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

2023-Aug-20 17:12:40.856579

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`
SHA-256 hash value of the acquisition data [hex]	`146bd749 7d8e2d61 a6e8559c 9342ee79 f6005a39 0ee4d776 ba43500d 00eb508d`

- 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	1.0.50
	built as	64-bit application
	built by	Intel C++ Compiler (`__INTEL_LLVM_COMPILER`: 20230200)
	linked libraries	Boost C++ 1.83.0
Analysis environment	Hostname	[redacted]
	CPU information	AMD Ryzen [redacted]
	Physical memory size	[redacted] MiB
	OS information	Windows 10 or greater 64-bit
	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]	Notes to H_{original}	$H_{\text{bitstring}}$ [b] [bit / 1 - bit]	Notes to $H_{\text{bitstring}}$
The Most Common Value Estimate	0.319651	see 3.1	0.151827	see 4.1
The Collision Estimate	—	—	0.0727058	see 4.2
The Markov Estimate	—	—	0.0916044	see 4.3
The Compression Estimate	—	—	0.0631355	see 4.4
The t-Tuple Estimate	0.29116	see 3.2	0.0322176	see 4.5
The Longest Repeated Substring (LRS) Estimate	0.519281	see 3.3	0.0648017	see 4.6
Multi Most Common in Window Prediction Estimate	0.319646	see 3.4	0.0419265	see 4.7
The Lag Prediction Estimate	0.466258	see 3.5	0.0420028	see 4.8
The MultiMMC Prediction Estimate	0.320277	see 3.6	0.0419265	see 4.9
The LZ78Y Prediction Estimate	0.330375	see 3.7	0.0419265	see 4.10
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

