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

2023-Feb-26 14:08:37.017921

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/biased-random-bytes.bin`

- 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	1.0.43
Analysis environment	Hostname	[redacted]
	CPU information	AMD Ryzen [redacted]
	Physical memory size	[redacted] MB
	OS information	Windows 10 or greater
	Username	[redacted]

1.3 Identification of analysis conditions

Table 3 Identification information of analysis conditions

Number of samples	1000000
Bits per sample	8
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 / 8 - bit]	$H_{\text{bitstring}}$ [b] [bit / 1 - bit]
The Most Common Value Estimate	0.319651	0.151827
The Collision Estimate	—	0.0727058
The Markov Estimate	—	0.0916044
The Compression Estimate	—	0.0631355
The t-Tuple Estimate	0.29116	0.0322176
The Longest Repeated Substring (LRS) Estimate	0.519281	0.0648017
Multi Most Common in Window Prediction Estimate	0.319646	0.0419265
The Lag Prediction Estimate	0.466258	0.0420028
The MultiMMC Prediction Estimate	0.320277	0.0419265
The LZ78Y Prediction Estimate	0.330375	0.0419265
The intial entropy source estimate [bit / 8 - bit] $H_I = \min(H_{\text{original}}, 8 \times H_{\text{bitstring}})$		0.257741

[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

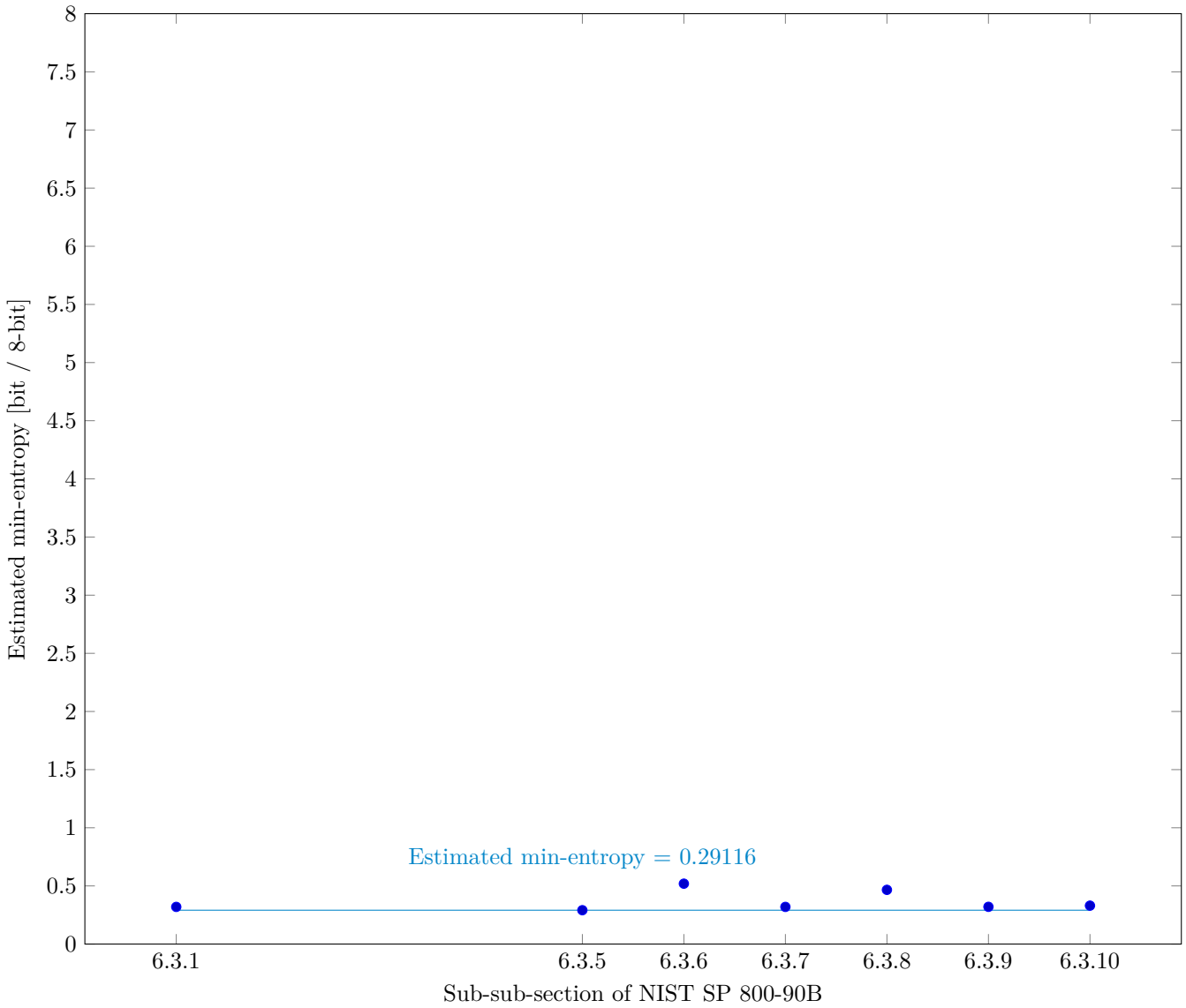


Fig. 1 Estimated Min-Entropy using §6.3 of NIST SP 800-90B

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

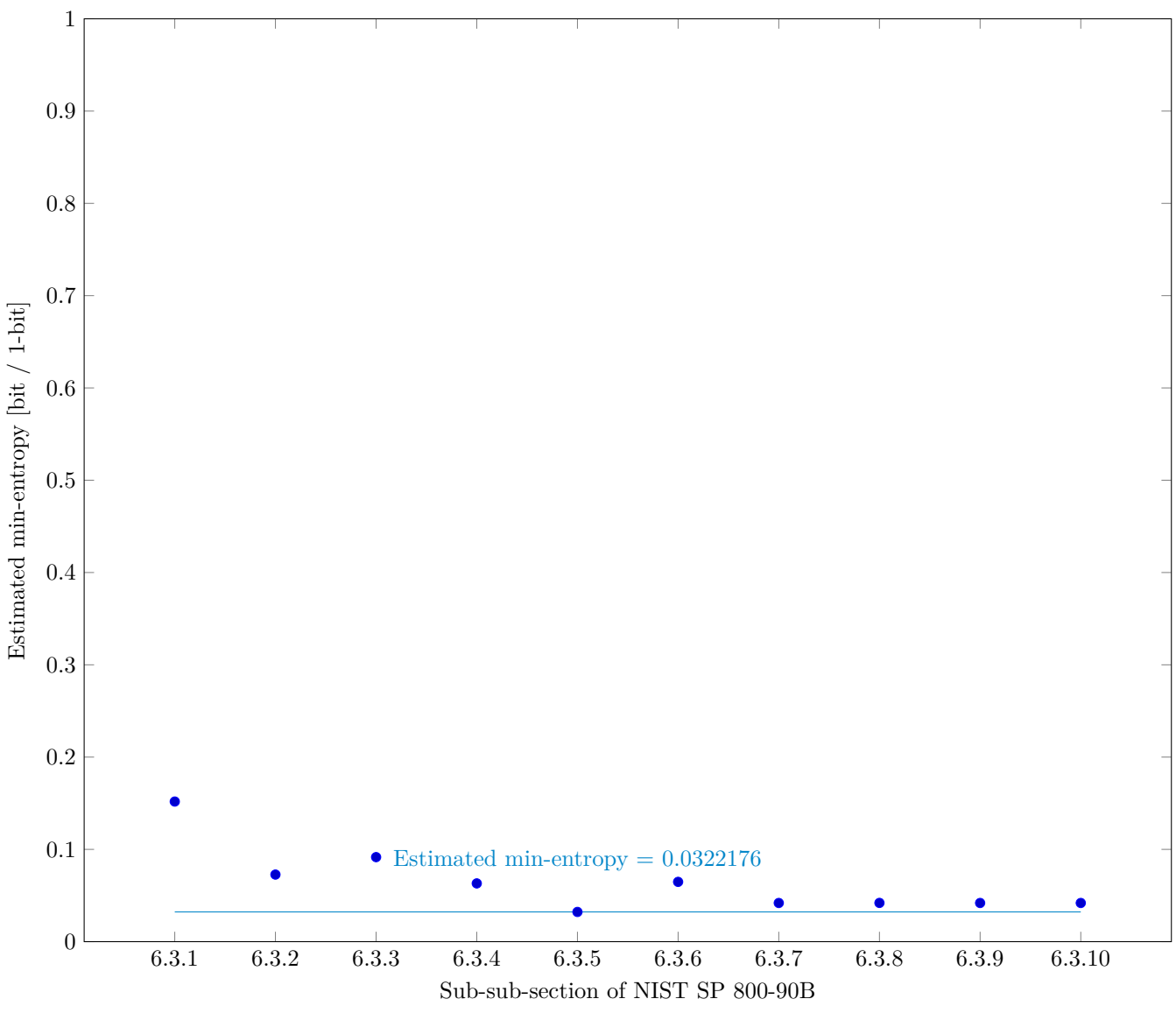
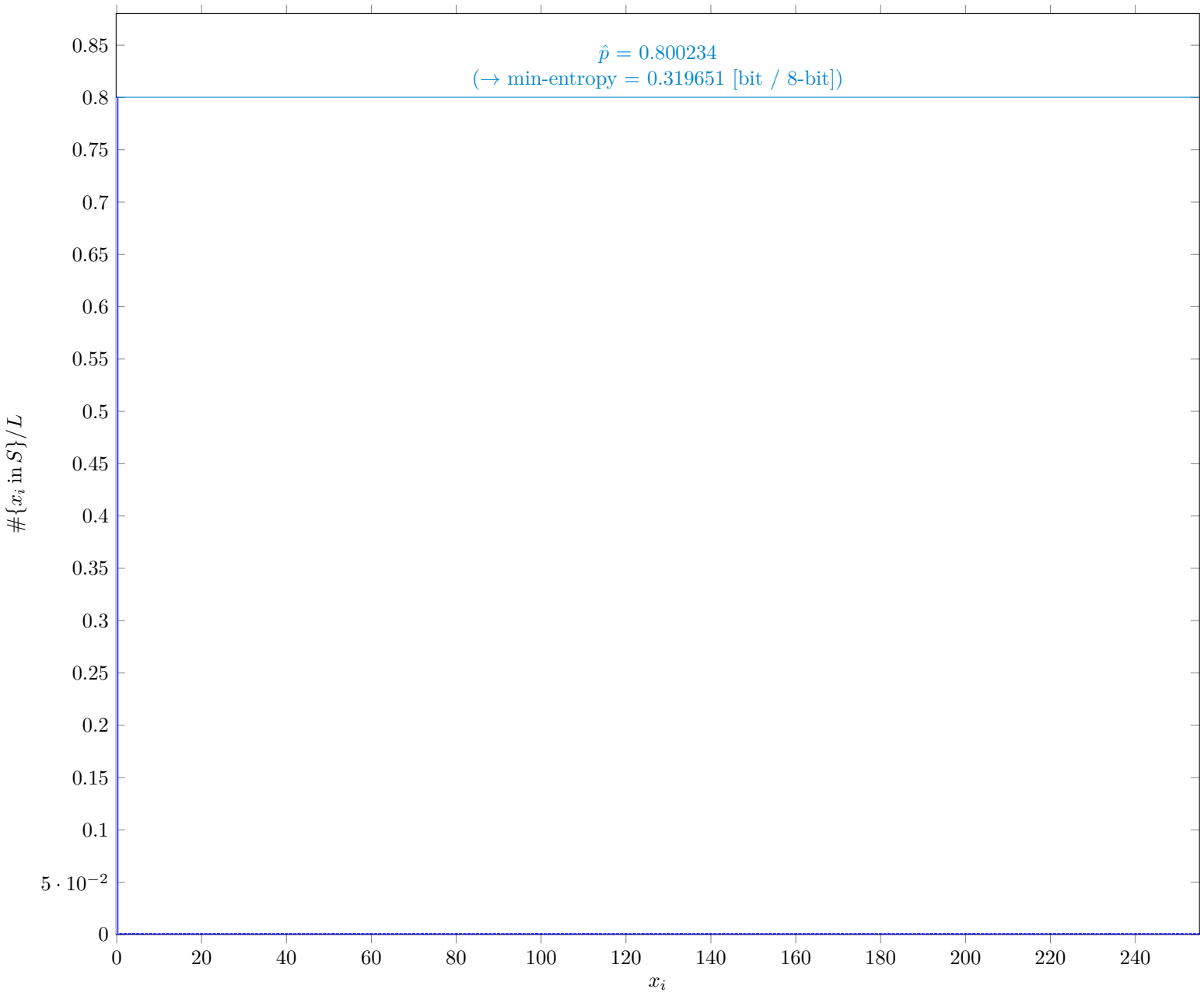


