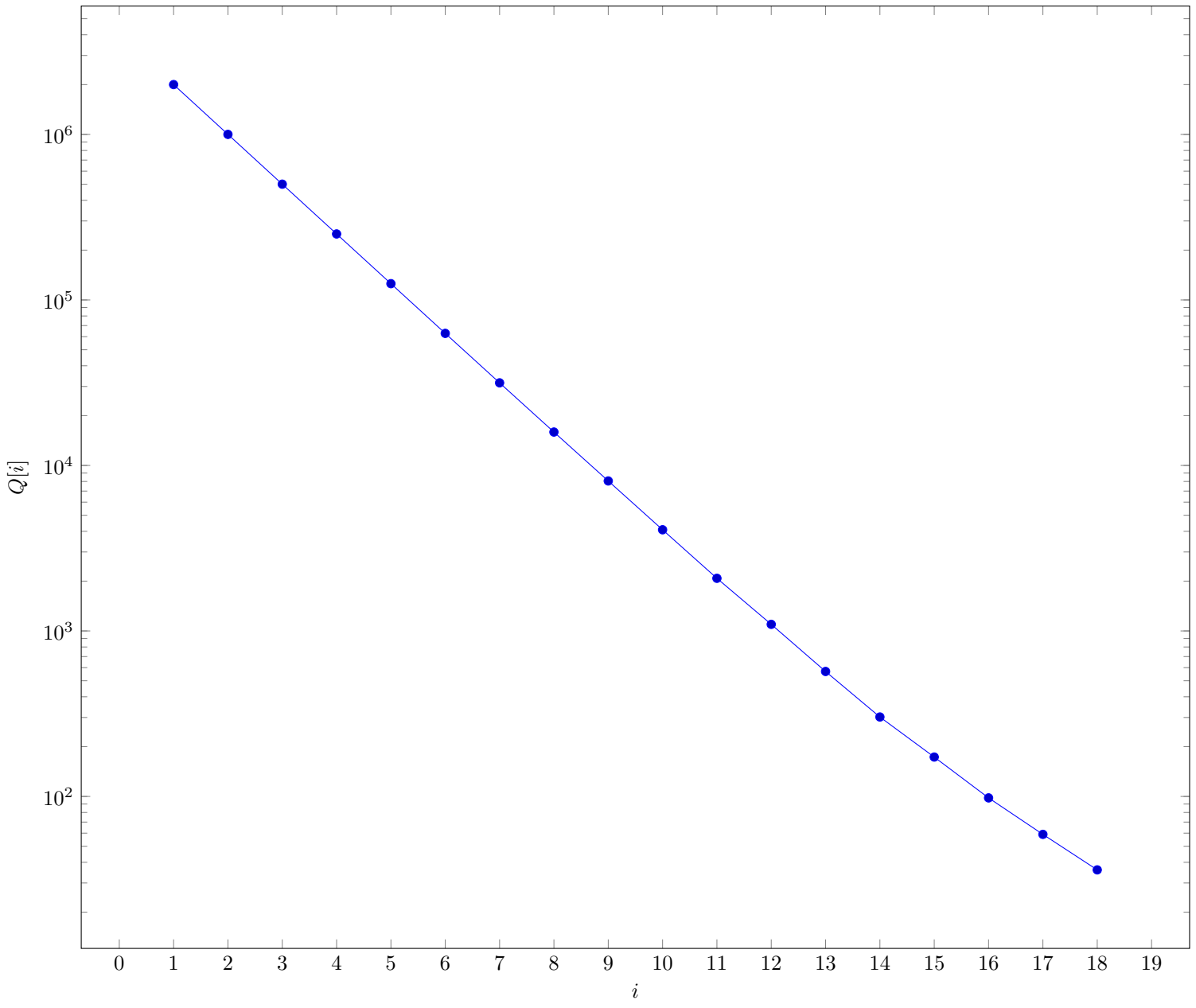

Fig. 18 Intermediate value $Q[i]$ in §6.3.5 of NIST SP 800-90B

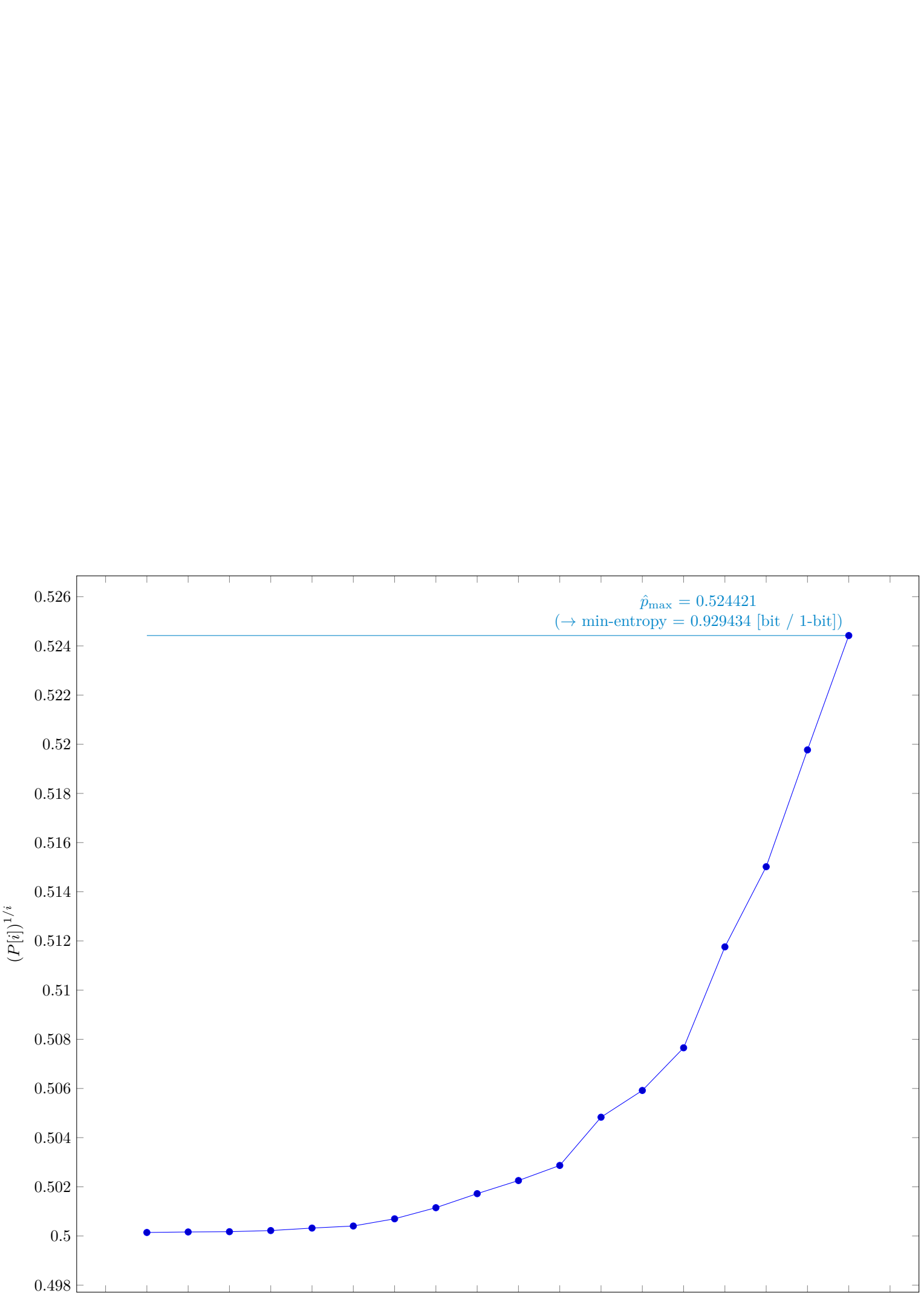

Fig. 19 $P[i]^{1/i}$ in §6.3.5 of NIST SP 800-90B

4.5.1 Supplemental information for traceability

Table 15 Supplemental information for traceability (NIST SP 800-90B Section 6.3.5)