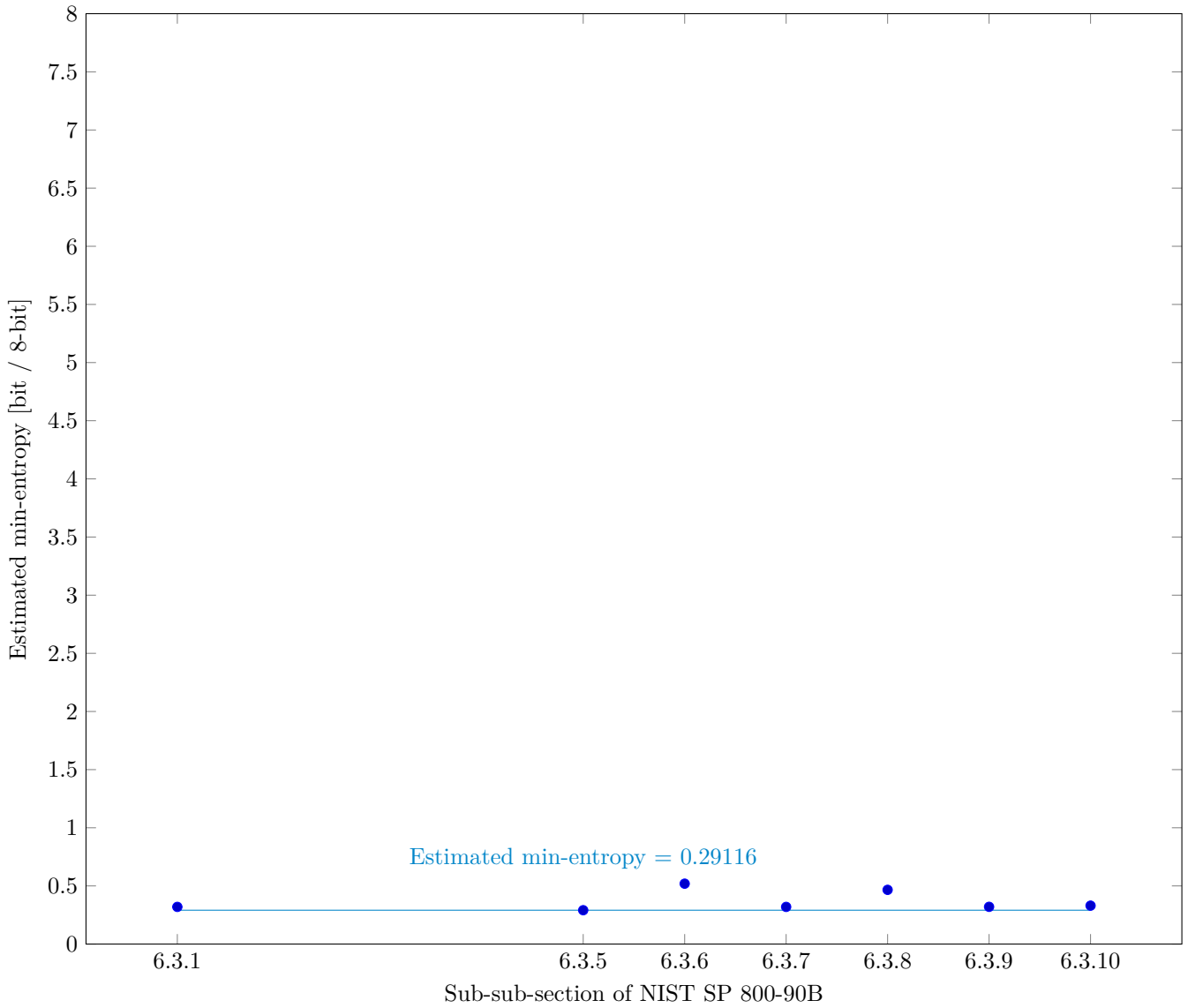

Fig. 1 Estimated Min-Entropy using §6.3 of NIST SP 800-90B

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

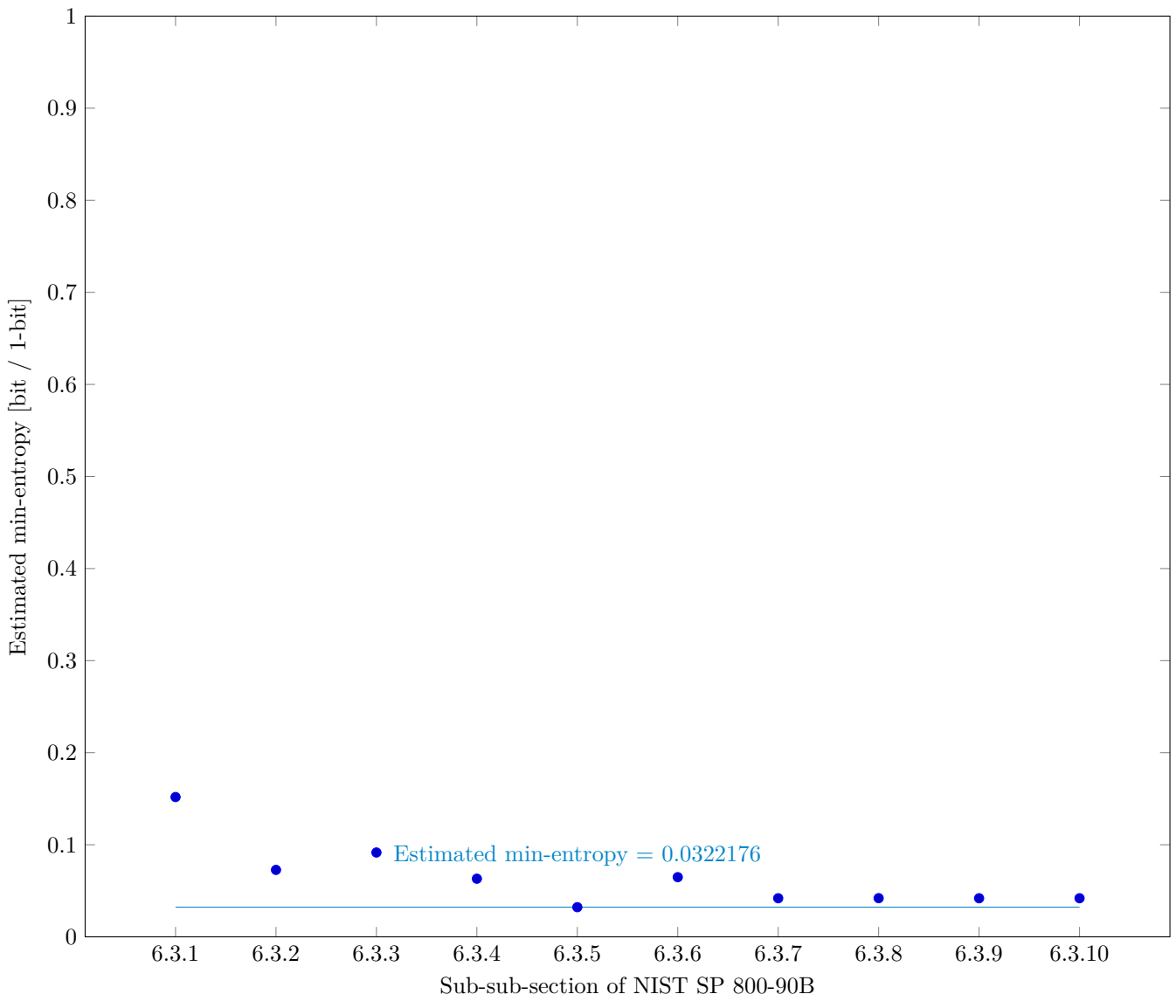

Fig. 2 Estimated Min-Entropy using §6.3 of NIST SP 800-90B

3 Detailed results of analysis from original samples

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

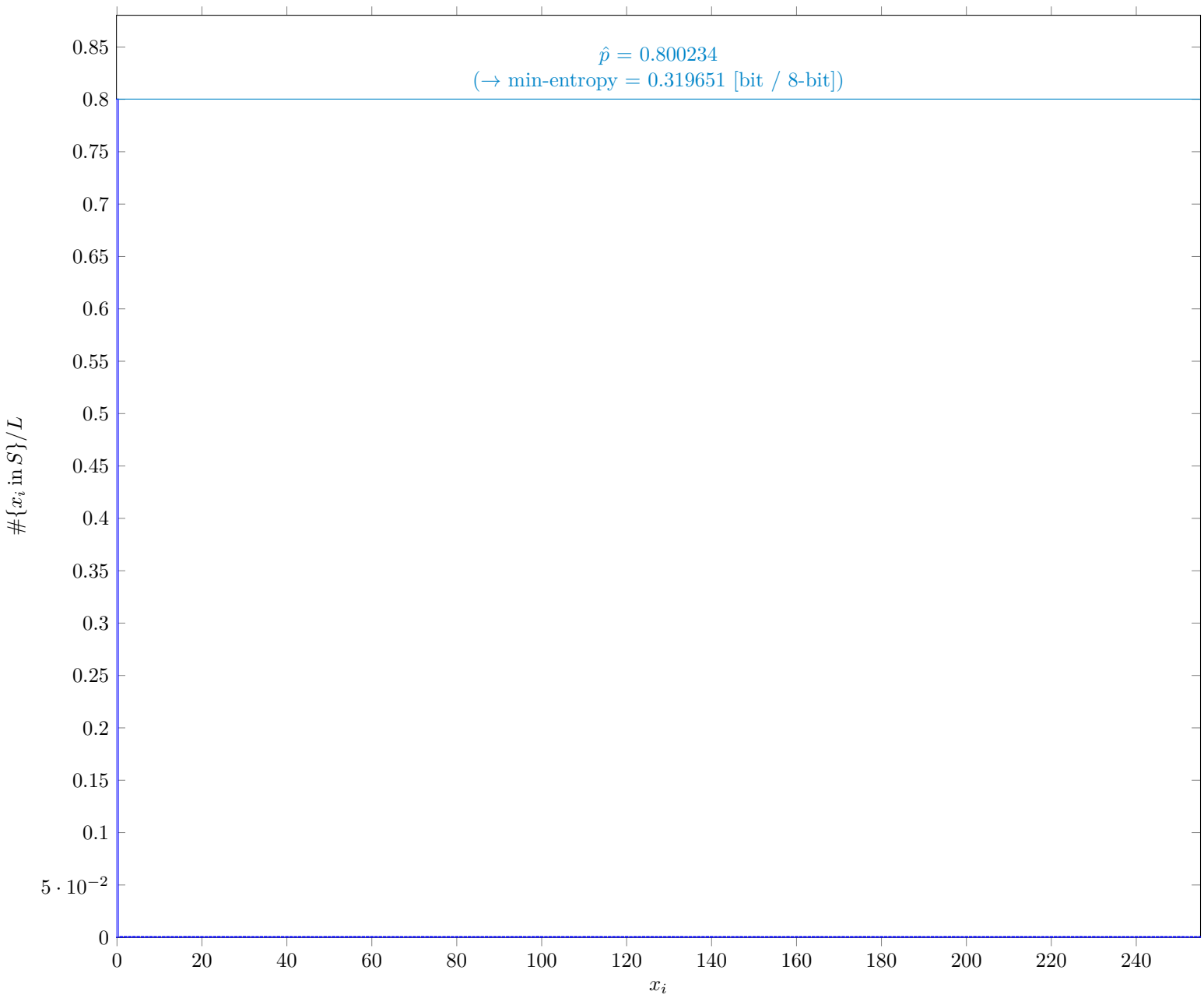

Fig. 3 Distribution of x_i

3.1.1 Supplemental information for traceability

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

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

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

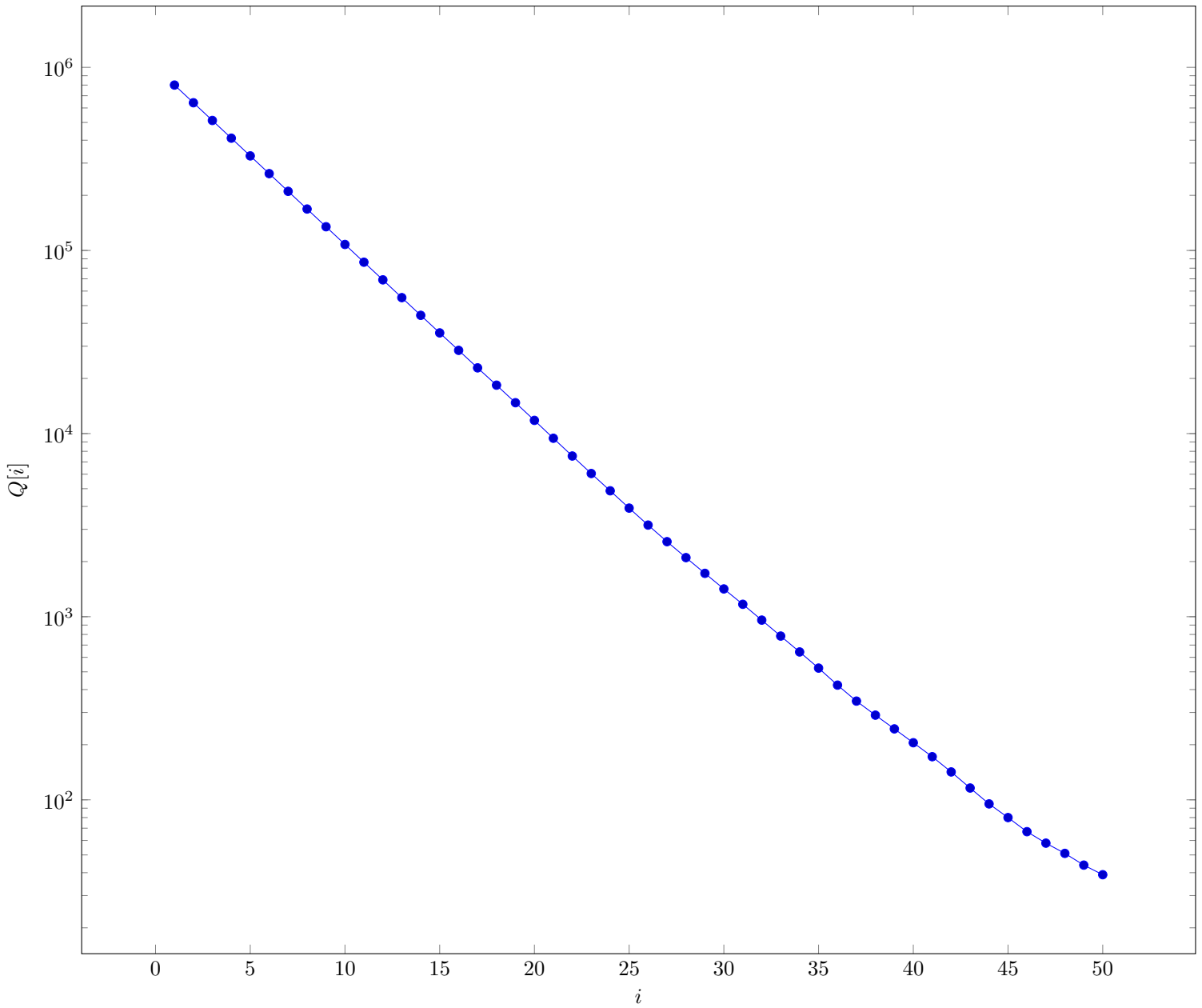

Fig. 4 Intermediate value $Q[i]$ in §6.3.5 of NIST SP 800-90B

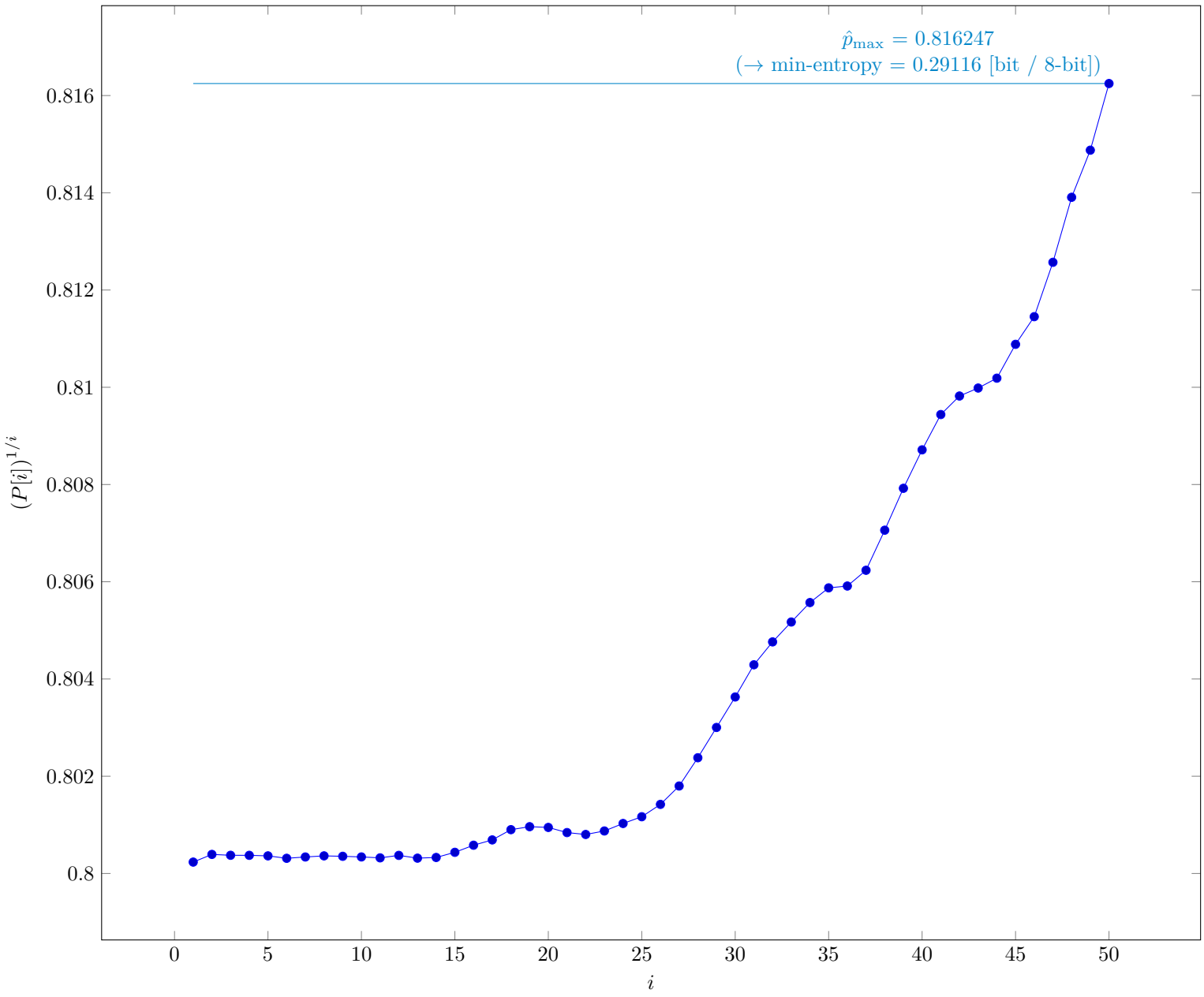

Fig. 5 $P[i]^{1/i}$ in §6.3.5 of NIST SP 800-90B

3.2.1 Supplemental information for traceability

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

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

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

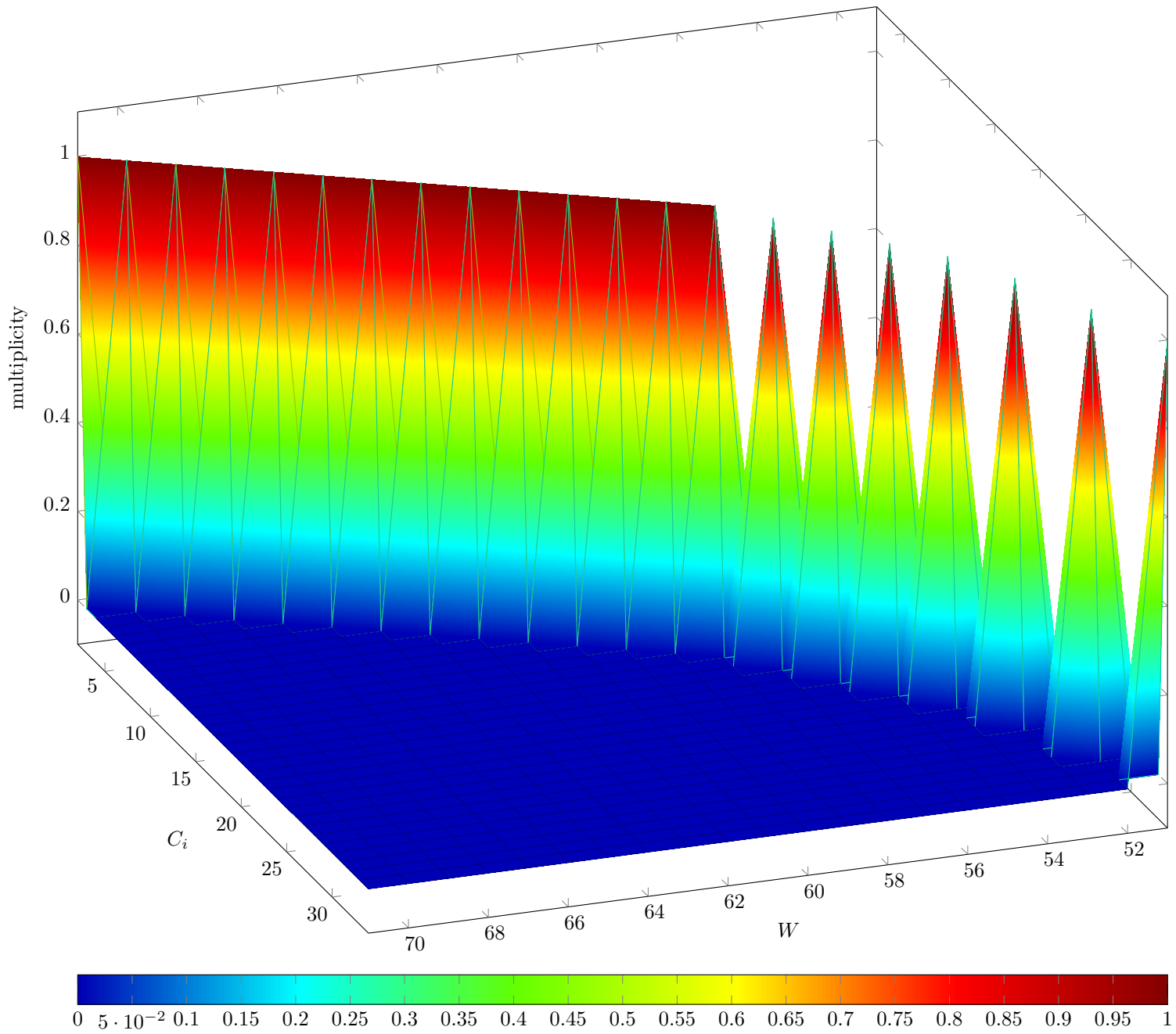

Fig. 6 Estimated W-tuple collision probability in Step 3 of §6.3.6 of NIST SP 800-90B