Fig. 2 Estimated Min-Entropy using §6.3 of NIST SP 800-90B

3 Detailed results of analysis from original samples

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



3.1.1 Supplemental information for traceability

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

Symbol	Value
mode	800234
$\hat{p}$	0.800234
p_u	0.801264

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

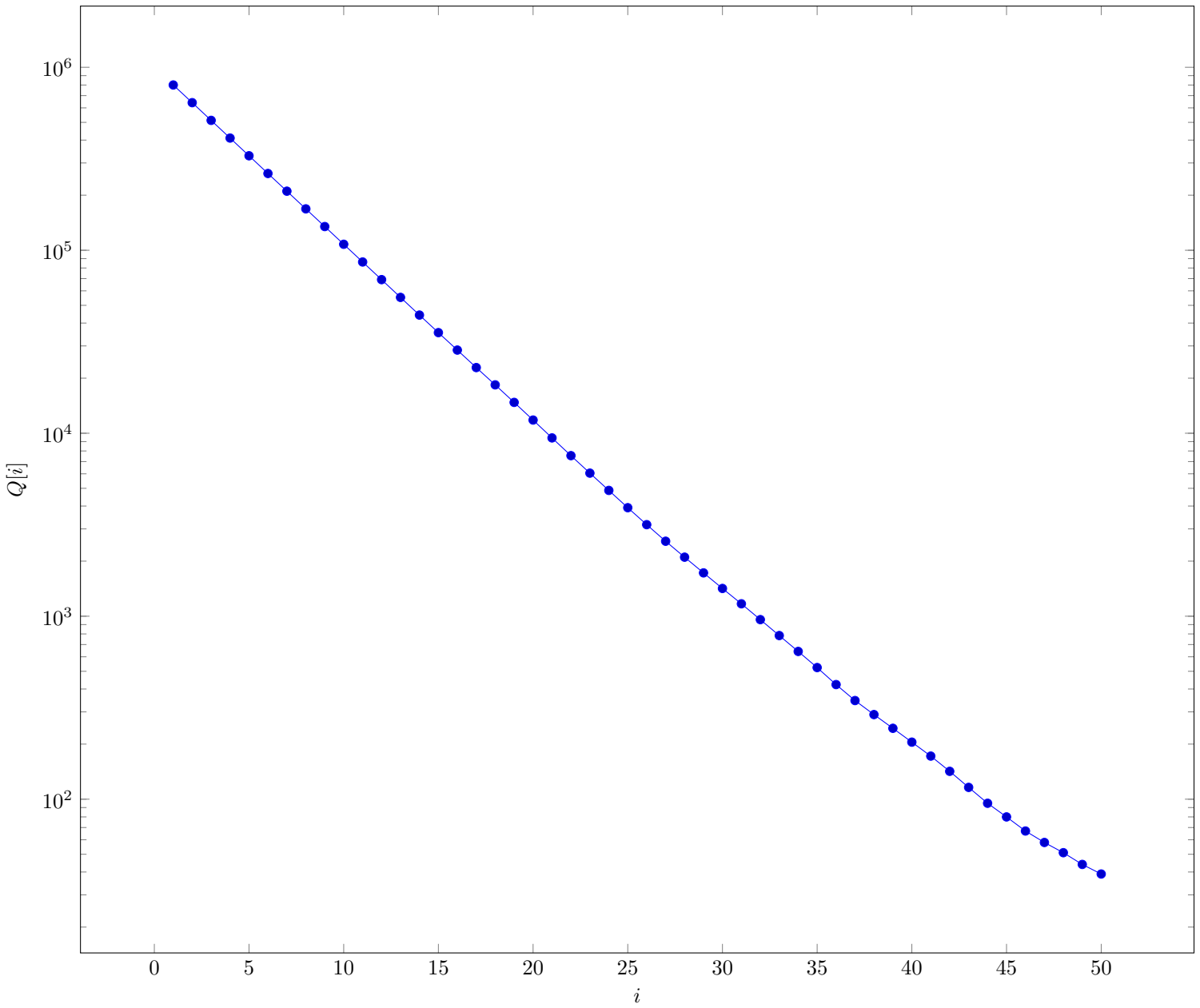


Fig. 3 Intermediate value $Q[i]$ in §6.3.5 of NIST SP 800-90B

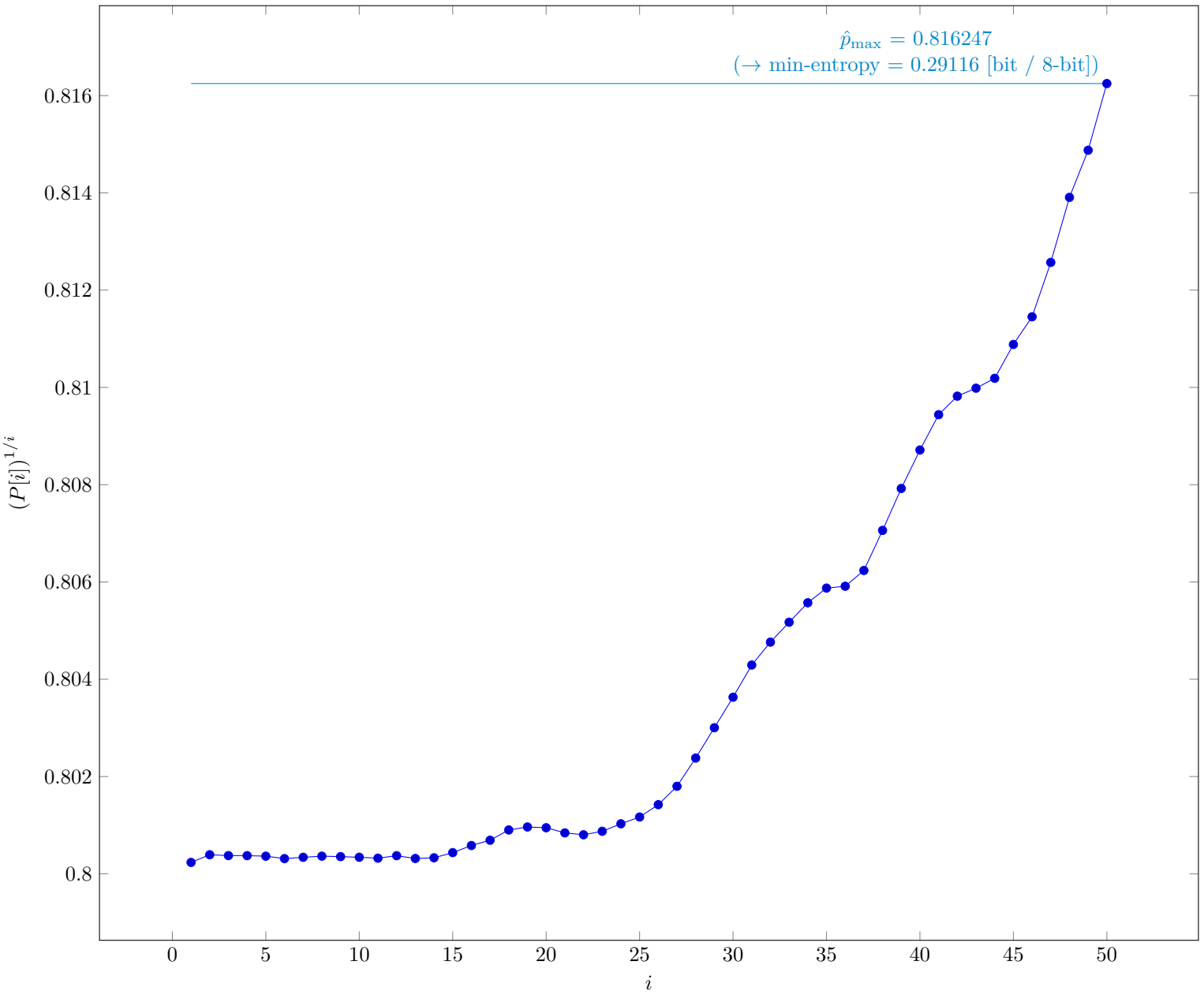


Fig. 4 $P[i]^{1/i}$ in §6.3.5 of NIST SP 800-90B

3.2.1 Supplemental information for traceability

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

Symbol	Value
t	50
$\hat{p}_{\max}$	0.816247
p_u	0.817245

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

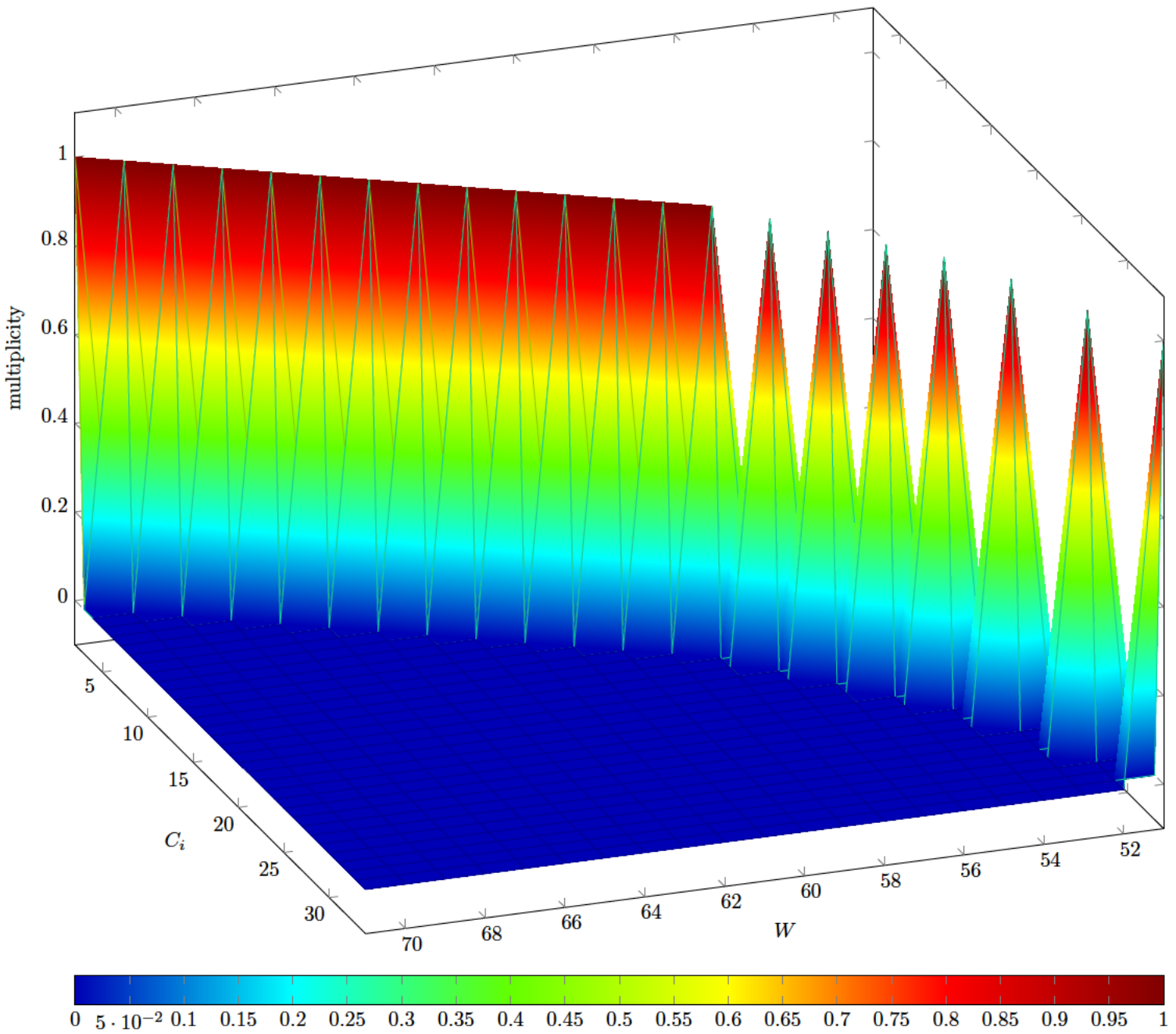


Fig. 5 Estimated W-tuple collision probability in Step 3 of §6.3.6 of NIST SP 800-90B