Symbol	Value
t	18
$\hat{p}_{\max}$	0.524421
p_u	0.525064

4.6 The LRS Estimate (NIST SP 800-90B Section 6.3.6)

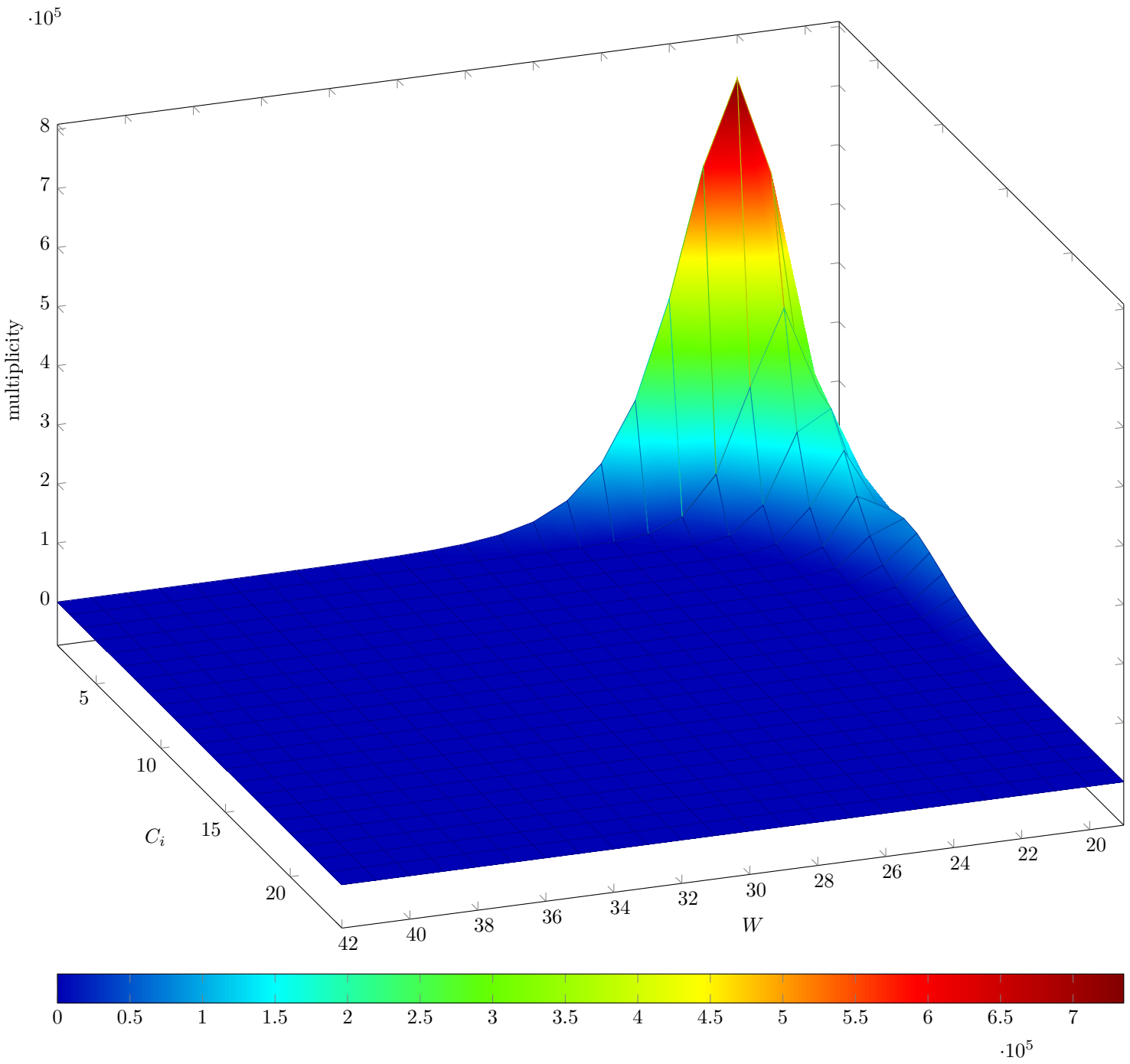

Fig. 20 Estimated W-tuple collision probability in Step 3 of §6.3.6 of NIST SP 800-90B