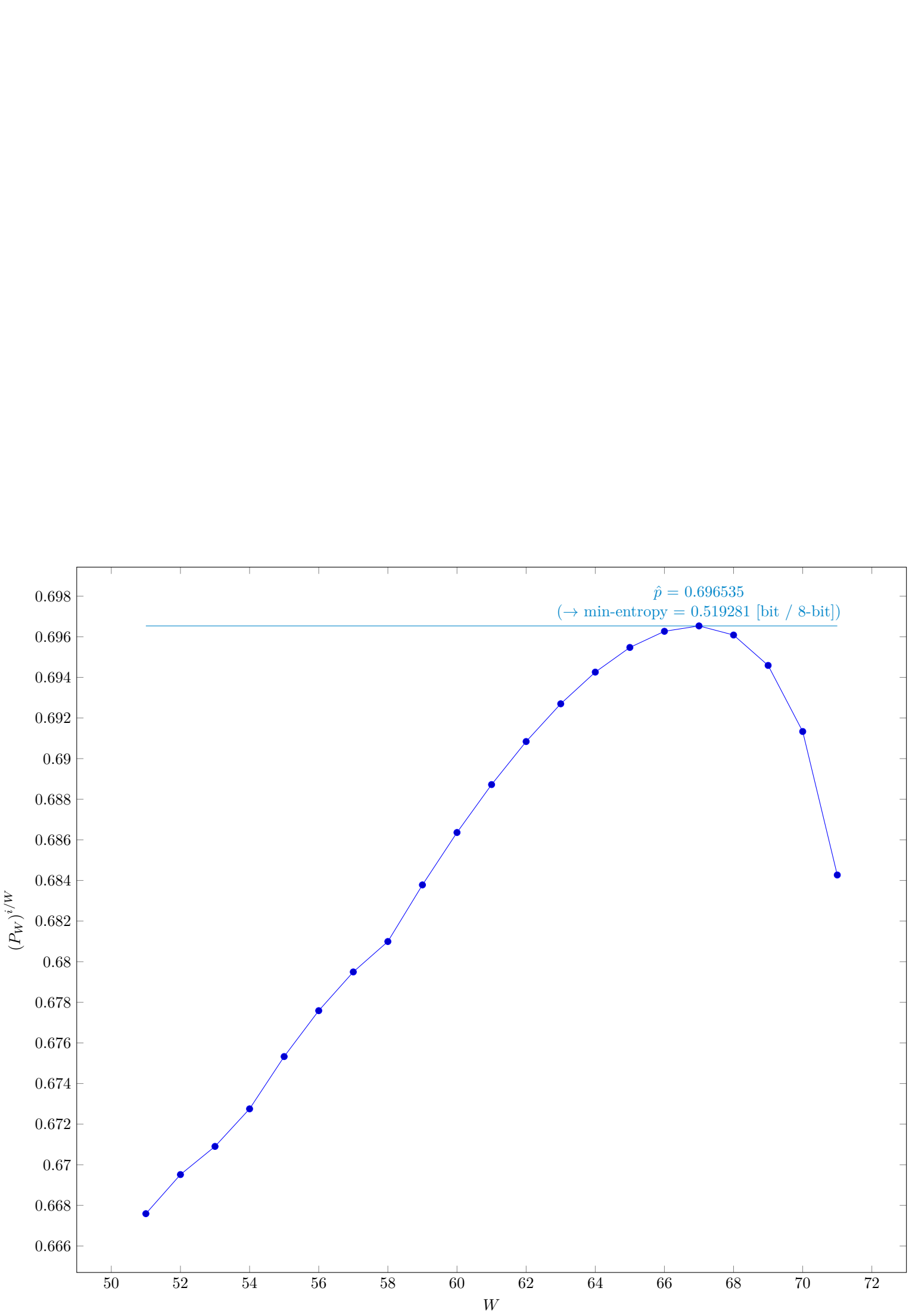

Fig. 7 Estimated average collision probability per string symbol in Step 3 of §6.3.6 of NIST SP 800-90B

3.3.1 Supplemental information for traceability

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

Symbol	Value
u	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)

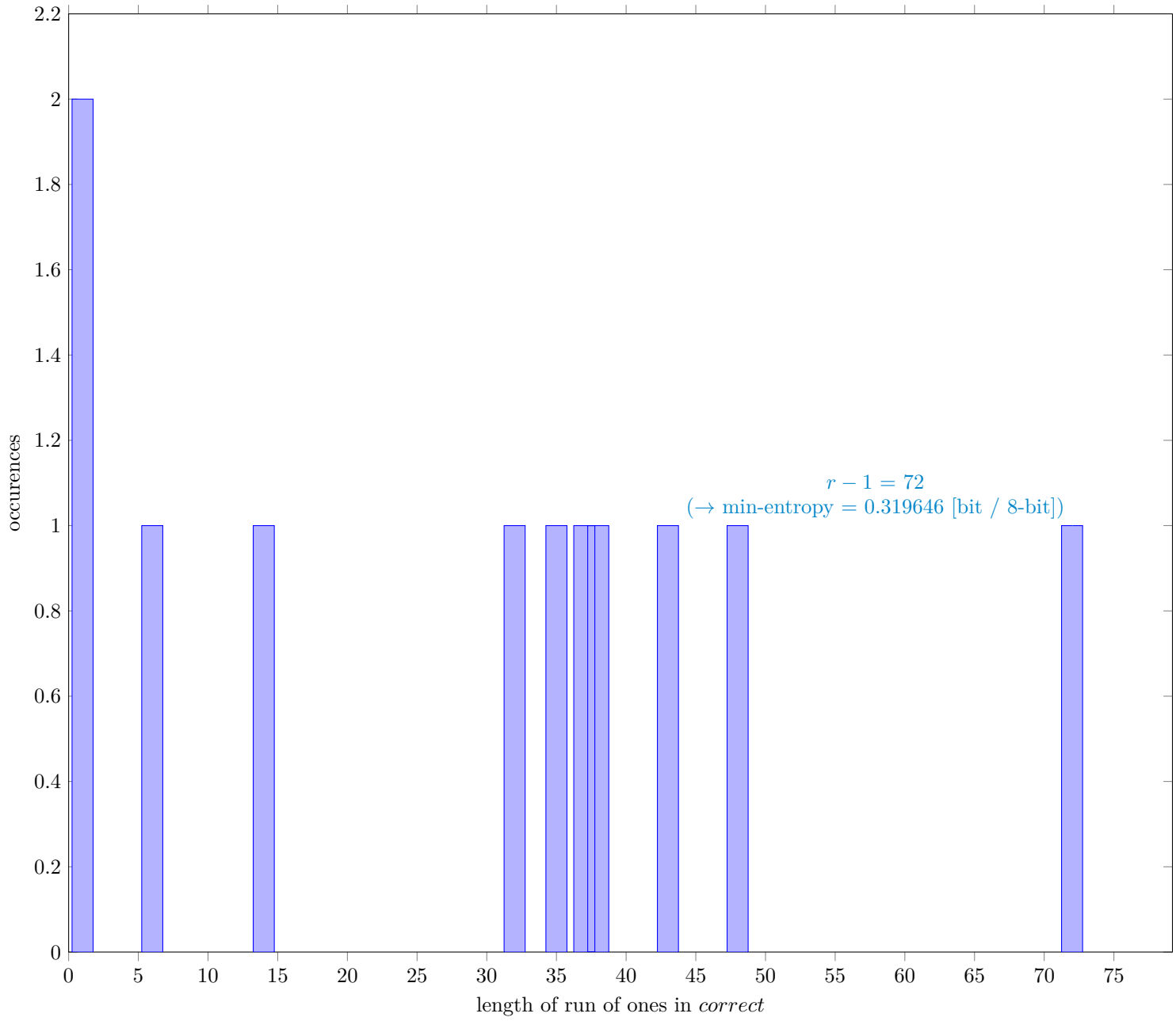

Fig. 8 Distribution of *correct*

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	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)