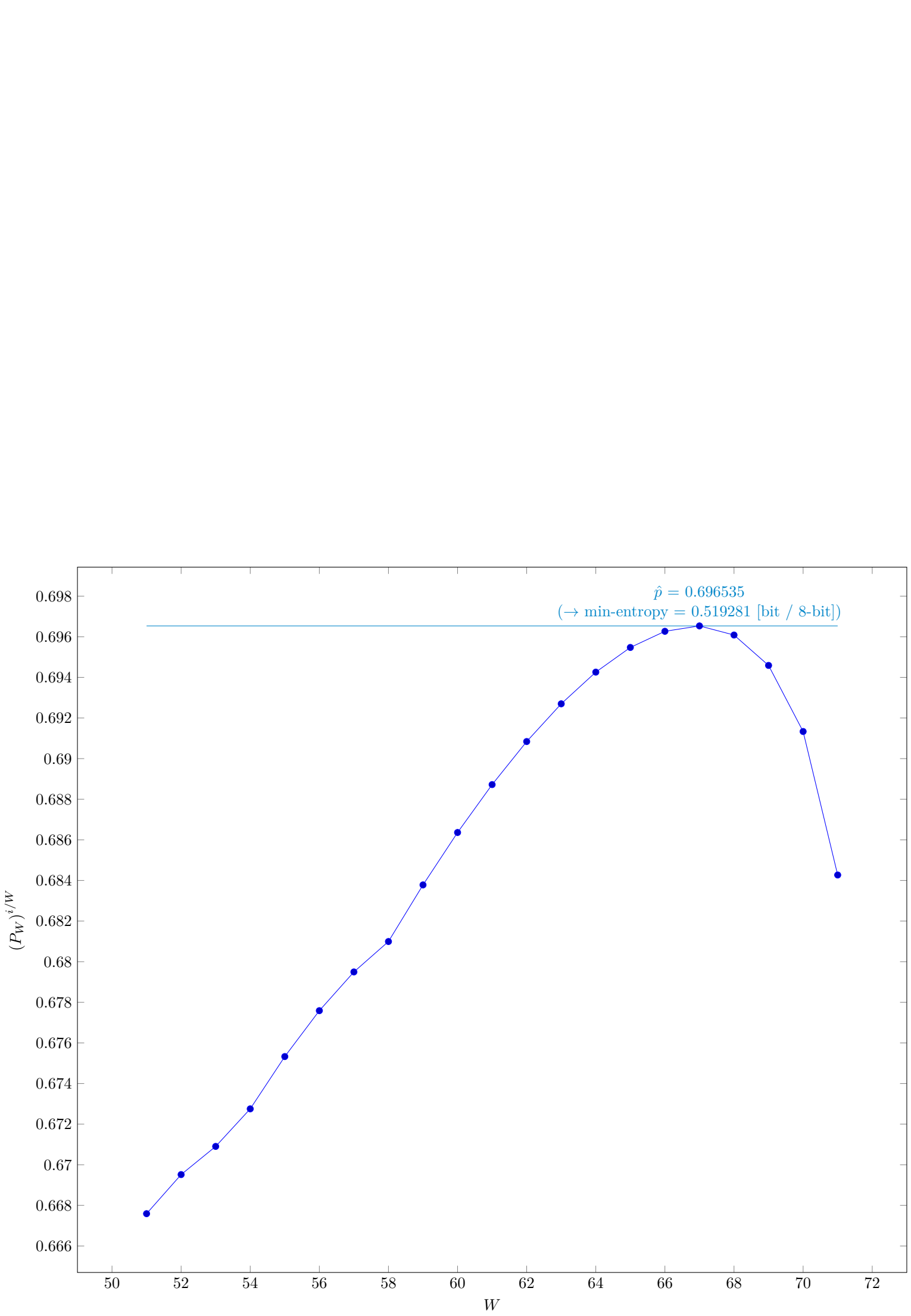


Fig. 6 Estimated average collision probability per string symbol in Step 3 of §6.3.6 of NIST SP 800-90B

3.3.1 Supplemental information for traceability

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

Symbol	Value
u	51
v	71
$\hat{p}$	0.696535
p_u	0.697719

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

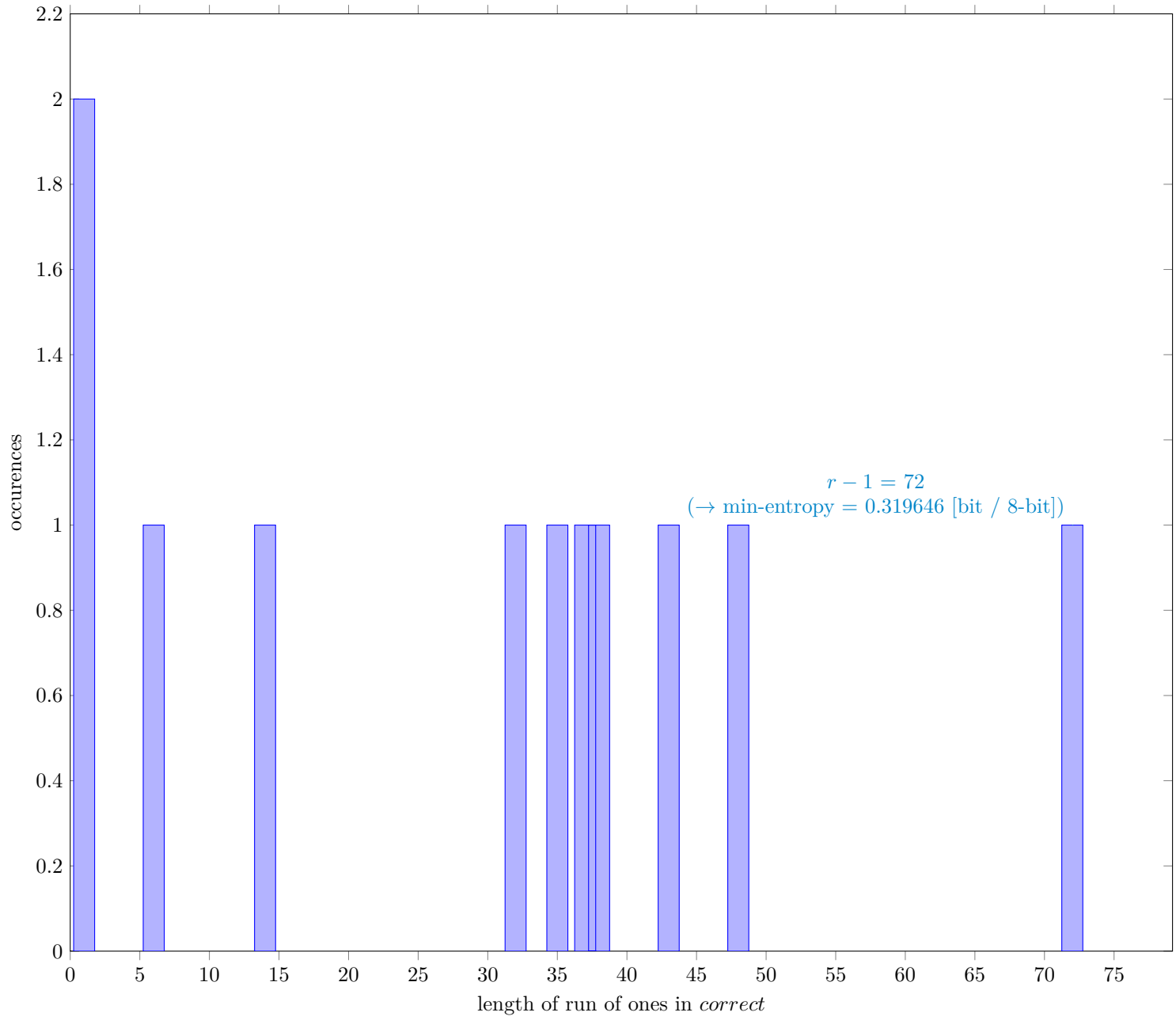


Fig. 7 Distribution of *correct*

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	800186
P_{global}	0.800236
P'_{global}	0.801266
r	73
P_{local}	0.794039

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

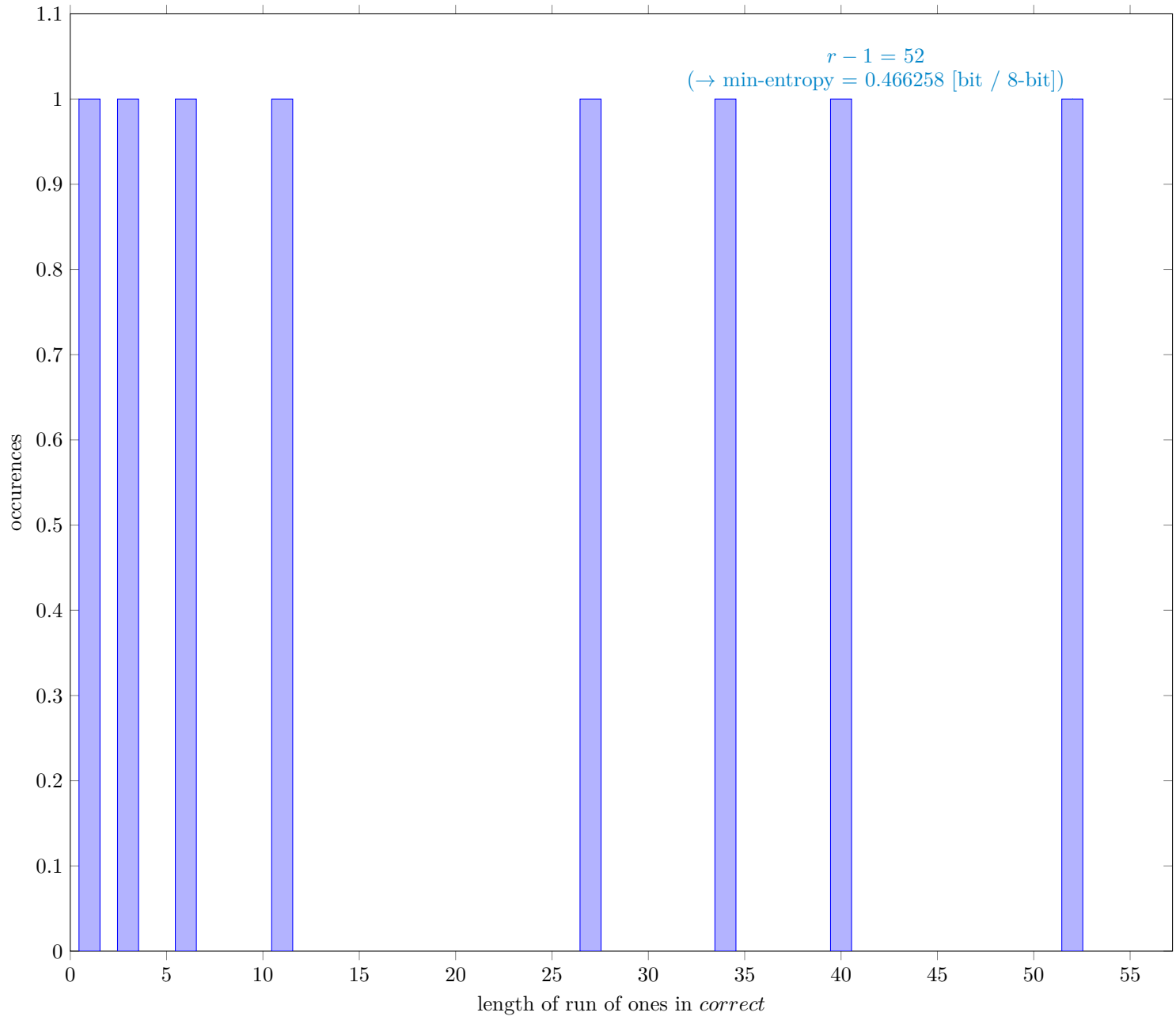


Fig. 8 Distribution of *correct*

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	639881
P_{global}	0.639882
P'_{global}	0.641118
r	53
P_{local}	0.723839