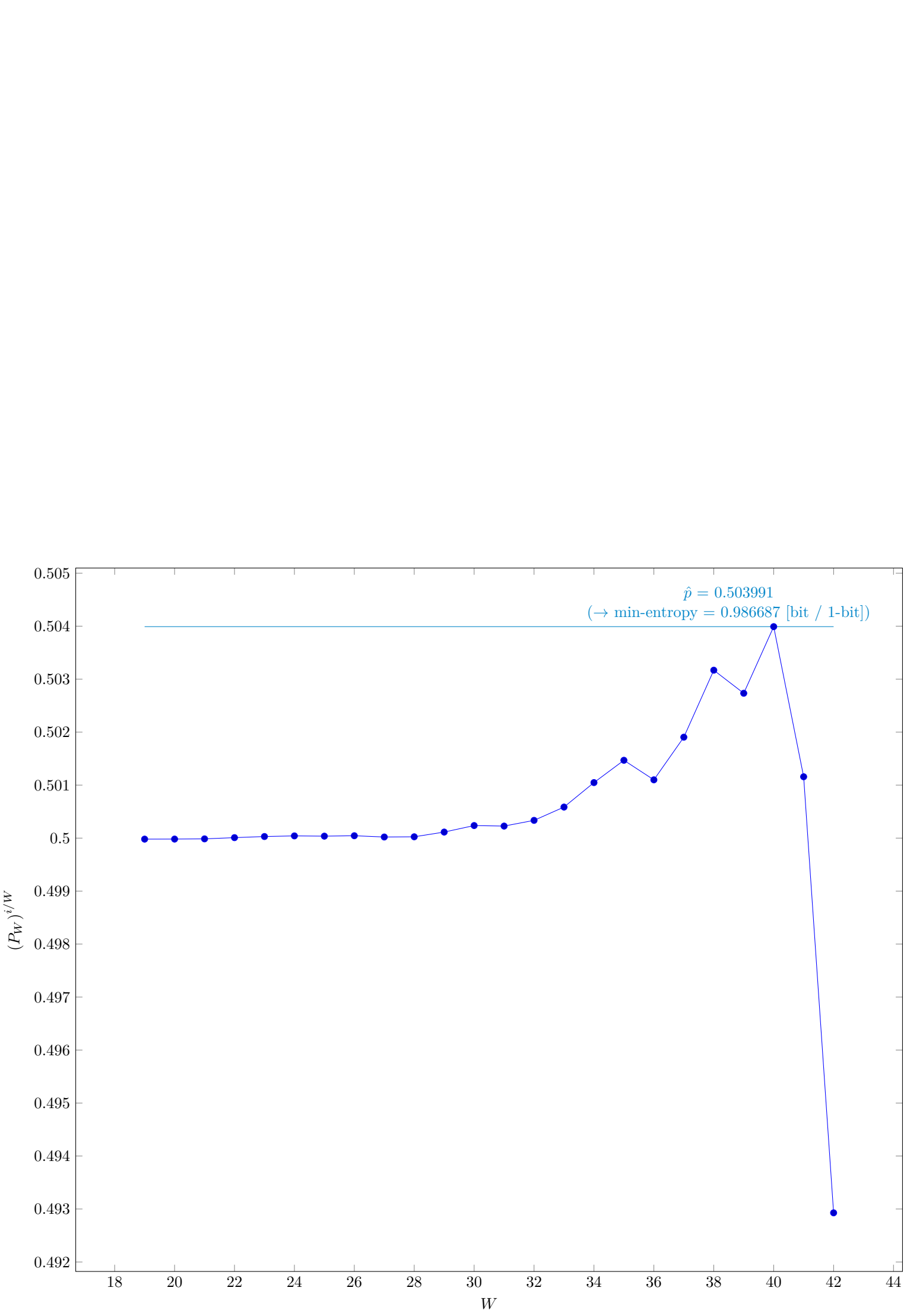

Fig. 21 Estimated average collision probability per string symbol in Step 3 of §6.3.6 of NIST SP 800-90B

4.6.1 Supplemental information for traceability

Table 16 Supplemental information for traceability (NIST SP 800-90B Section 6.3.6)