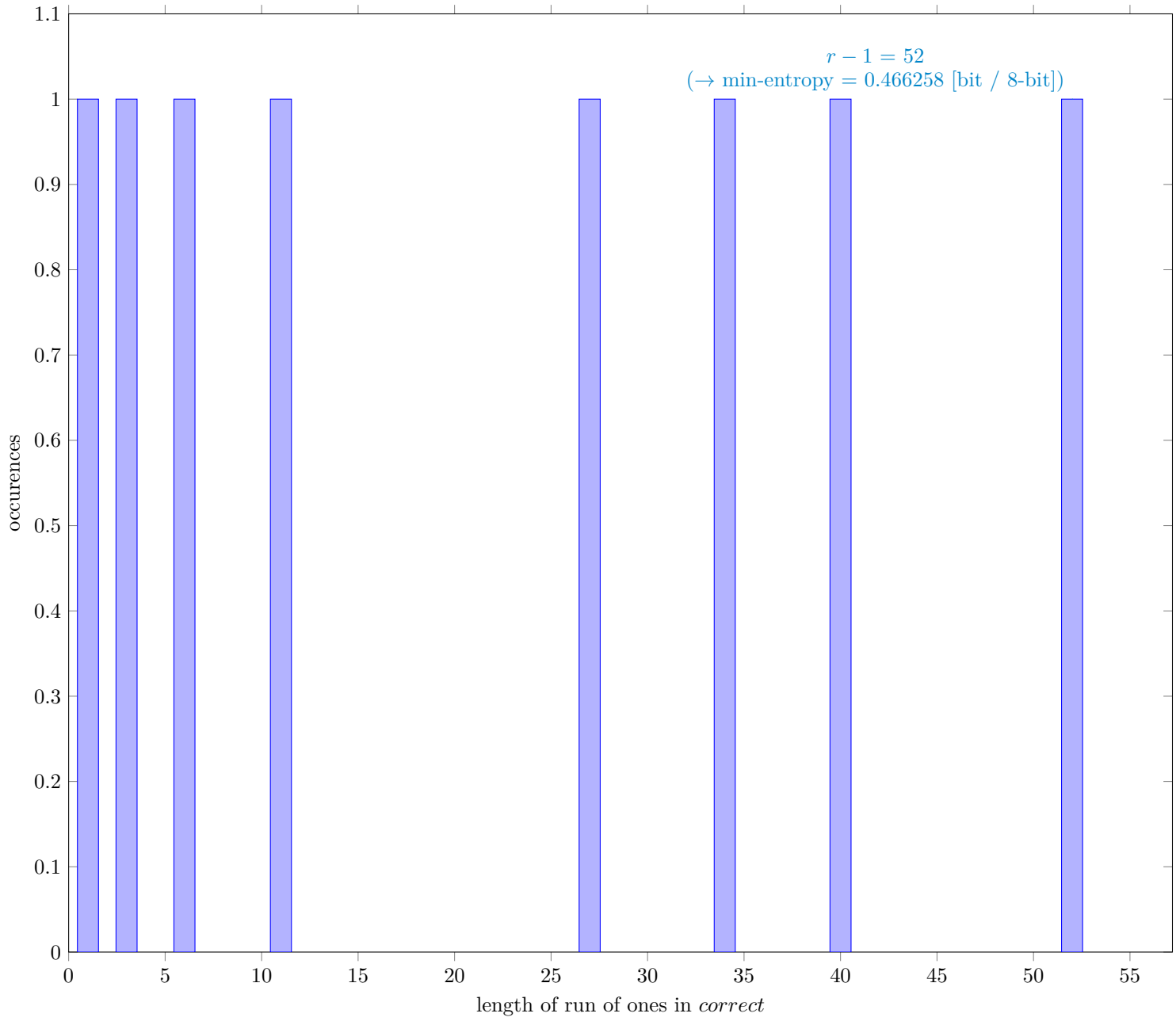

Fig. 9 Distribution of *correct*

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	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)

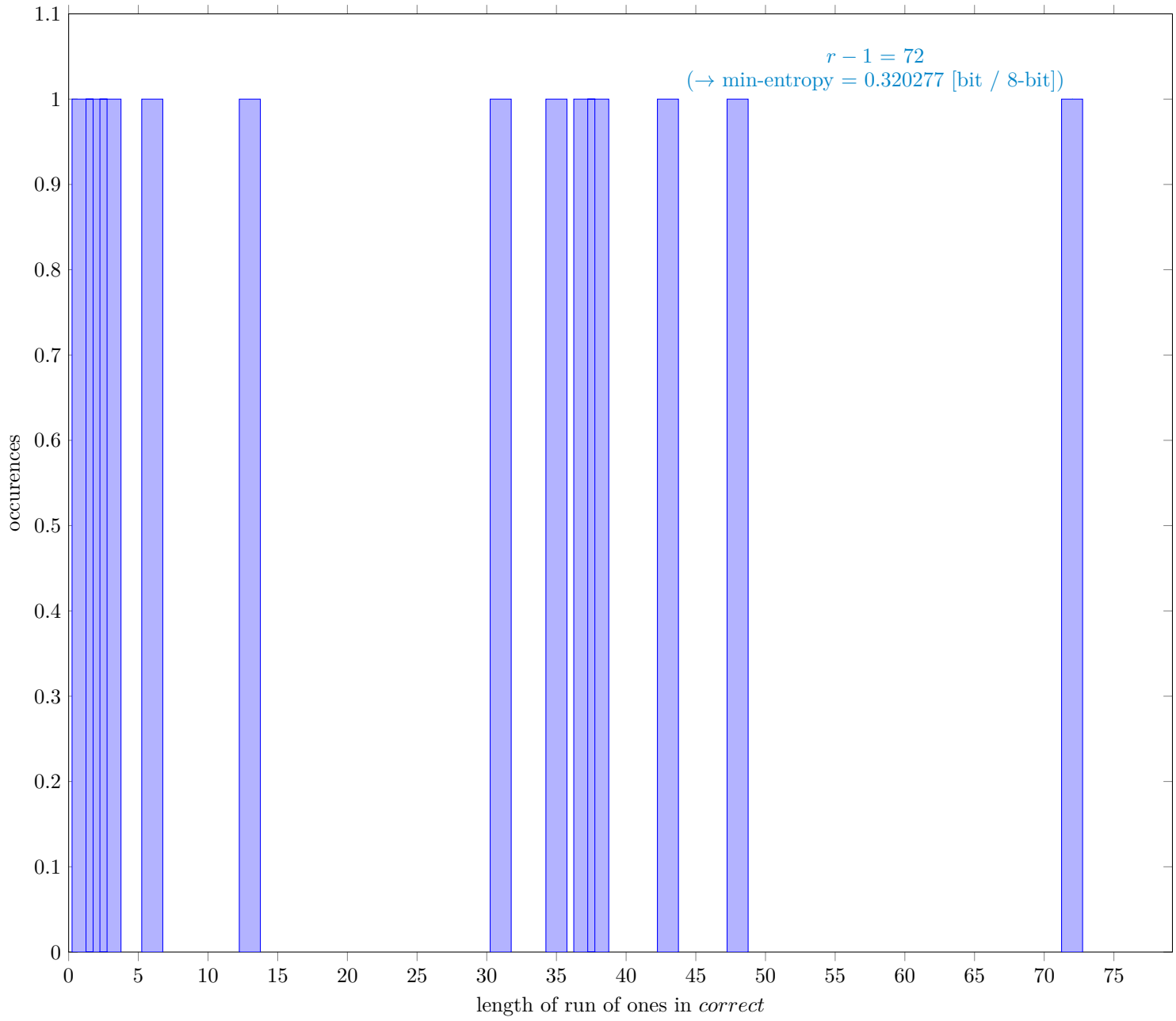

Fig. 10 Distribution of *correct*

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	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)

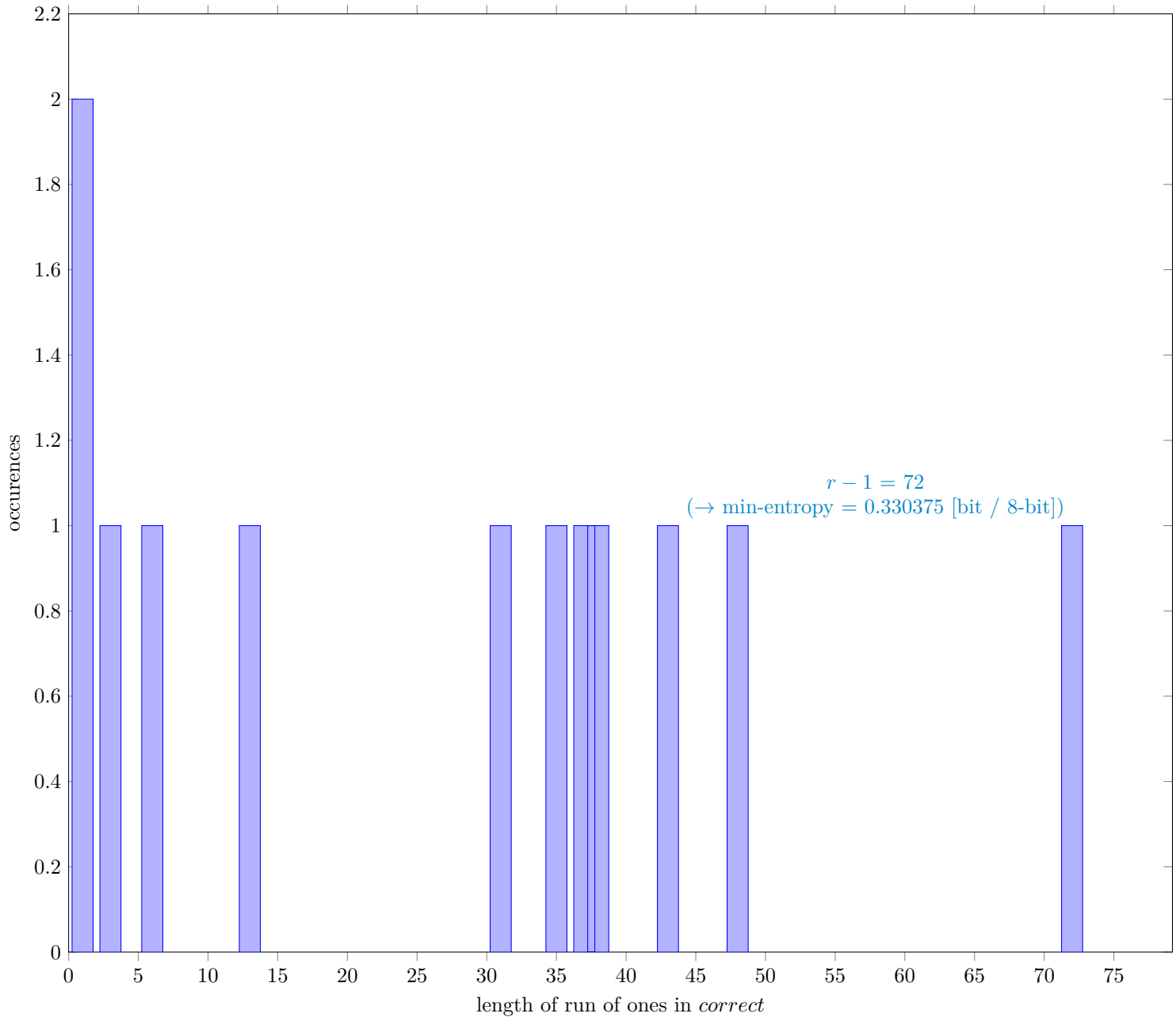

Fig. 11 Distribution of *correct*

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	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)