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

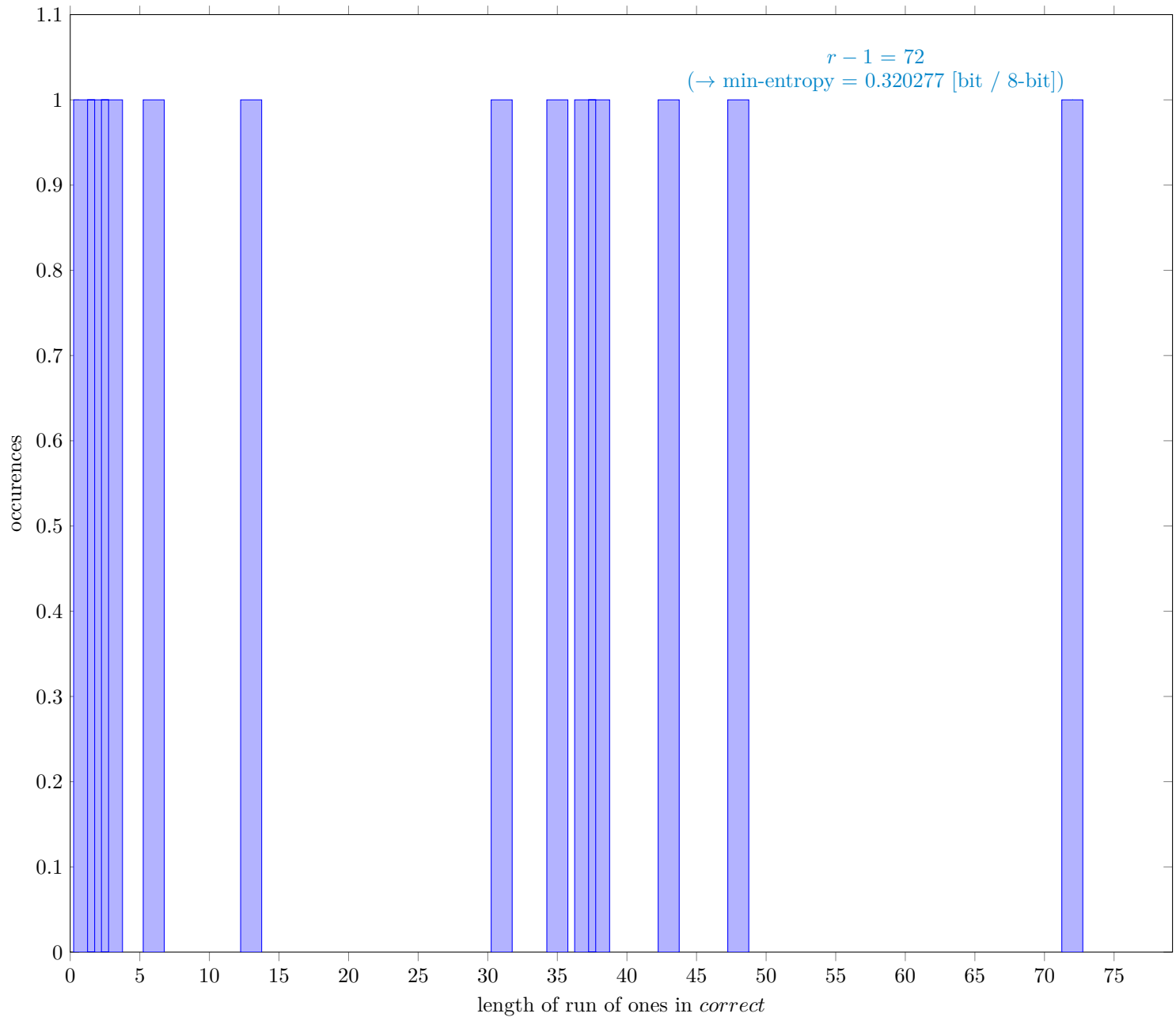


Fig. 9 Distribution of *correct*

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	799884
P_{global}	0.799886
P'_{global}	0.800916
r	73
P_{local}	0.794038

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

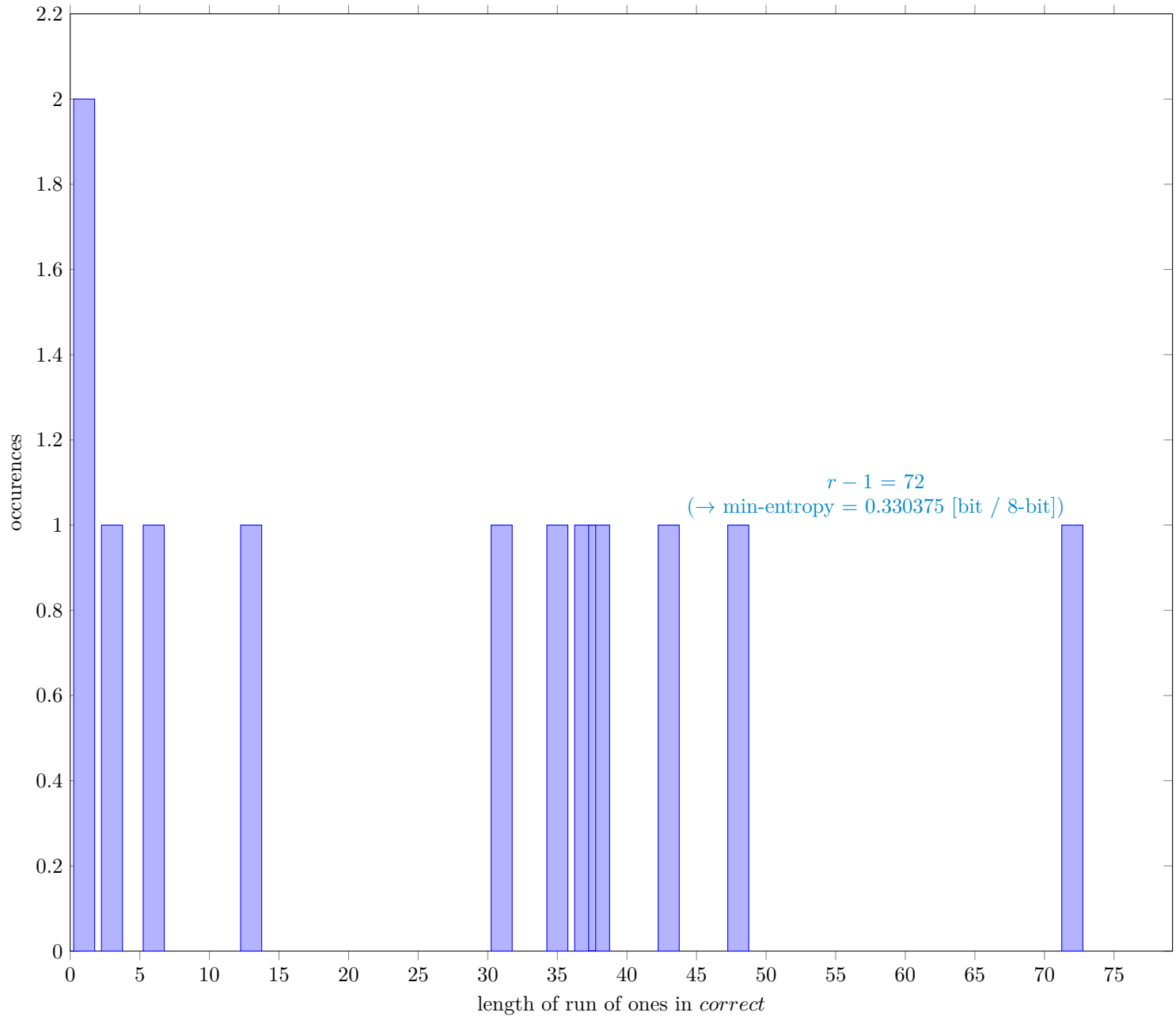


Fig. 10 Distribution of *correct*

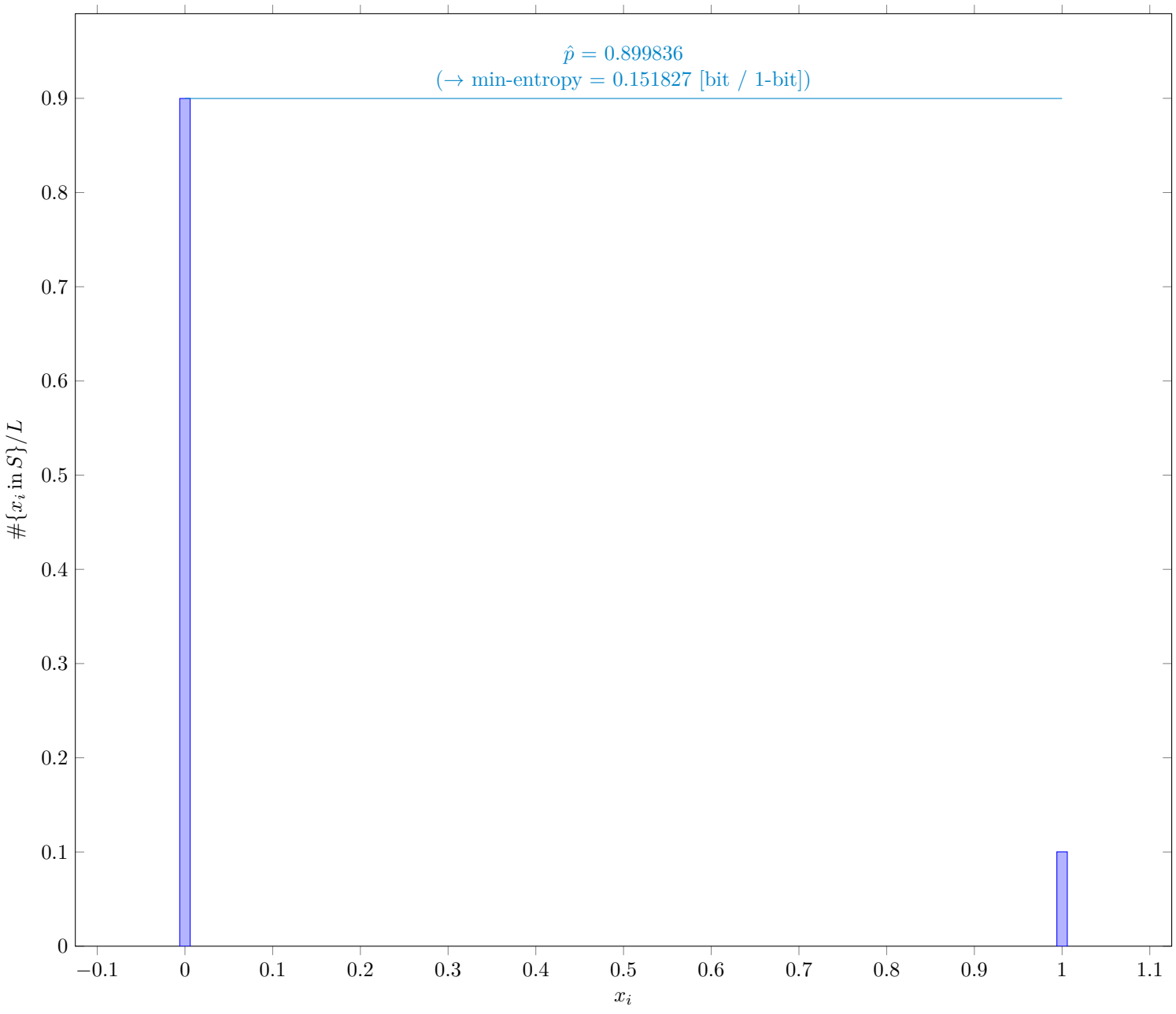
3.7.1 Supplemental information for traceability

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

Symbol	Value
N	999983
C	794275
P_{global}	0.794289
P'_{global}	0.79533
r	73
P_{local}	0.794038

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)