Symbol	Value
u	19
v	42
$\hat{p}$	0.503991
p_u	0.504635

4.7 Multi Most Common in Window Prediction Estimate (NIST SP 800-90B Section 6.3.7)

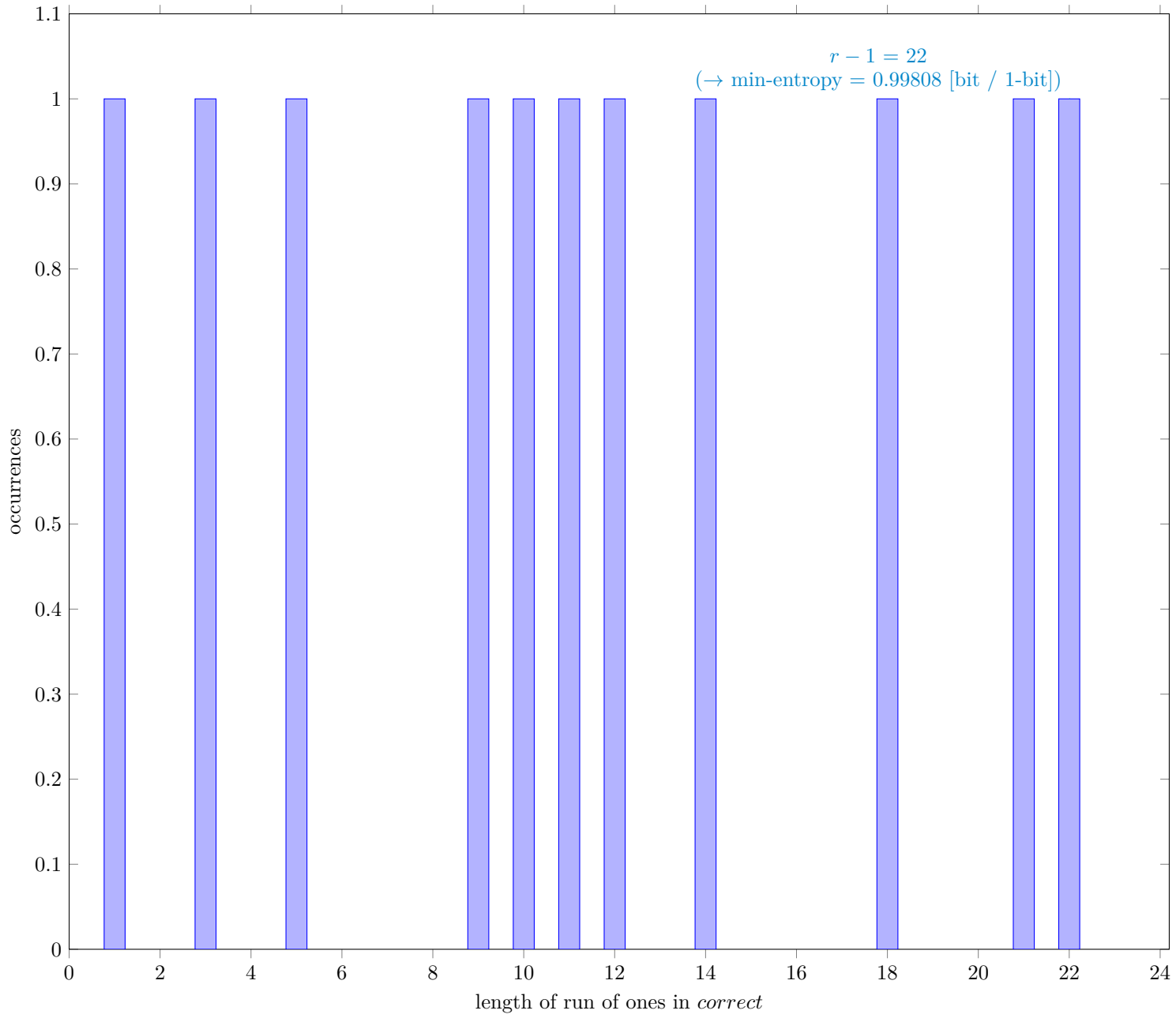

Fig. 22 Distribution of *correct*

4.7.1 Supplemental information for traceability

Table 17 Supplemental information for traceability (NIST SP 800-90B Section 6.3.7)

Symbol	Value
N	3999937
C	2000056
P_{global}	0.500022
P'_{global}	0.500666
r	23
P_{local}	0.433329