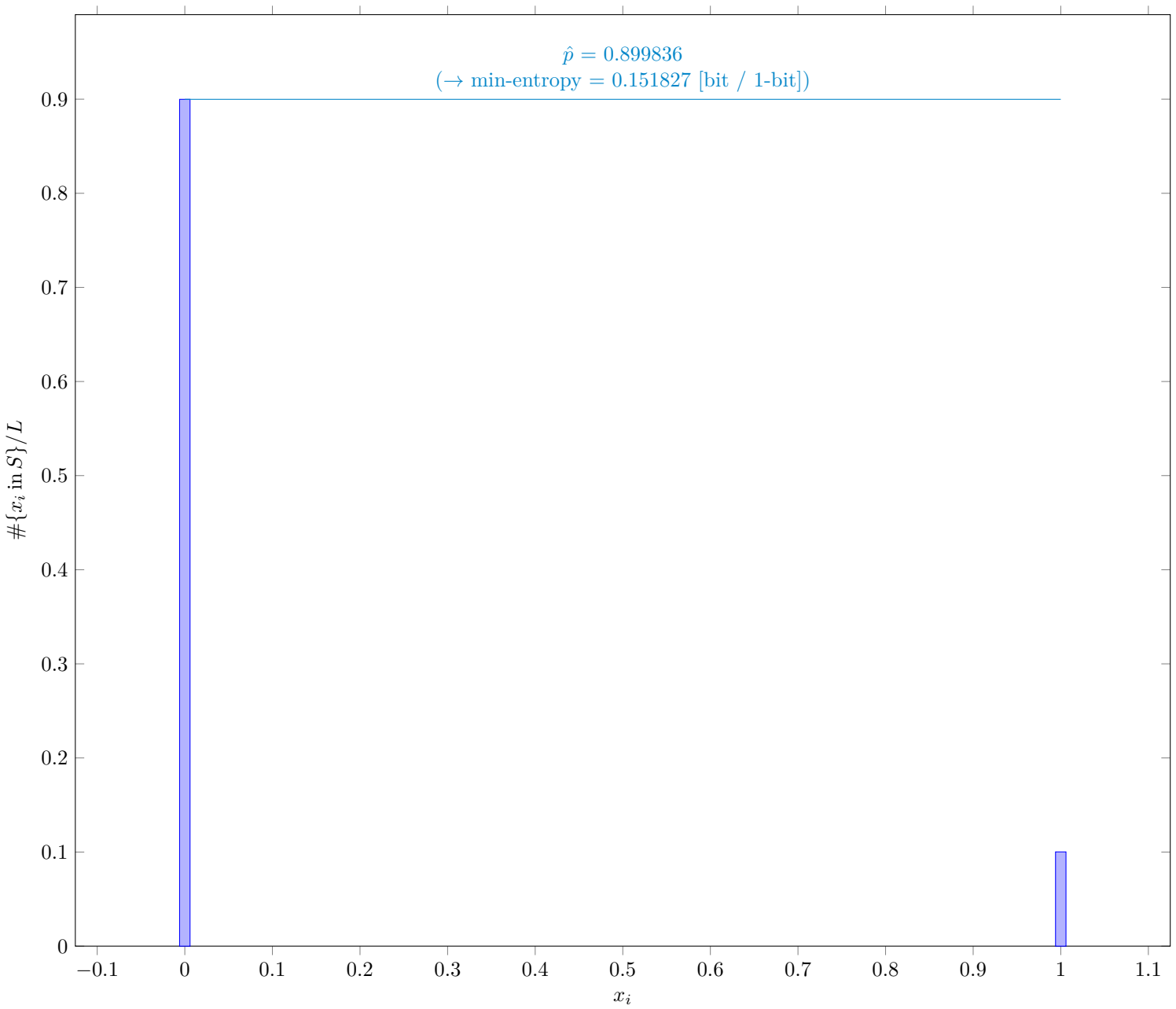

Fig. 12 Distribution of x_i

4.1.1 Supplemental information for traceability

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

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

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

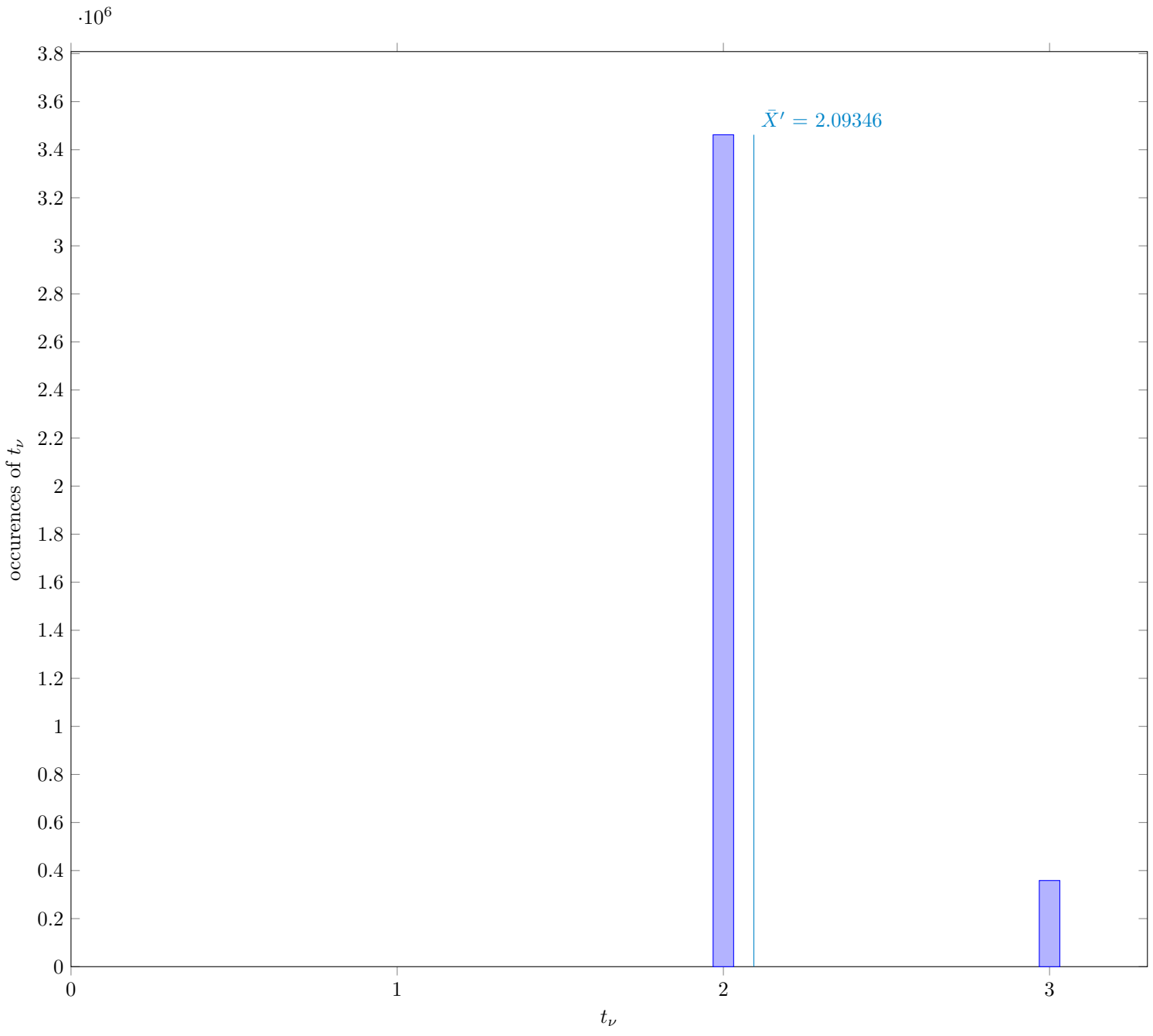

Fig. 13 Distribution of intermediate value t_ν

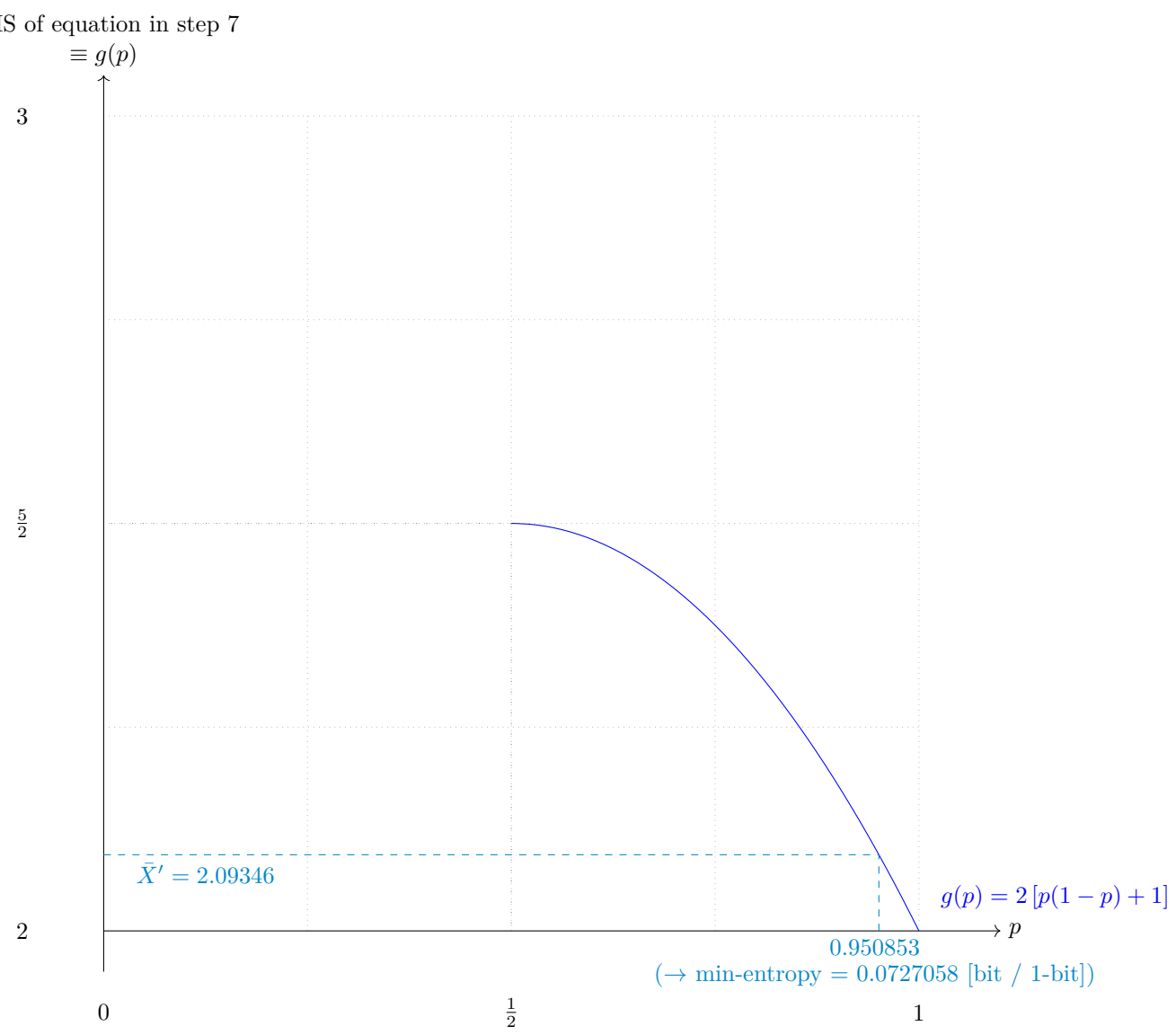

Fig. 14 Solution to the equation in step 7

4.2.1 Supplemental information for traceability

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

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

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

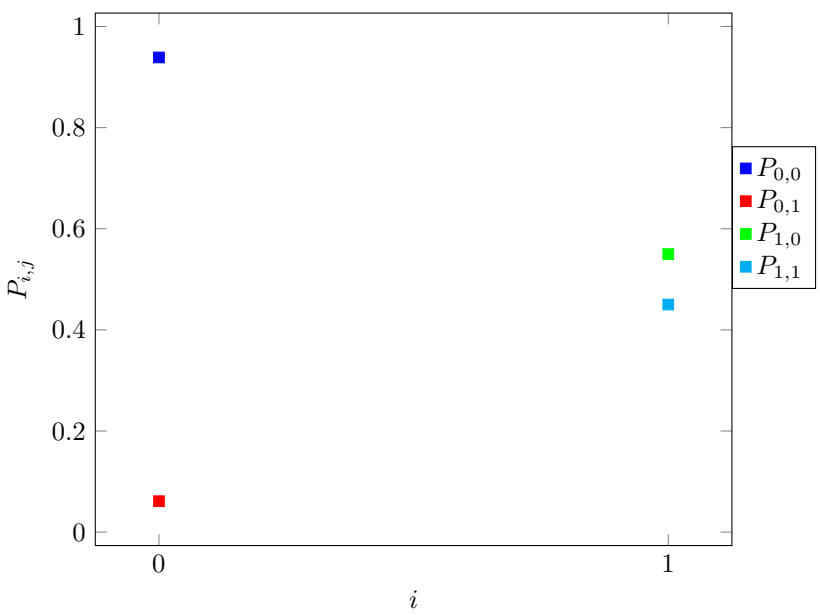