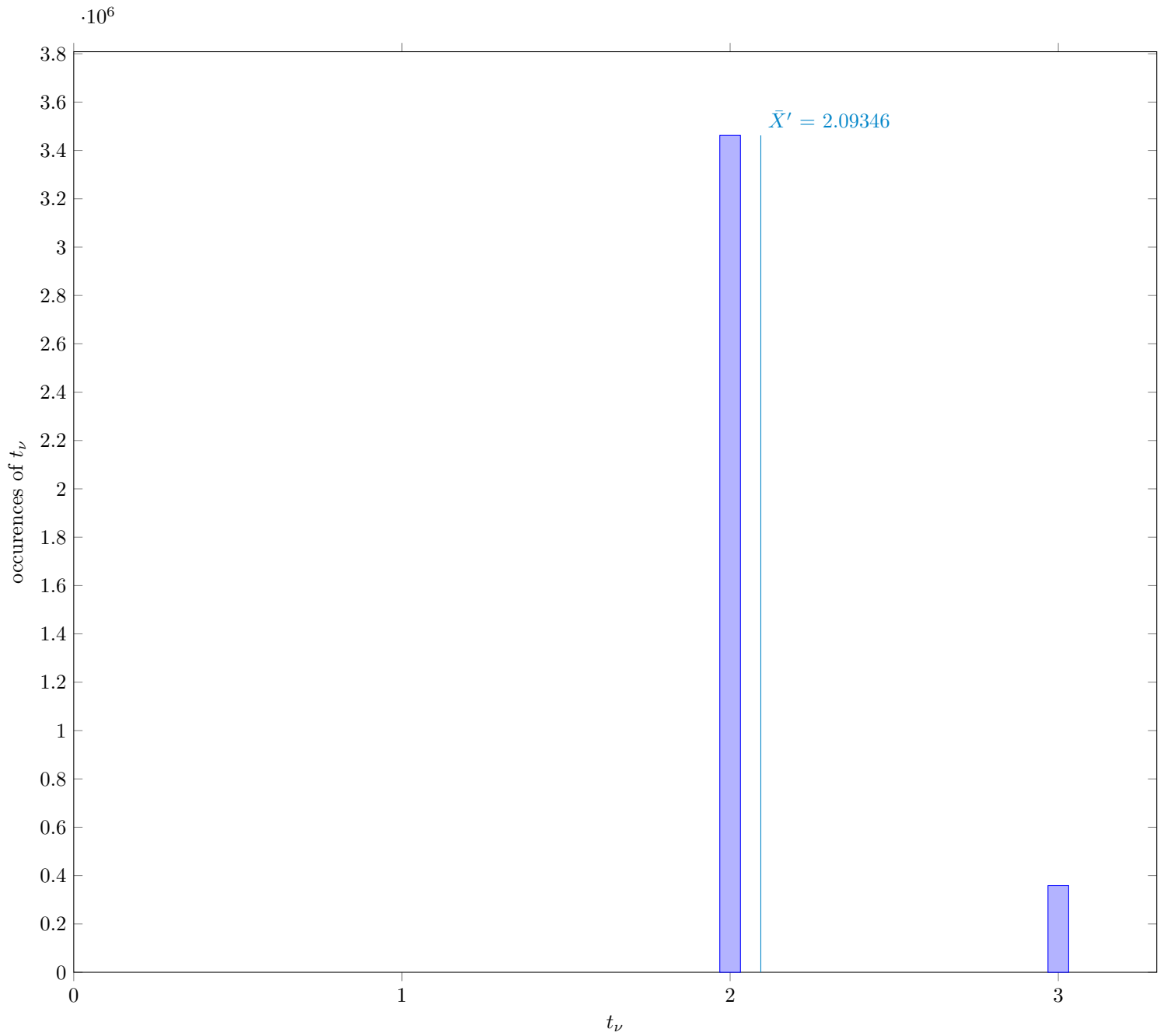


4.1.1 Supplemental information for traceability

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

Symbol	Value
mode	7198690
$\hat{p}$	0.899836
p_u	0.90011

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



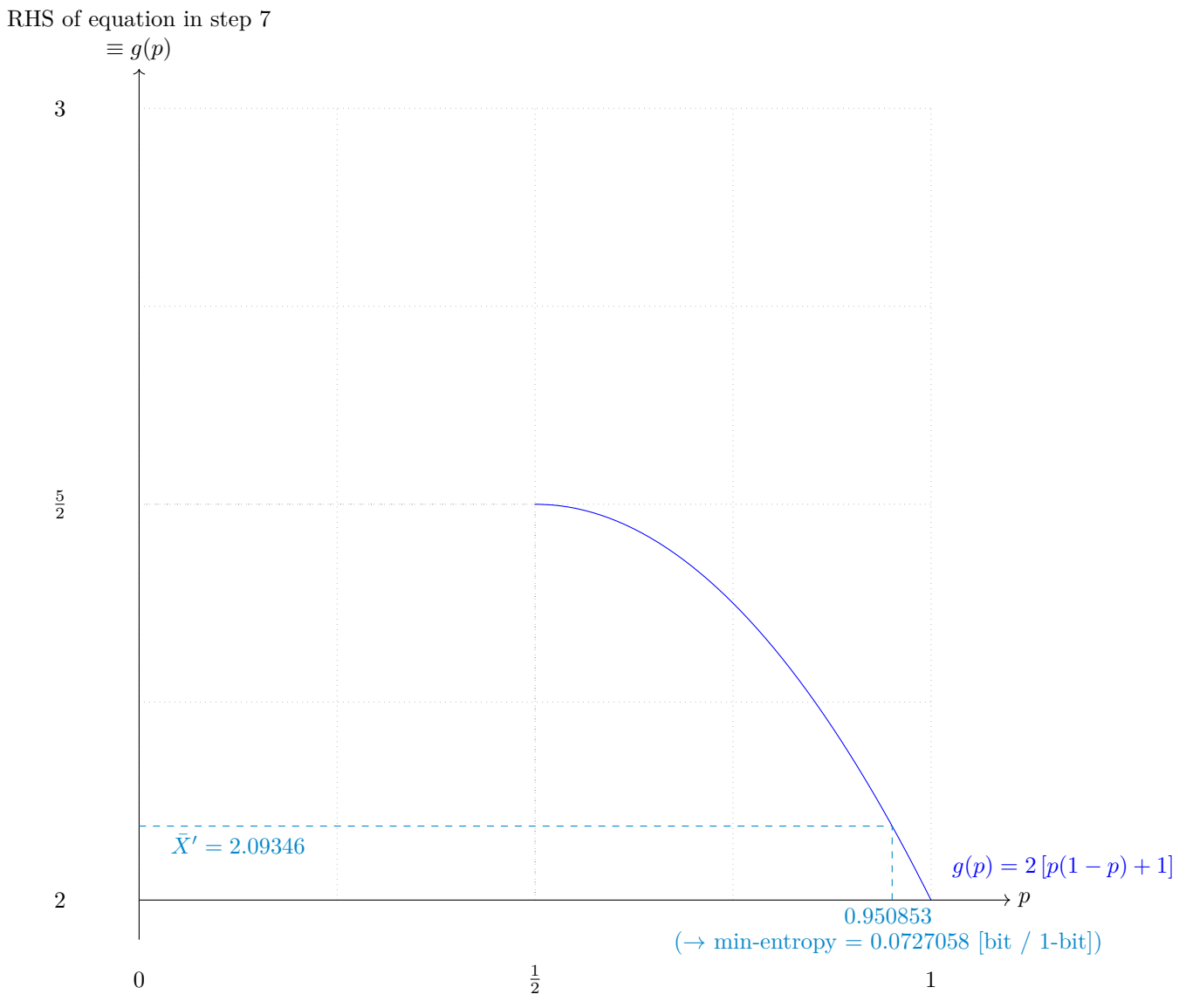


Fig. 11 Solution to the equation in step 7

4.2.1 Supplemental information for traceability

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

Symbol	Value
p	0.950853
$\bar{X}$	2.09385
$\bar{X}'$	2.09346
$\hat{\sigma}$	0.291617