4.8 Lag Prediction Estimate (NIST SP 800-90B Section 6.3.8)

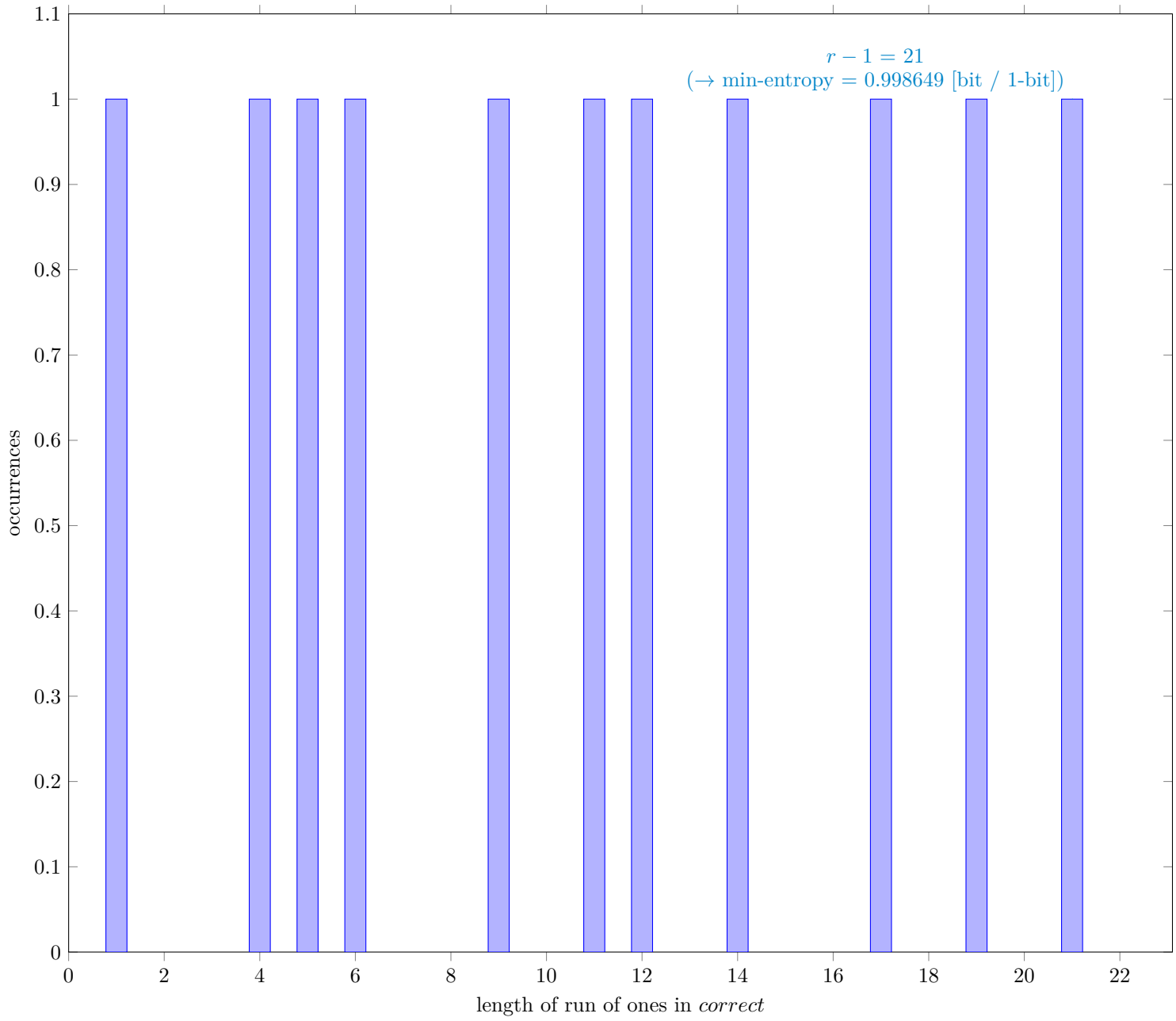

Fig. 23 Distribution of *correct*

4.8.1 Supplemental information for traceability

Table 18 Supplemental information for traceability (NIST SP 800-90B Section 6.3.8)

Symbol	Value
N	3999999
C	1999298
P_{global}	0.499825
P'_{global}	0.500469
r	22
P_{local}	0.416615