Fig. 15 Transition probability $P_{i,j}$ of §6.3.3 of NIST SP 800-90B

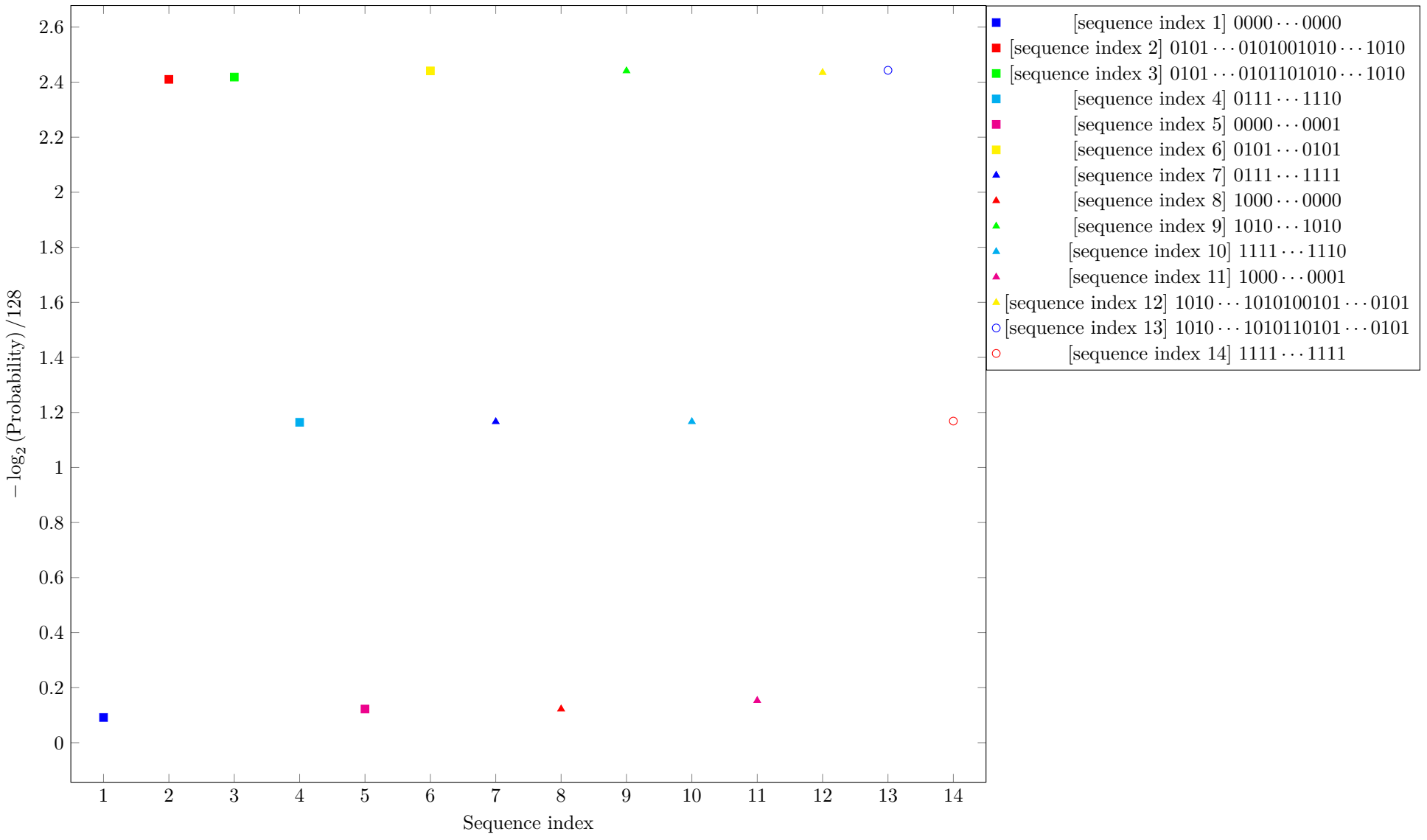

Fig. 16 Estimated Min-Entropy using §6.3.3 of NIST SP 800-90B

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

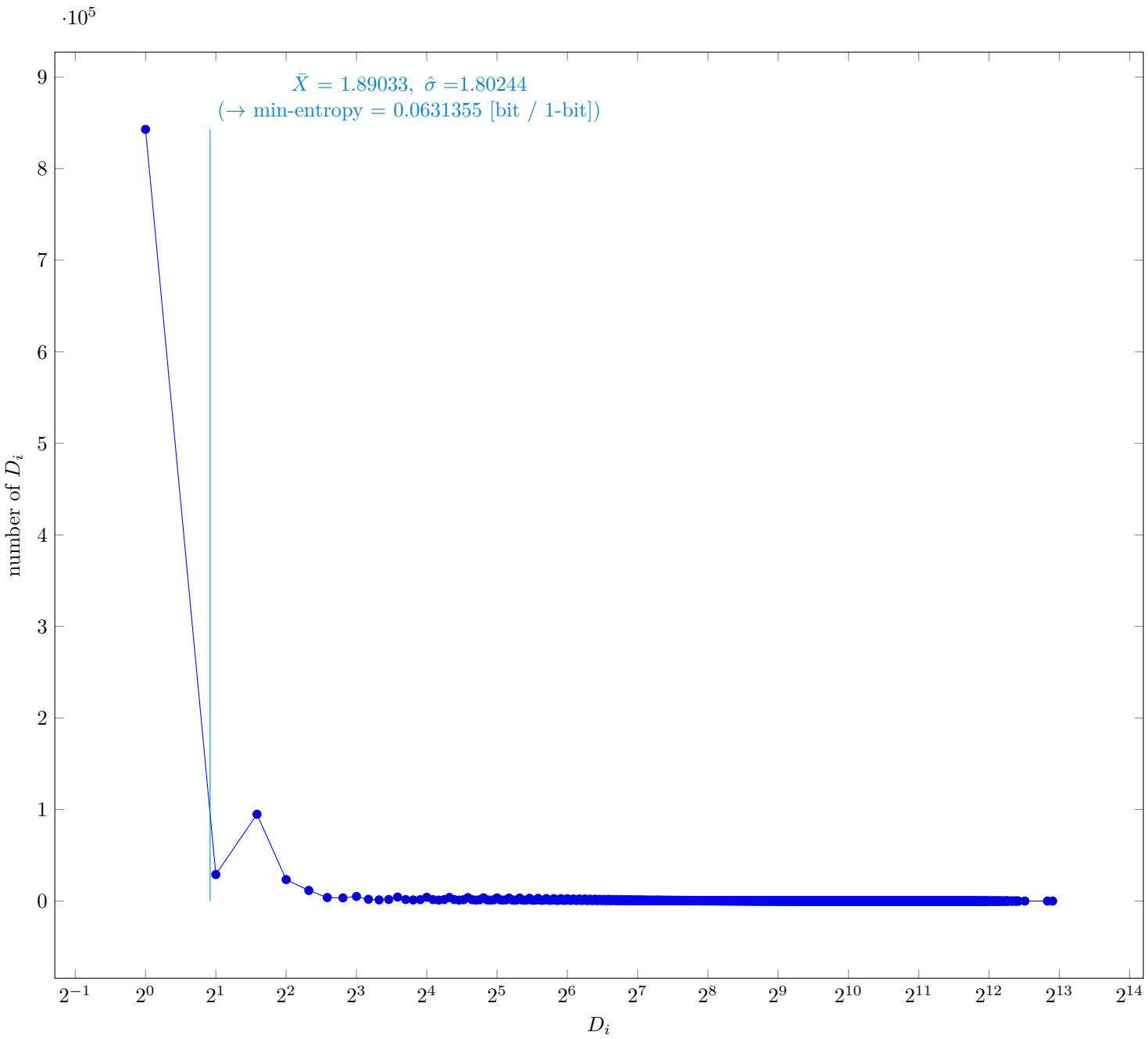

Fig. 17 Distribution of intermediate value D_i

4.4.1 Supplemental information for traceability

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

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

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

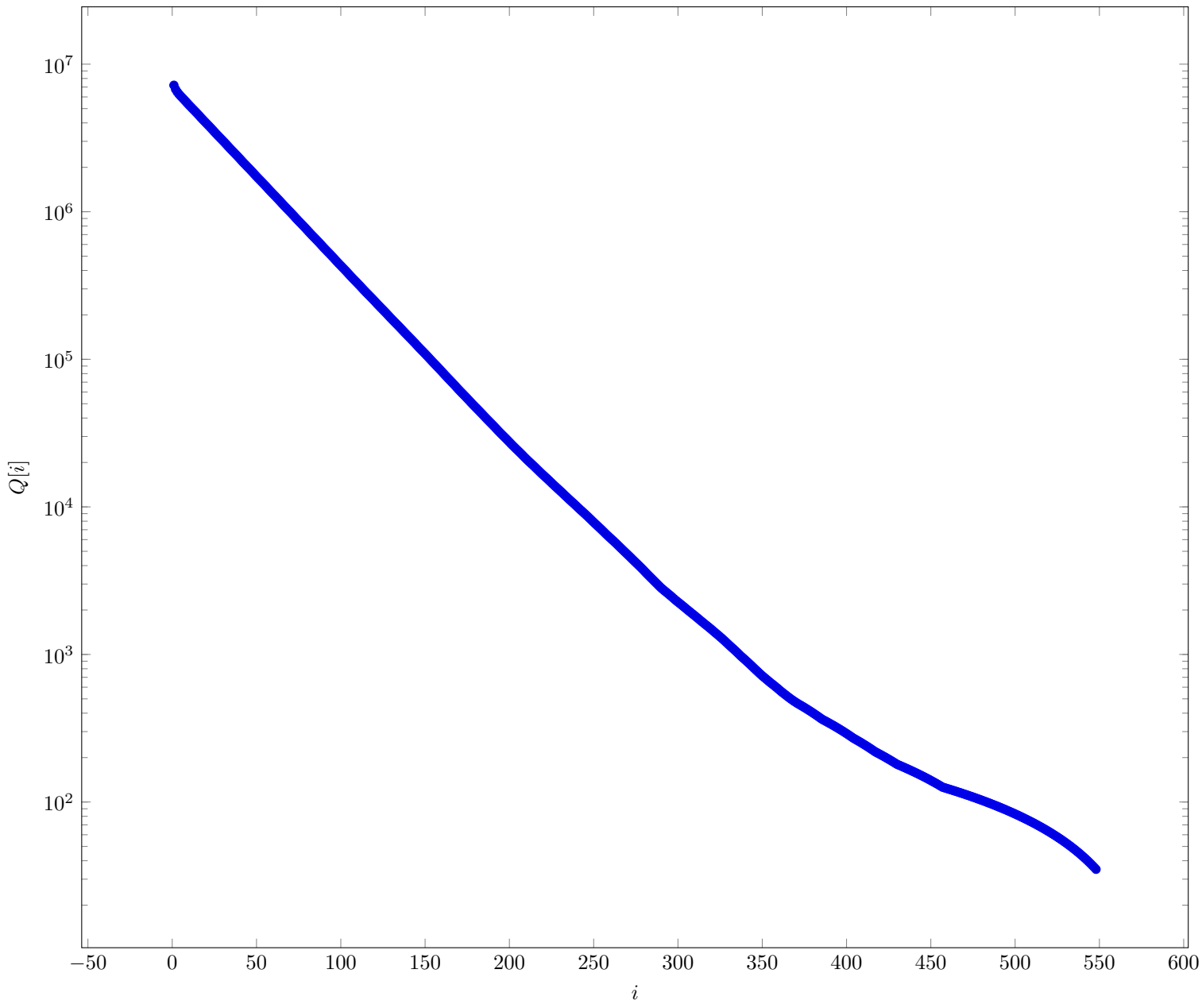

Fig. 18 Intermediate value $Q[i]$ in §6.3.5 of NIST SP 800-90B

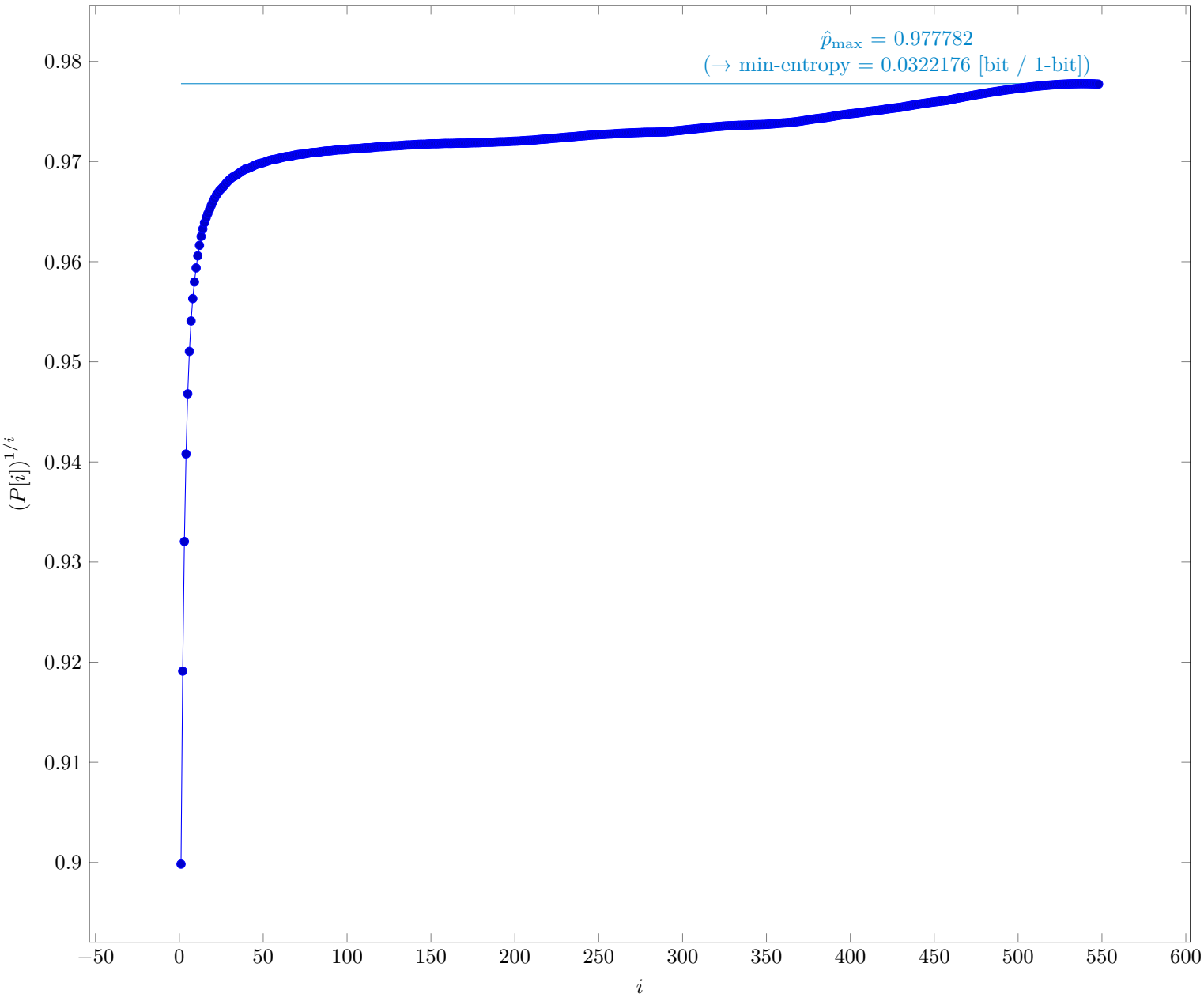