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

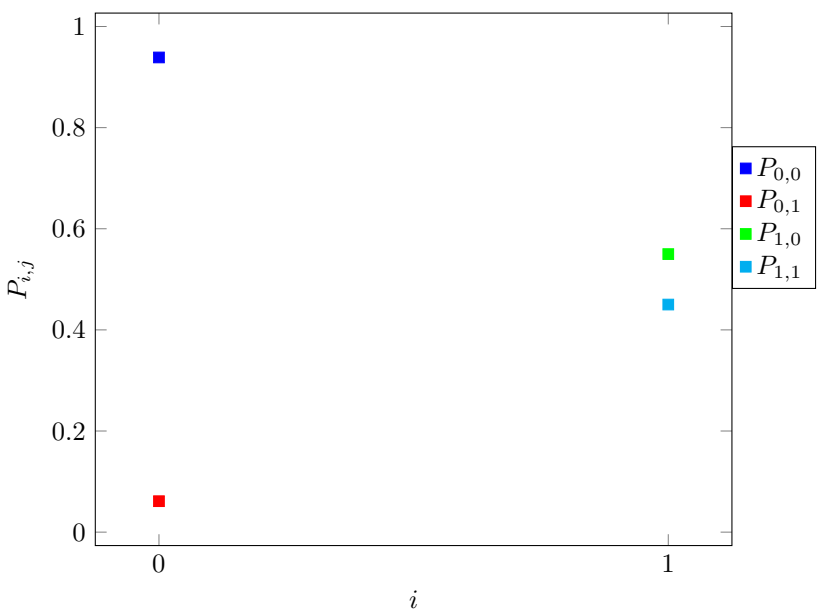


Fig. 12 Transition probability $P_{i,j}$ of §6.3.3 of NIST SP 800-90B

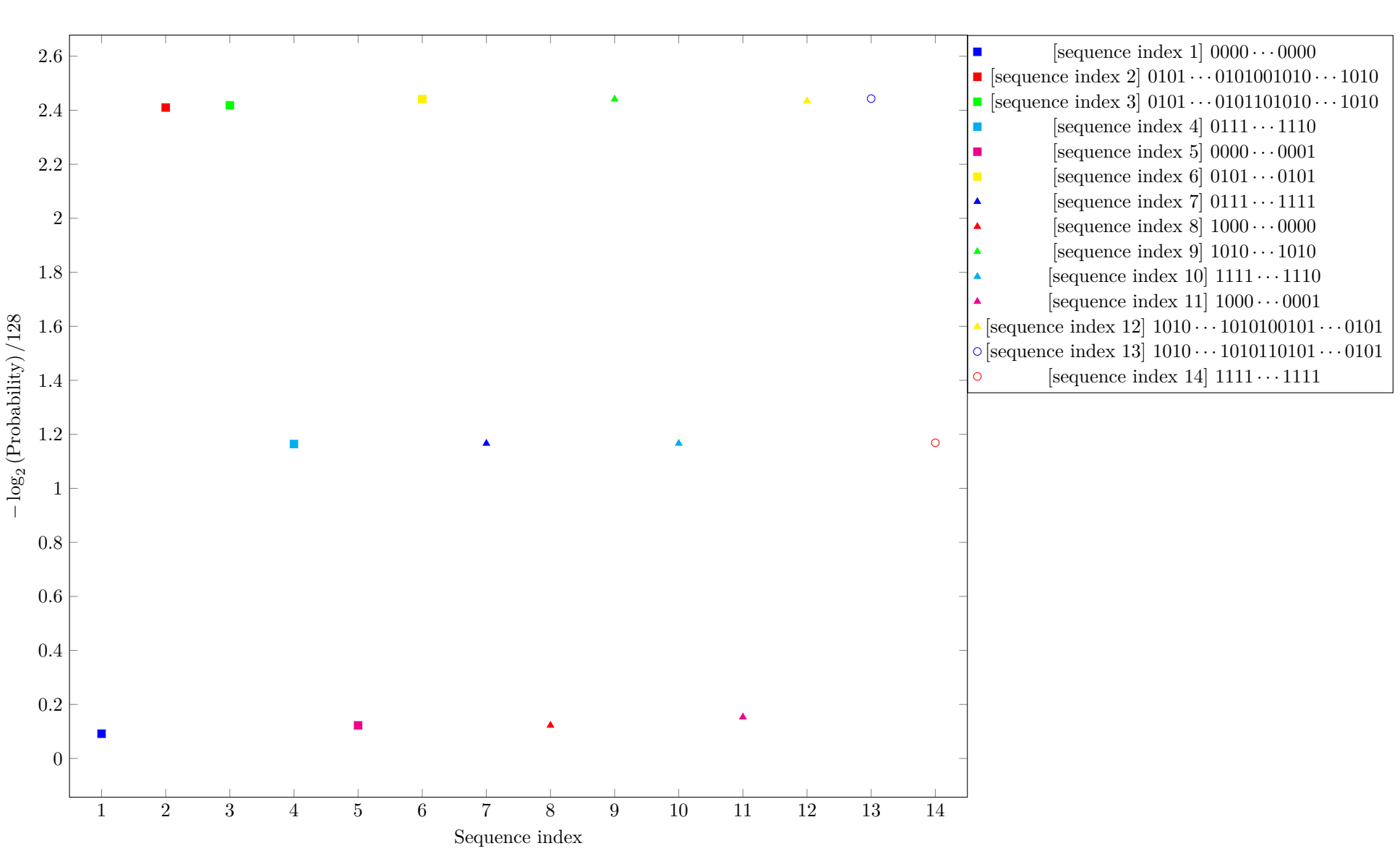
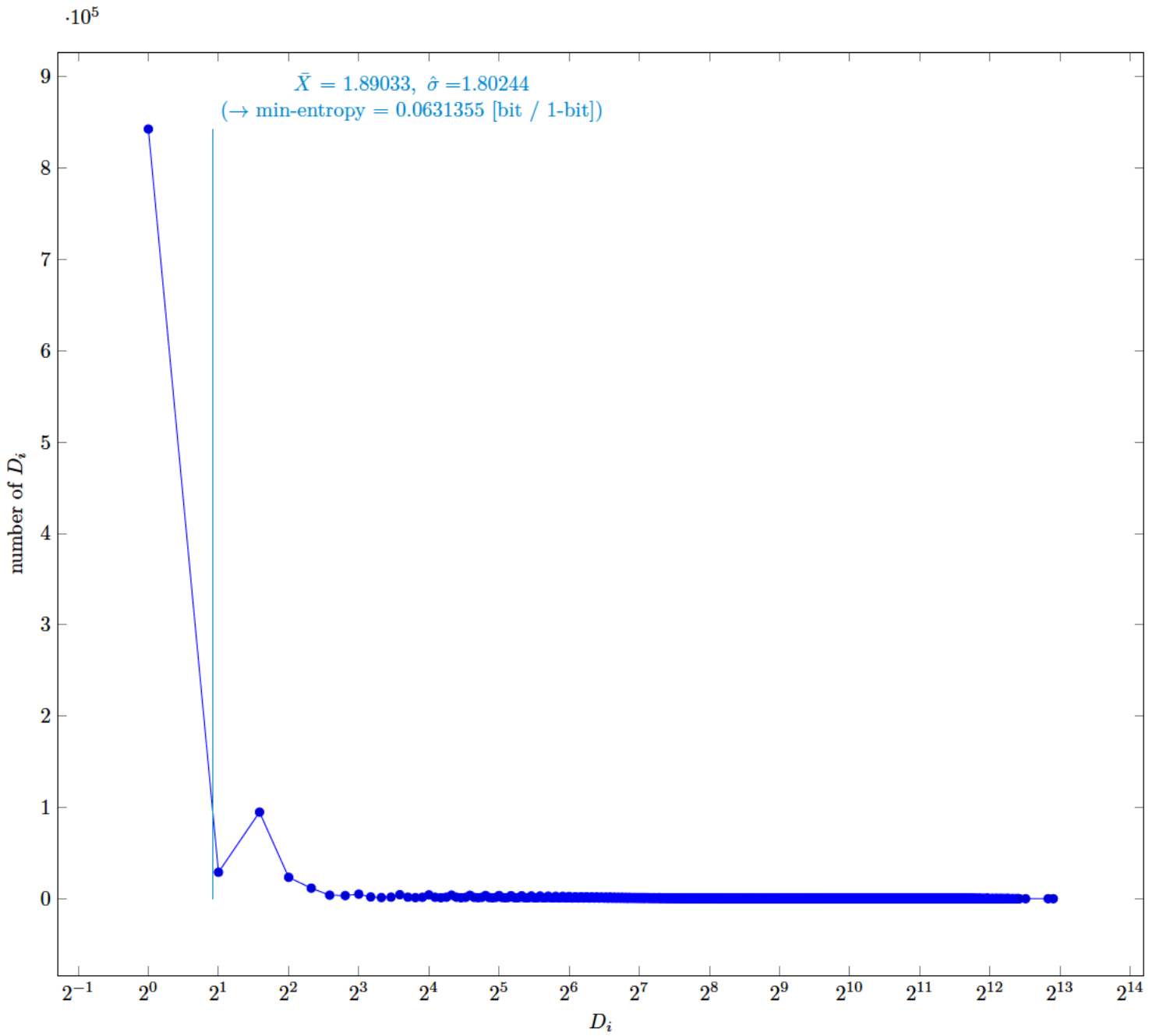


Fig. 13 Estimated Min-Entropy using §6.3.3 of NIST SP 800-90B

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



4.4.1 Supplemental information for traceability

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

Symbol	Value
p	0.76907
$\bar{X}$	1.89033
$\hat{\sigma}$	1.80244
$\bar{X}'$	1.88631