4.9 The MultiMMC Prediction Estimate (NIST SP 800-90B Section 6.3.9)

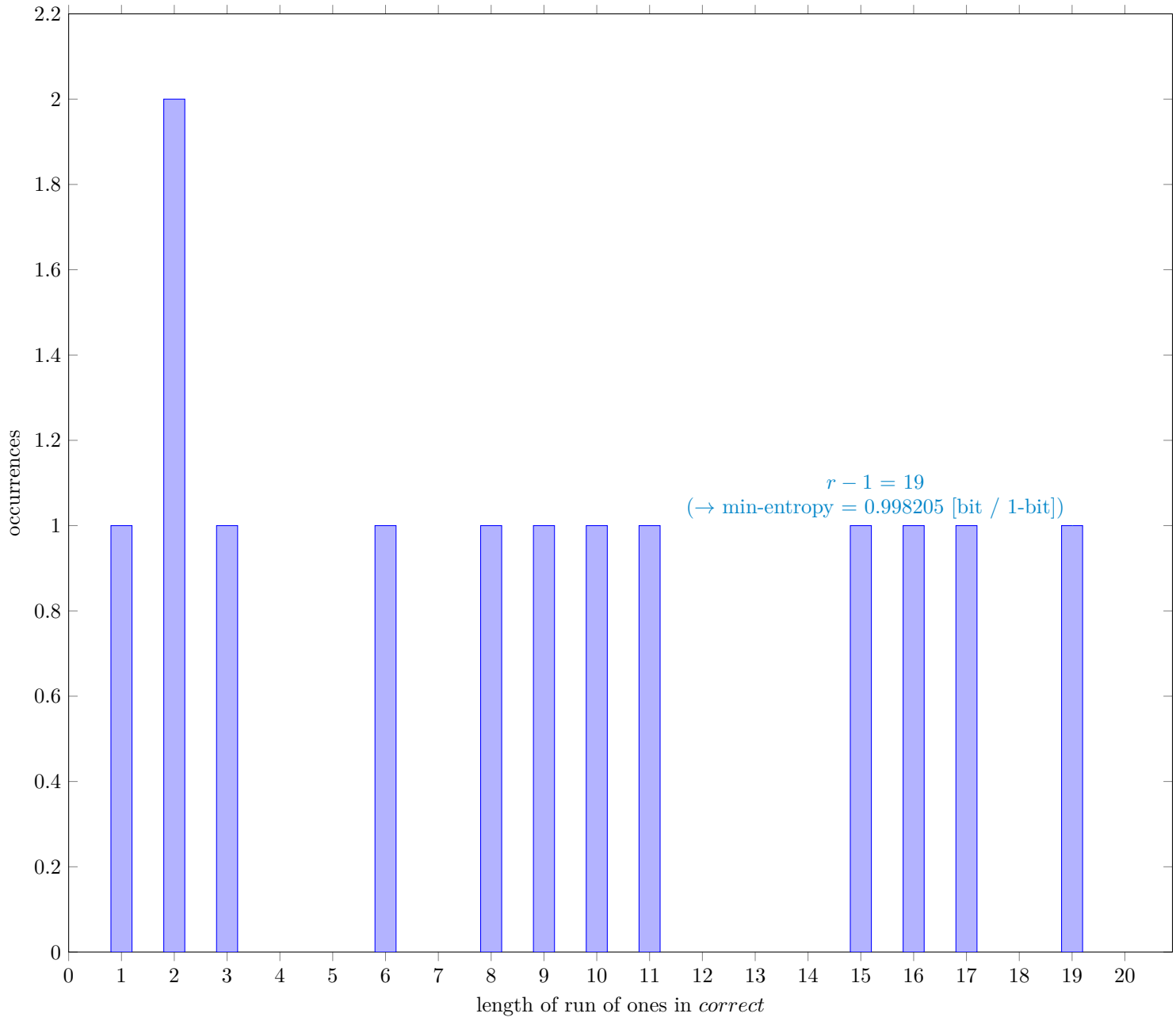

Fig. 24 Distribution of *correct*

4.9.1 Supplemental information for traceability

Table 19 Supplemental information for traceability (NIST SP 800-90B Section 6.3.9)

Symbol	Value
N	3999998
C	1999913
P_{global}	0.499978
P'_{global}	0.500622
r	20
P_{local}	0.380545