Fig. 19 $P[i]^{1/i}$ in §6.3.5 of NIST SP 800-90B

4.5.1 Supplemental information for traceability

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

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

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

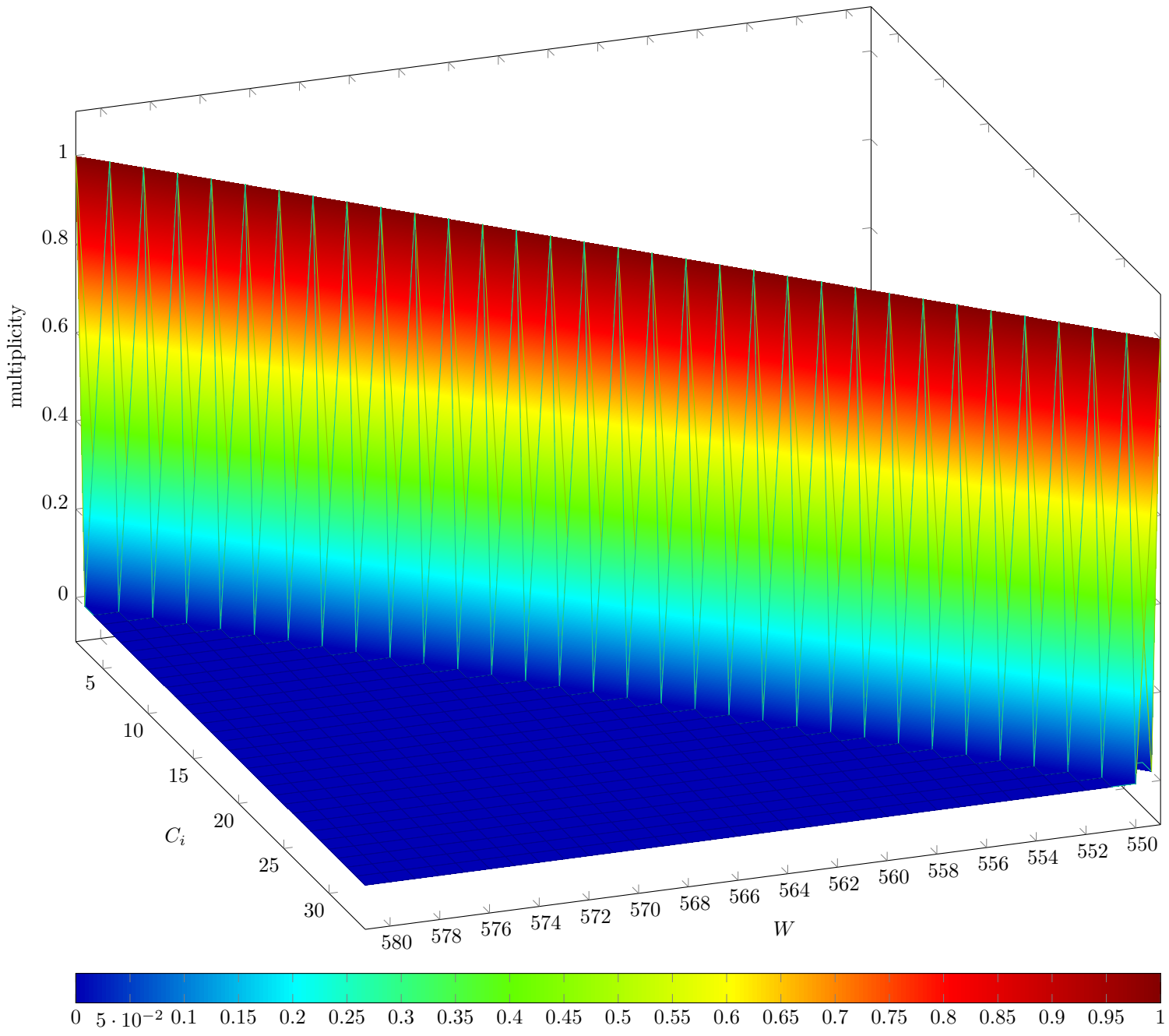

Fig. 20 Estimated W-tuple collision probability in Step 3 of §6.3.6 of NIST SP 800-90B

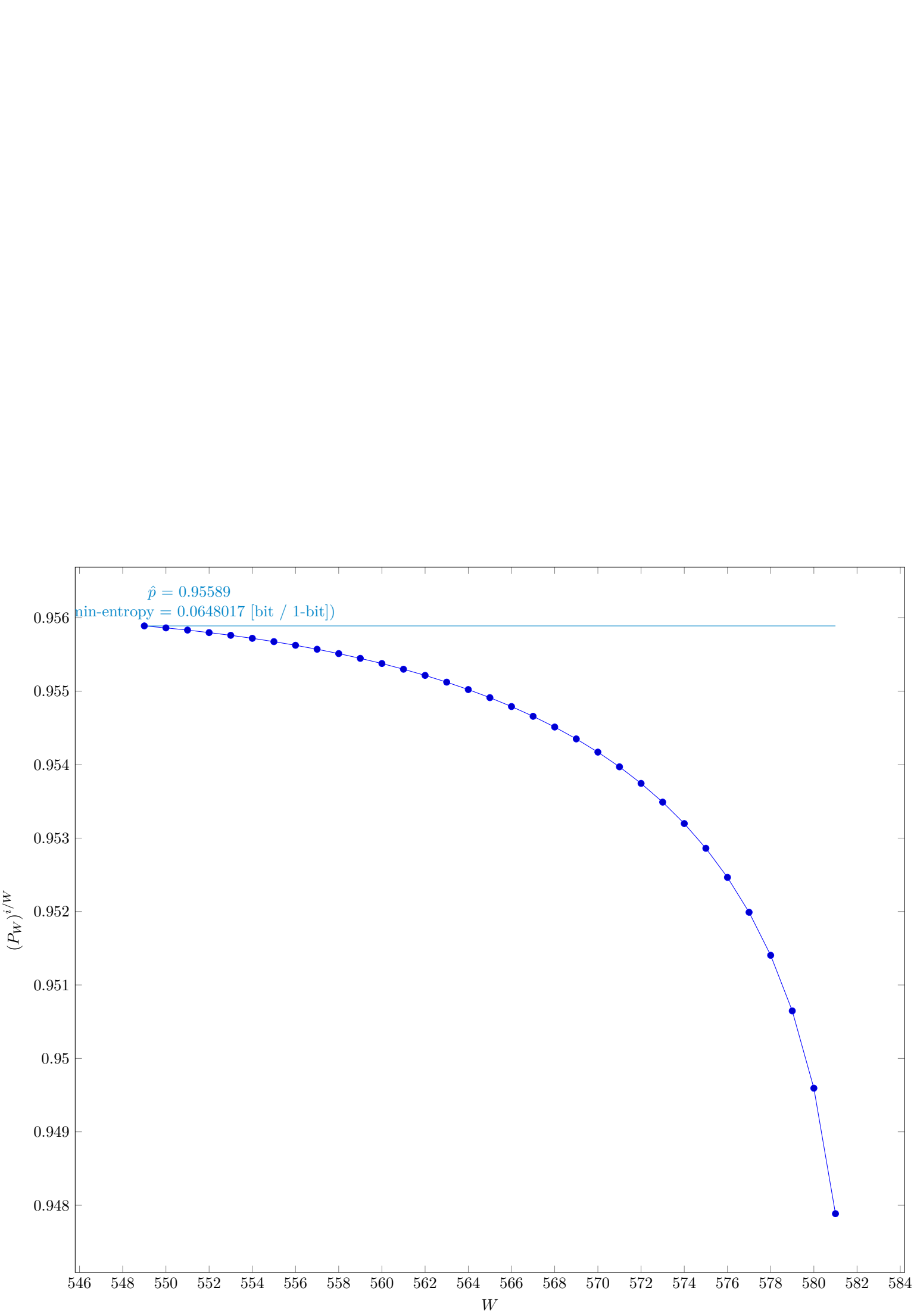

Fig. 21 Estimated average collision probability per string symbol in Step 3 of §6.3.6 of NIST SP 800-90B

4.6.1 Supplemental information for traceability

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

Symbol	Value
u	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)