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

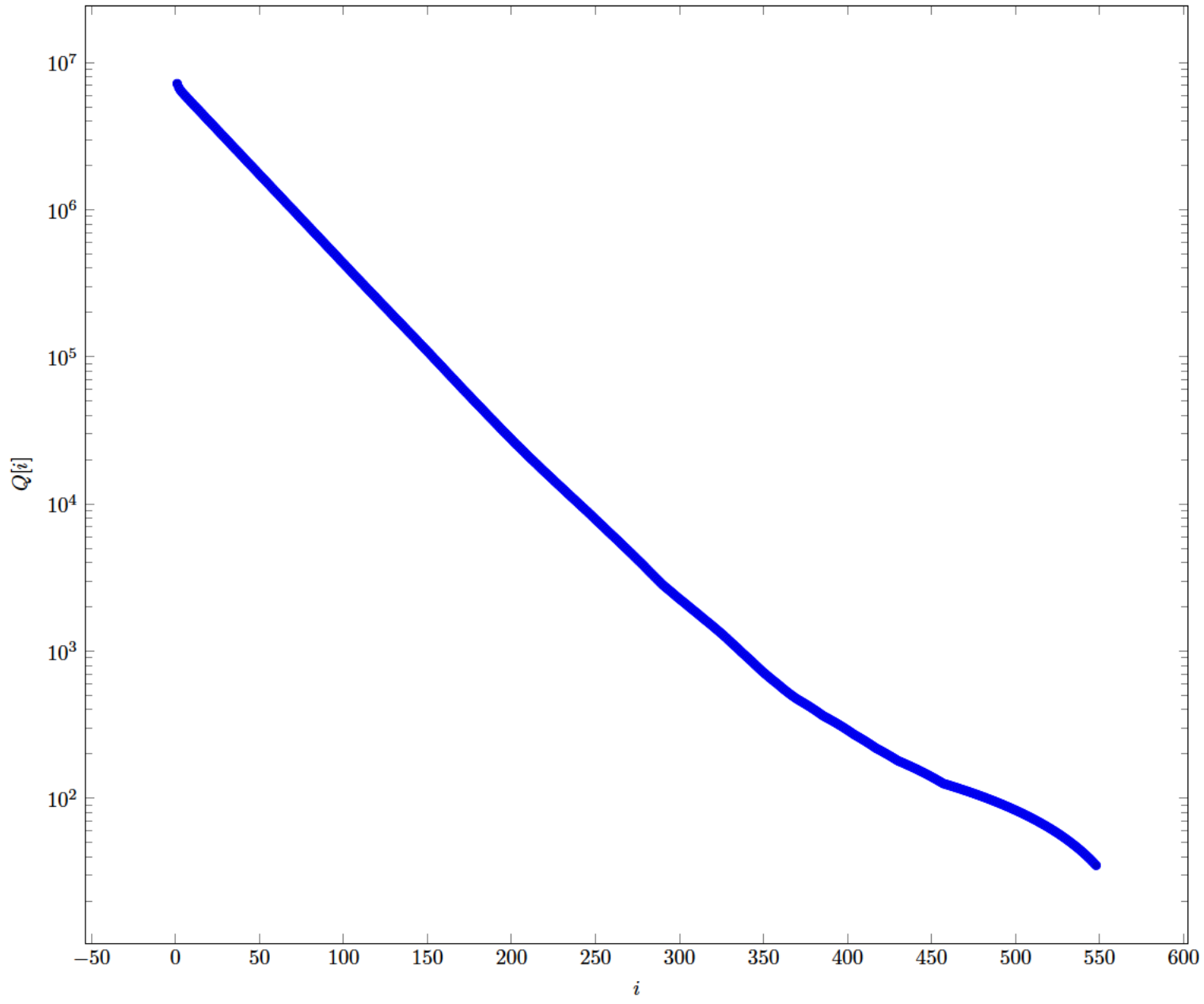


Fig. 14 Intermediate value $Q[i]$ in §6.3.5 of NIST SP 800-90B

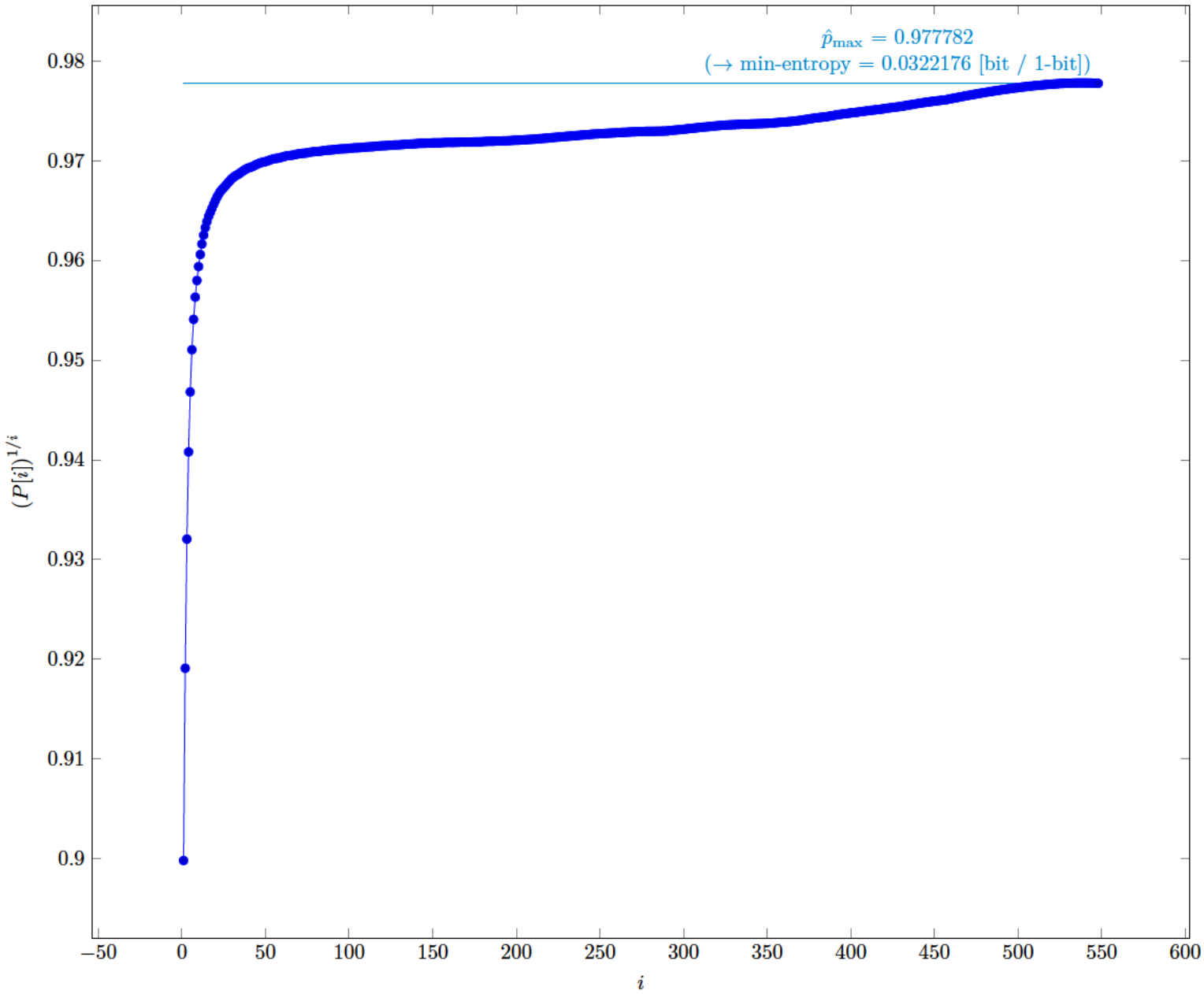


Fig. 15 $P[i]^{1/i}$ in §6.3.5 of NIST SP 800-90B

4.5.1 Supplemental information for traceability

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

Symbol	Value
t	548
$\hat{p}_{\max}$	0.977782
p_u	0.977916

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

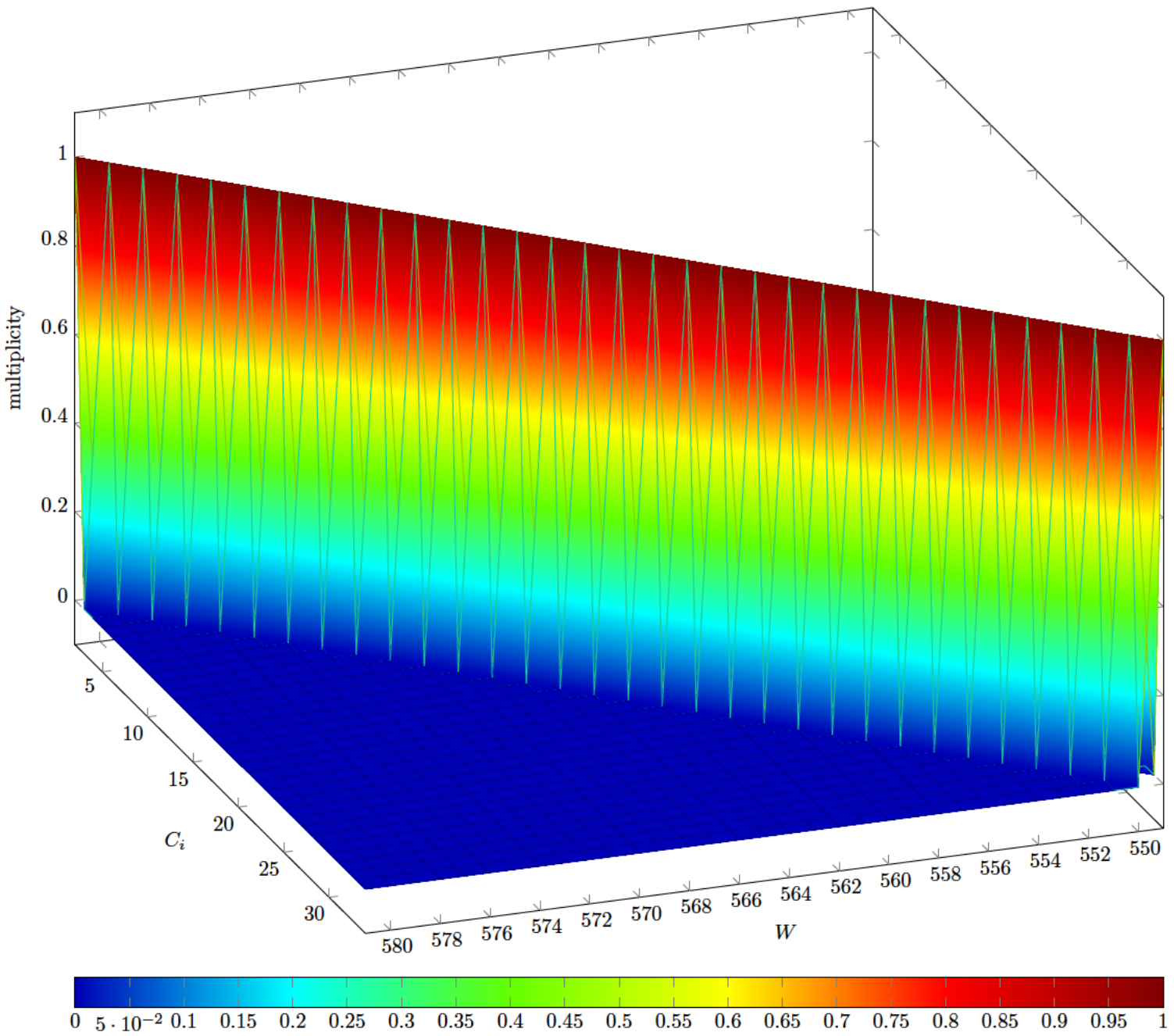


Fig. 16 Estimated W-tuple collision probability in Step 3 of §6.3.6 of NIST SP 800-90B

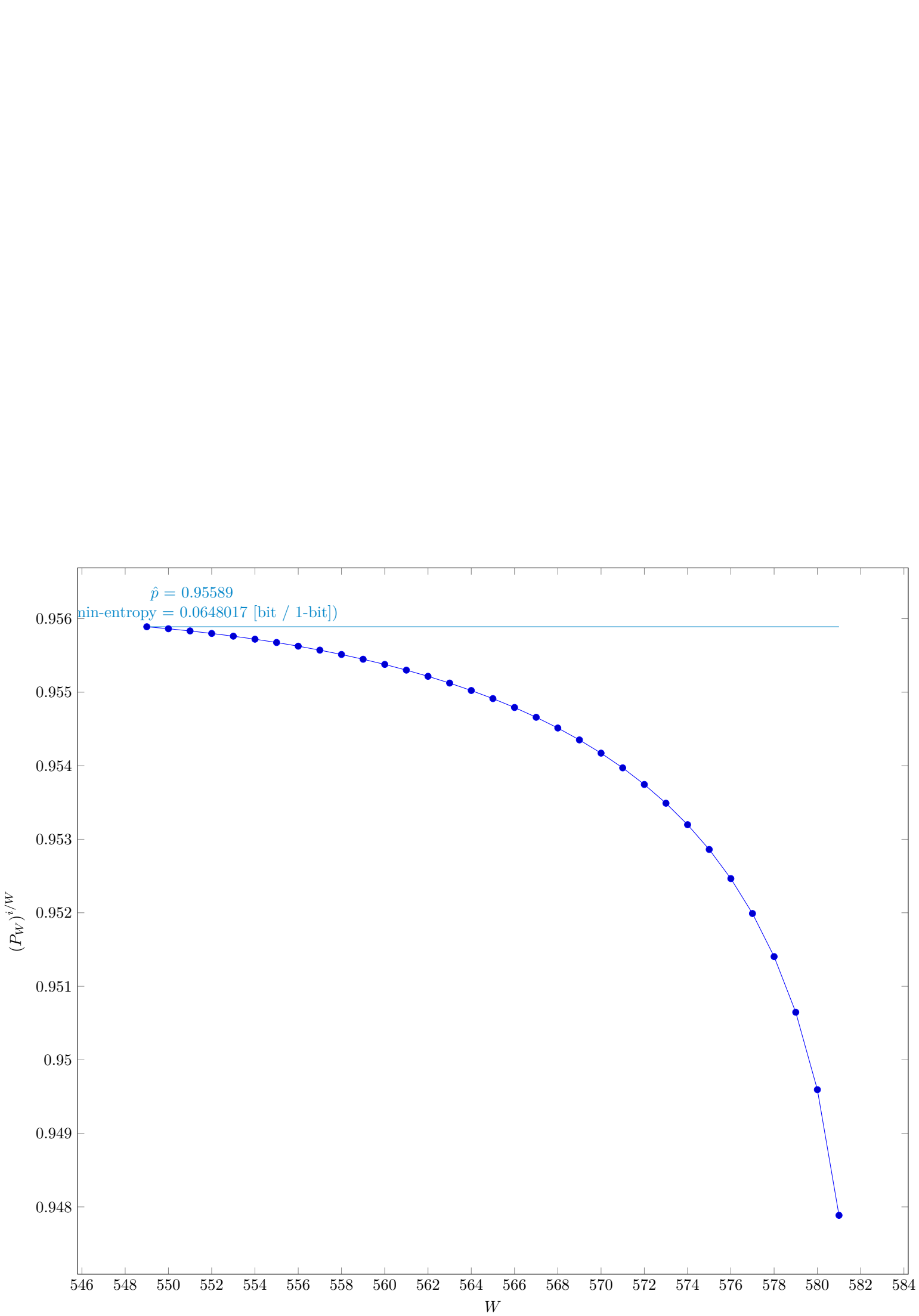


Fig. 17 Estimated average collision probability per string symbol in Step 3 of §6.3.6 of NIST SP 800-90B

4.6.1 Supplemental information for traceability

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

Symbol	Value
u	549
v	581
$\hat{p}$	0.95589
p_u	0.956077

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

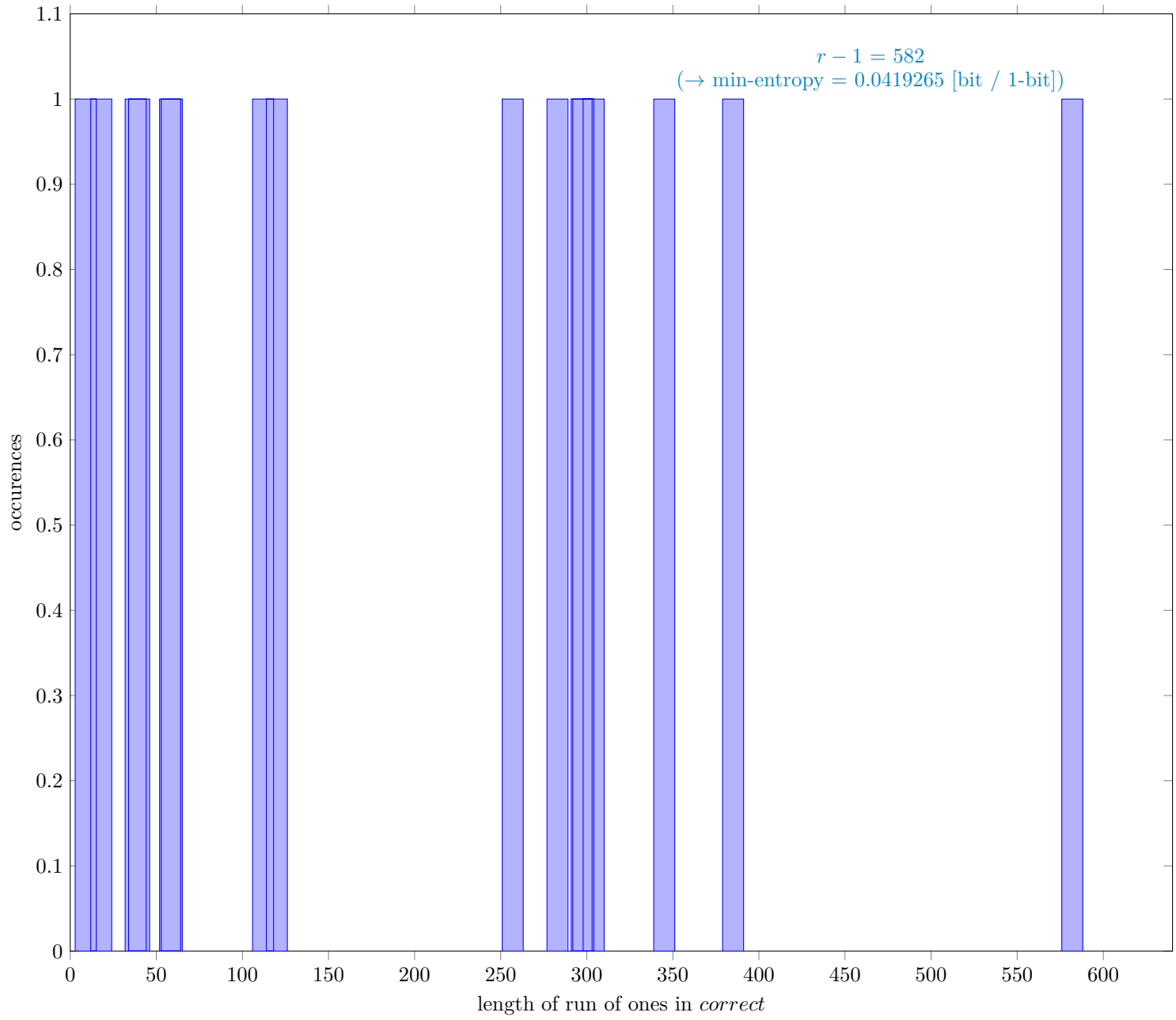


Fig. 18 Distribution of *correct*

4.7.1 Supplemental information for traceability

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

Symbol	Value
N	7999937
C	7198538
P_{global}	0.899824
P'_{global}	0.900098
r	583
P_{local}	0.971357