4.10 The LZ78Y Prediction Estimate (NIST SP 800-90B Section 6.3.10)

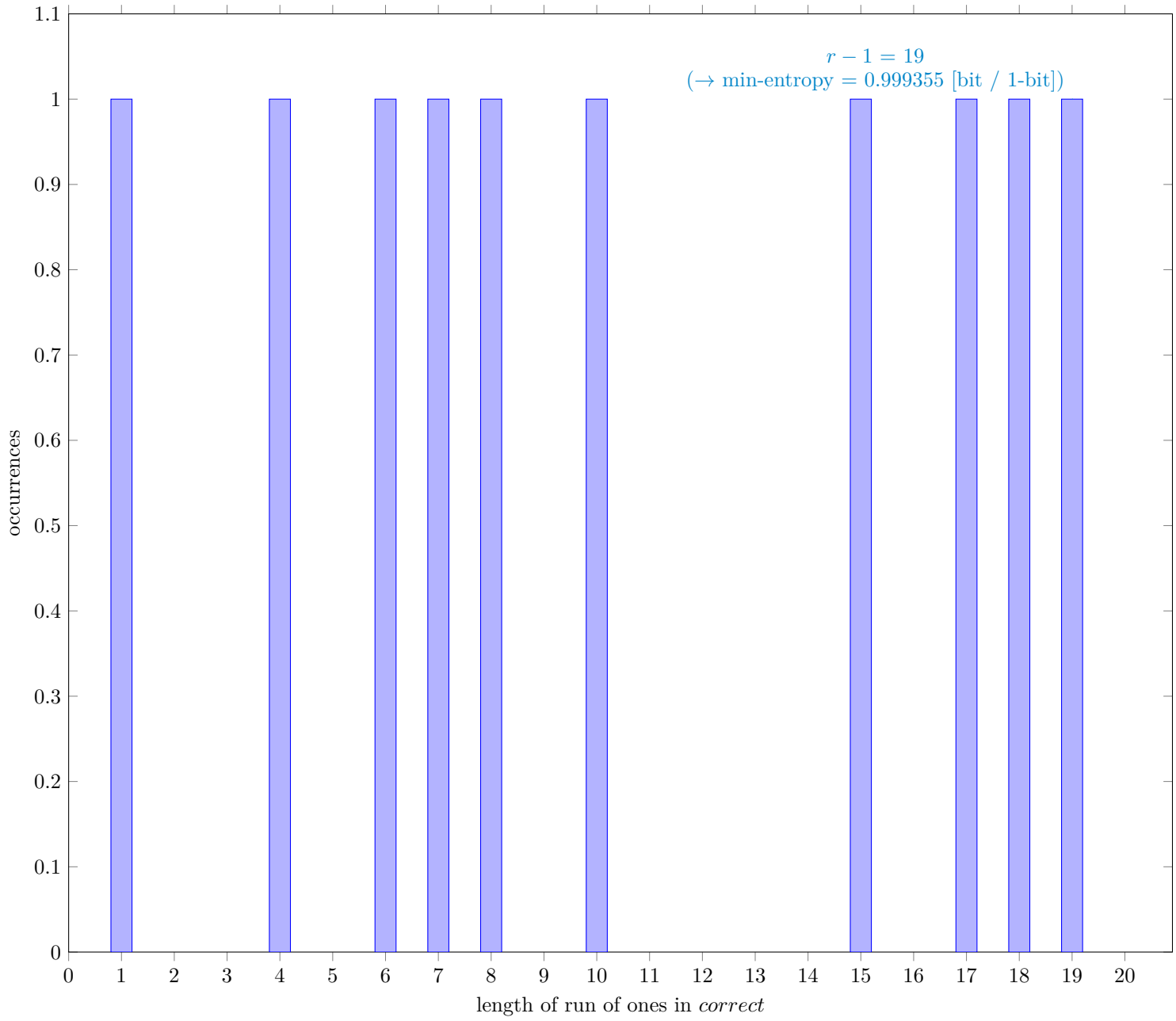

Fig. 25 Distribution of *correct*

4.10.1 Supplemental information for traceability

Table 20 Supplemental information for traceability (NIST SP 800-90B Section 6.3.10)

Symbol	Value
N	3999983
C	1998310
P_{global}	0.49958
P'_{global}	0.500224
r	20
P_{local}	0.380545

4 References

[1] Meltem Sönmez Turan, Elaine Barker, John Kelsey, Kerry A. McKay, Mary L. Baish, Mike Boyle, *Recommendation for the Entropy Sources Used for Random Bit Generation*, NIST Special Publication 800-90B, Jan. 2018 `https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-90B.pdf`

[2] G. Sakurai, *Proposed list of corrections for NIST SP 800-90B 6.3 Estimators*, Dec. 2022 `https://github.com/g-g-sakura/AnotherEntropyEstimationTool/blob/main/documentation/ProposedListOfCorrections_SP800-90B.pdf`