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

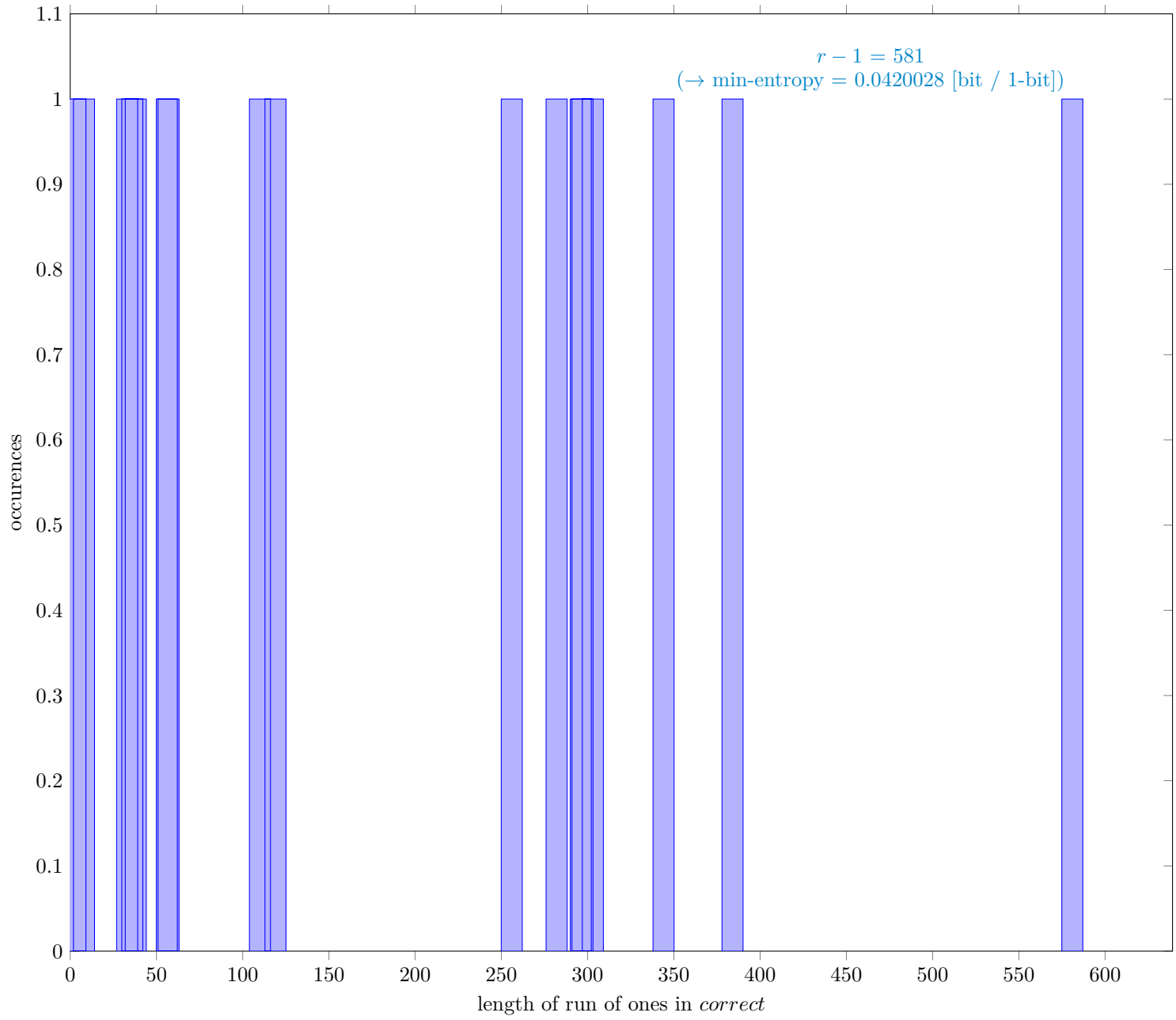


Fig. 19 Distribution of *correct*

4.8.1 Supplemental information for traceability

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

Symbol	Value
N	7999999
C	7118705
P_{global}	0.889838
P'_{global}	0.890123
r	582
P_{local}	0.971306

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

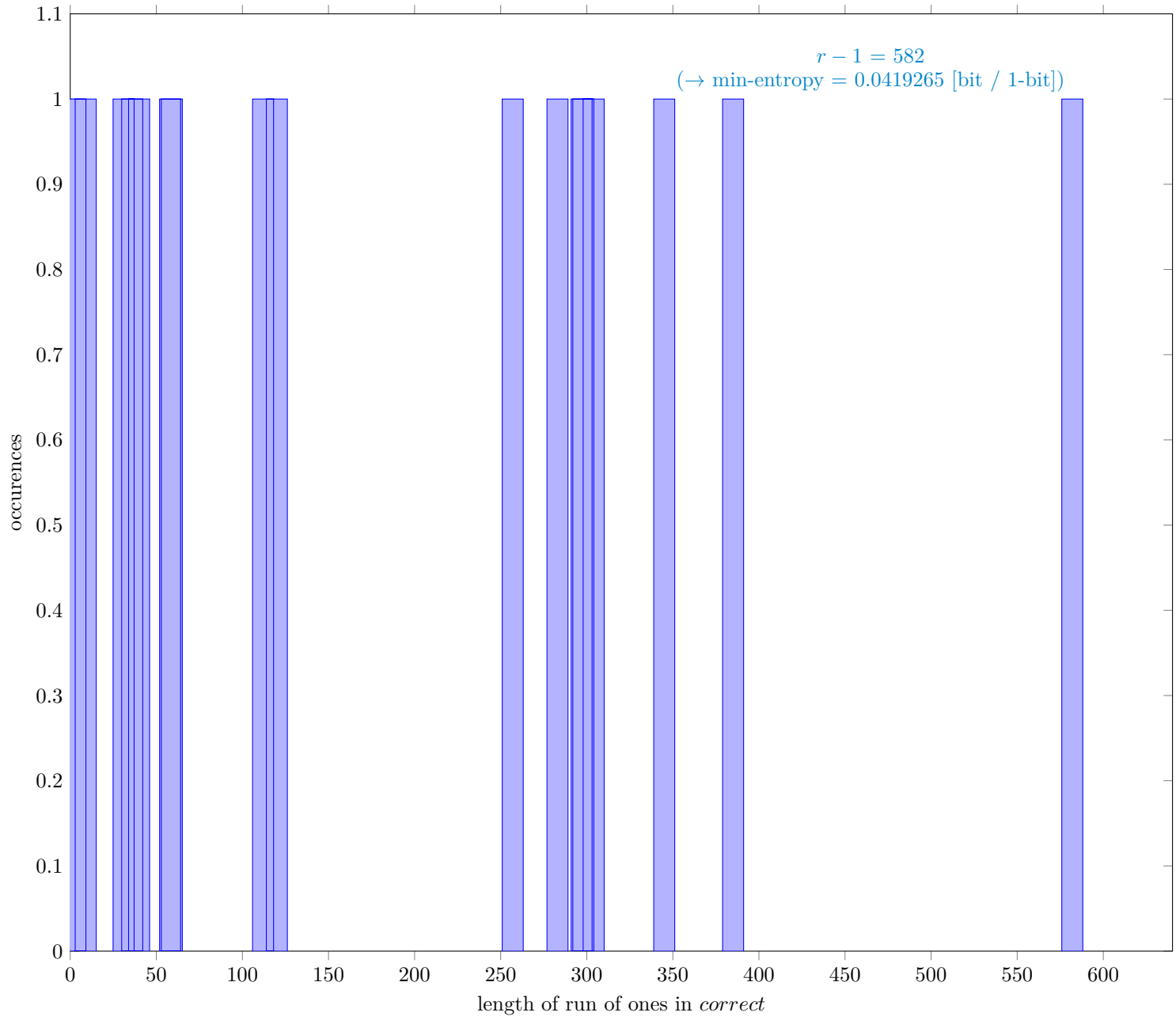


Fig. 20 Distribution of *correct*

4.9.1 Supplemental information for traceability

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

Symbol	Value
N	7999998
C	7198679
P_{global}	0.899835
P'_{global}	0.900109
r	583
P_{local}	0.971357

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

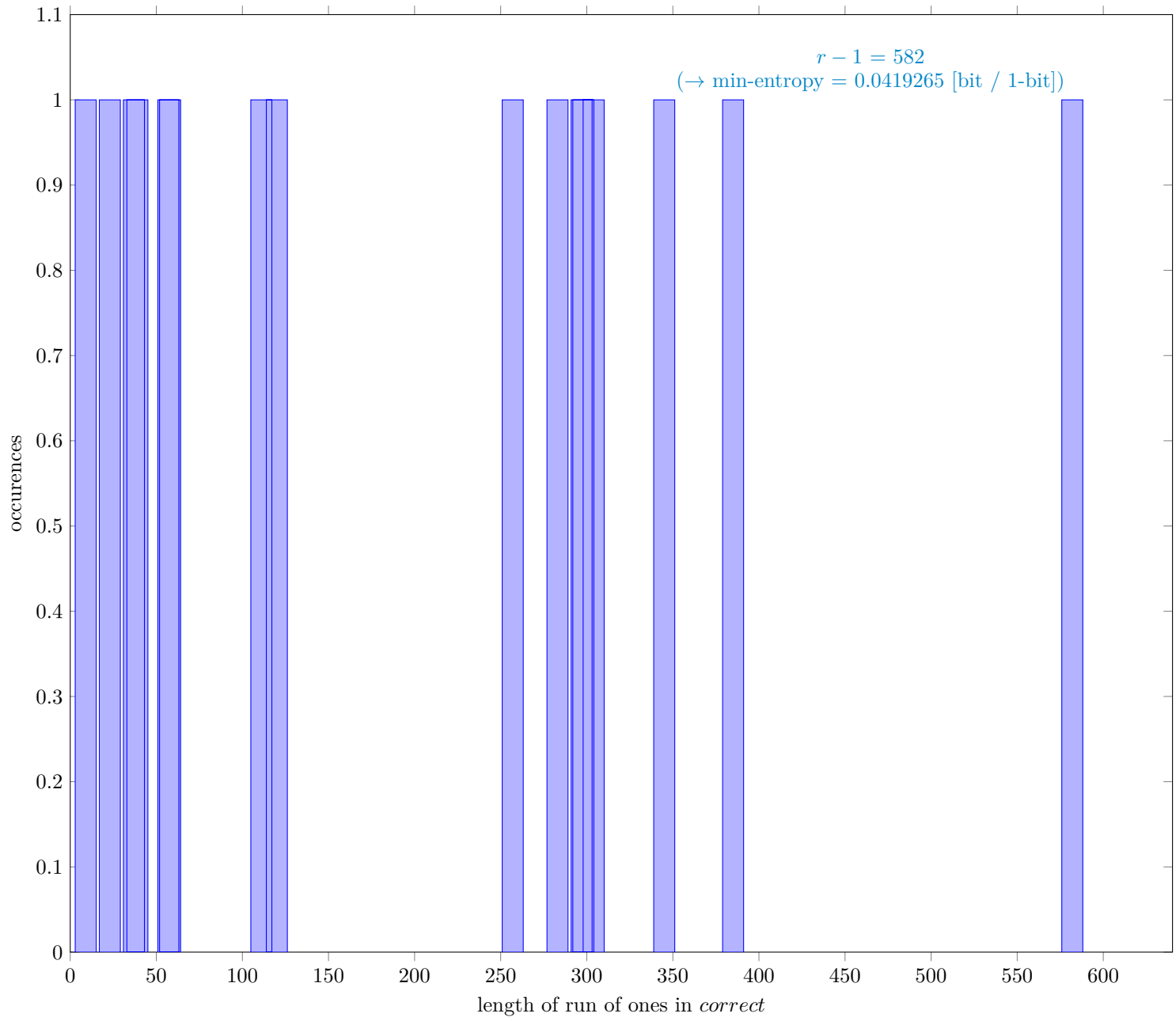


Fig. 21 Distribution of *correct*

4.10.1 Supplemental information for traceability

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

Symbol	Value
N	7999983
C	7198670
P_{global}	0.899836
P'_{global}	0.900109
r	583
P_{local}	0.971357

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

[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`