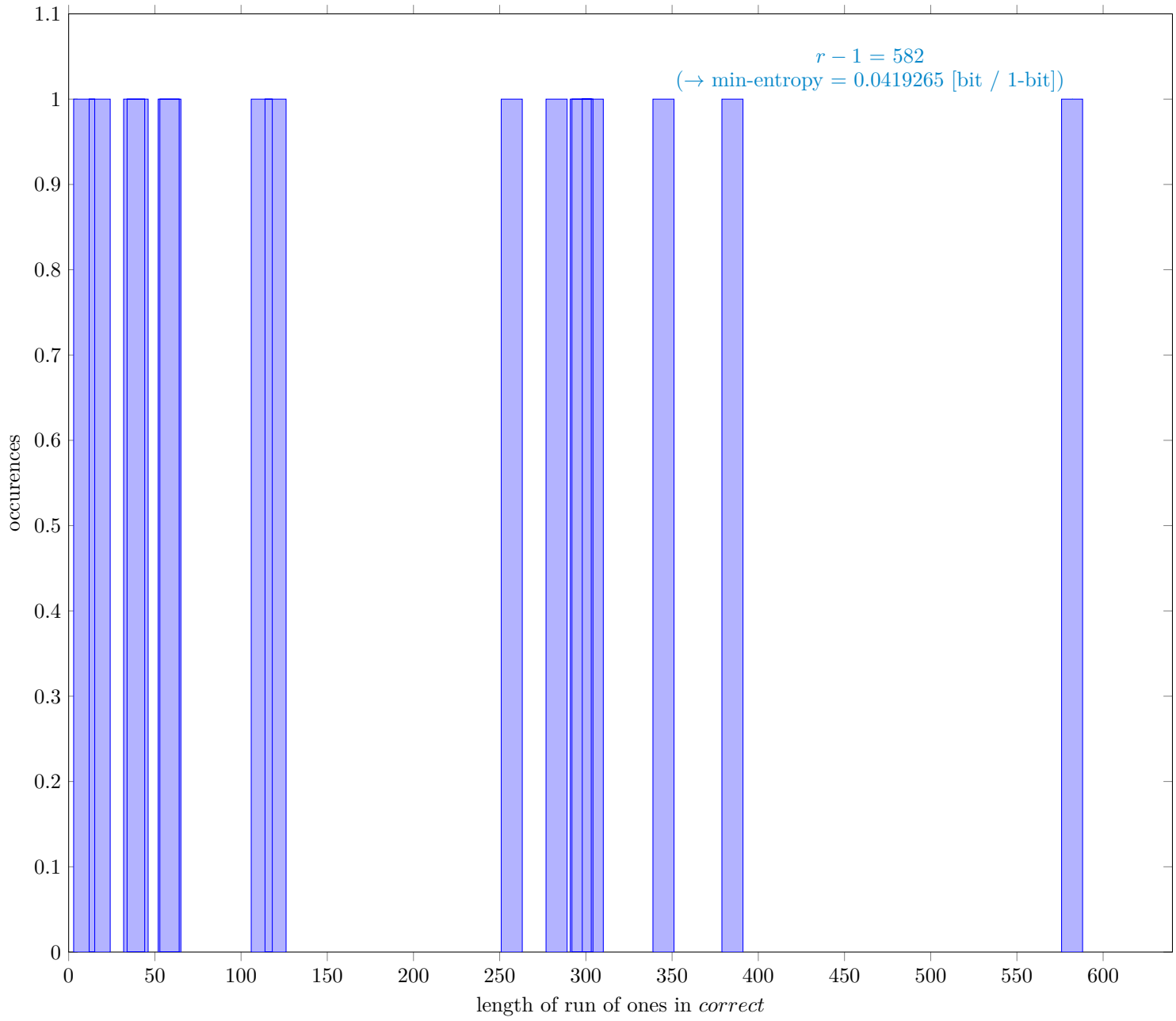

Fig. 22 Distribution of *correct*

4.7.1 Supplemental information for traceability

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

Symbol	Value
N	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)

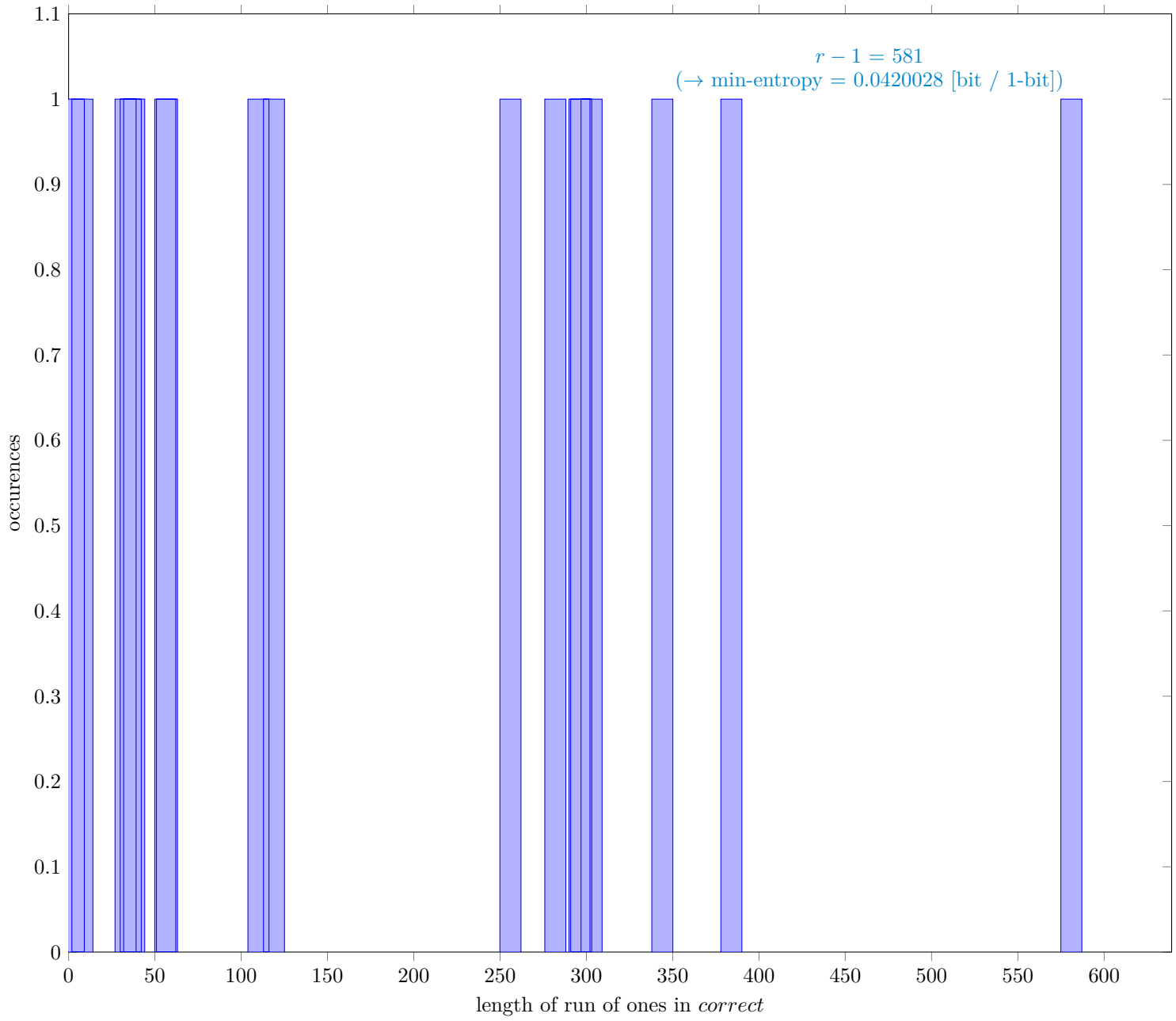

Fig. 23 Distribution of *correct*

4.8.1 Supplemental information for traceability

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

Symbol	Value
N	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)

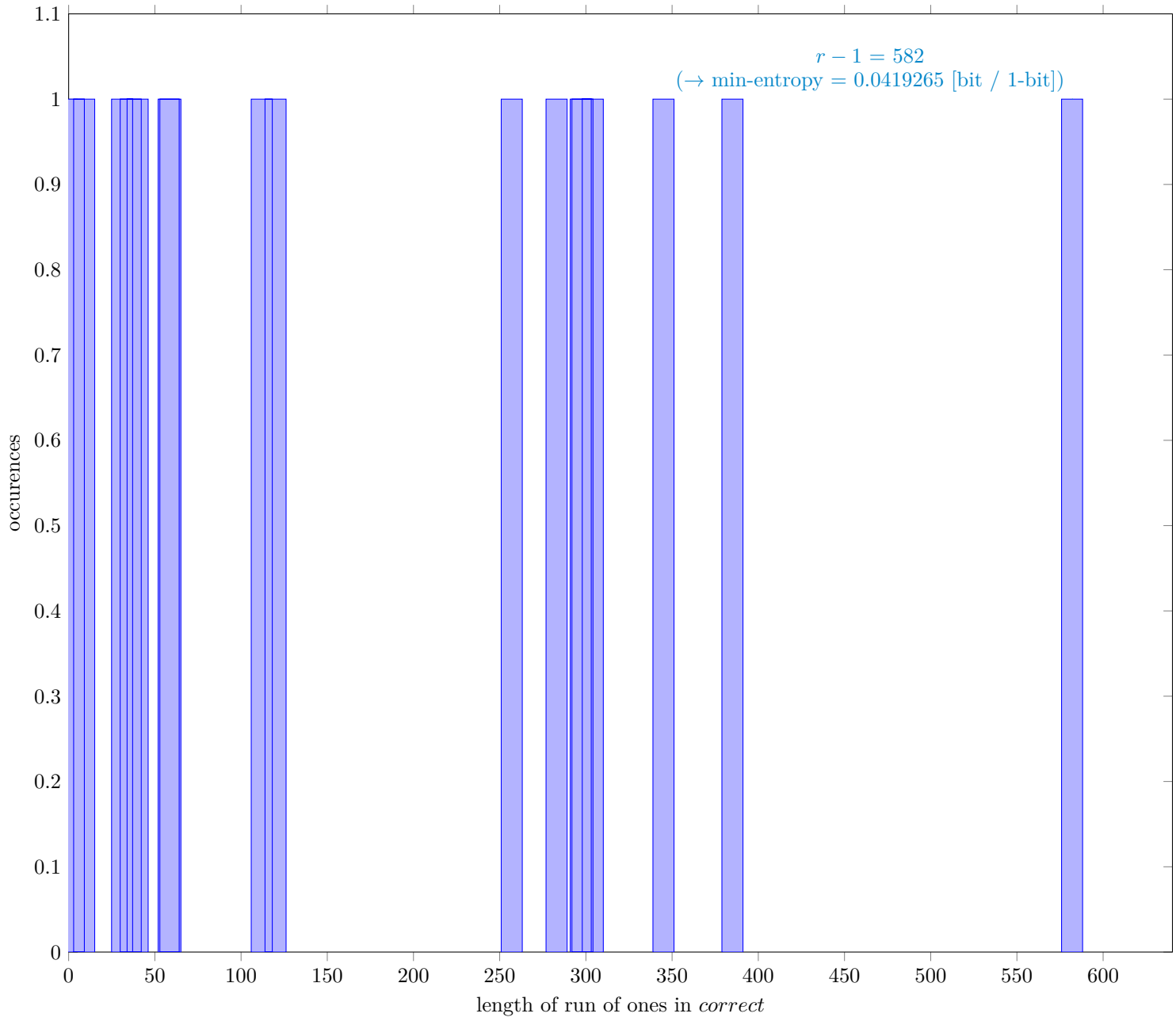

Fig. 24 Distribution of *correct*

4.9.1 Supplemental information for traceability

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

Symbol	Value
N	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)

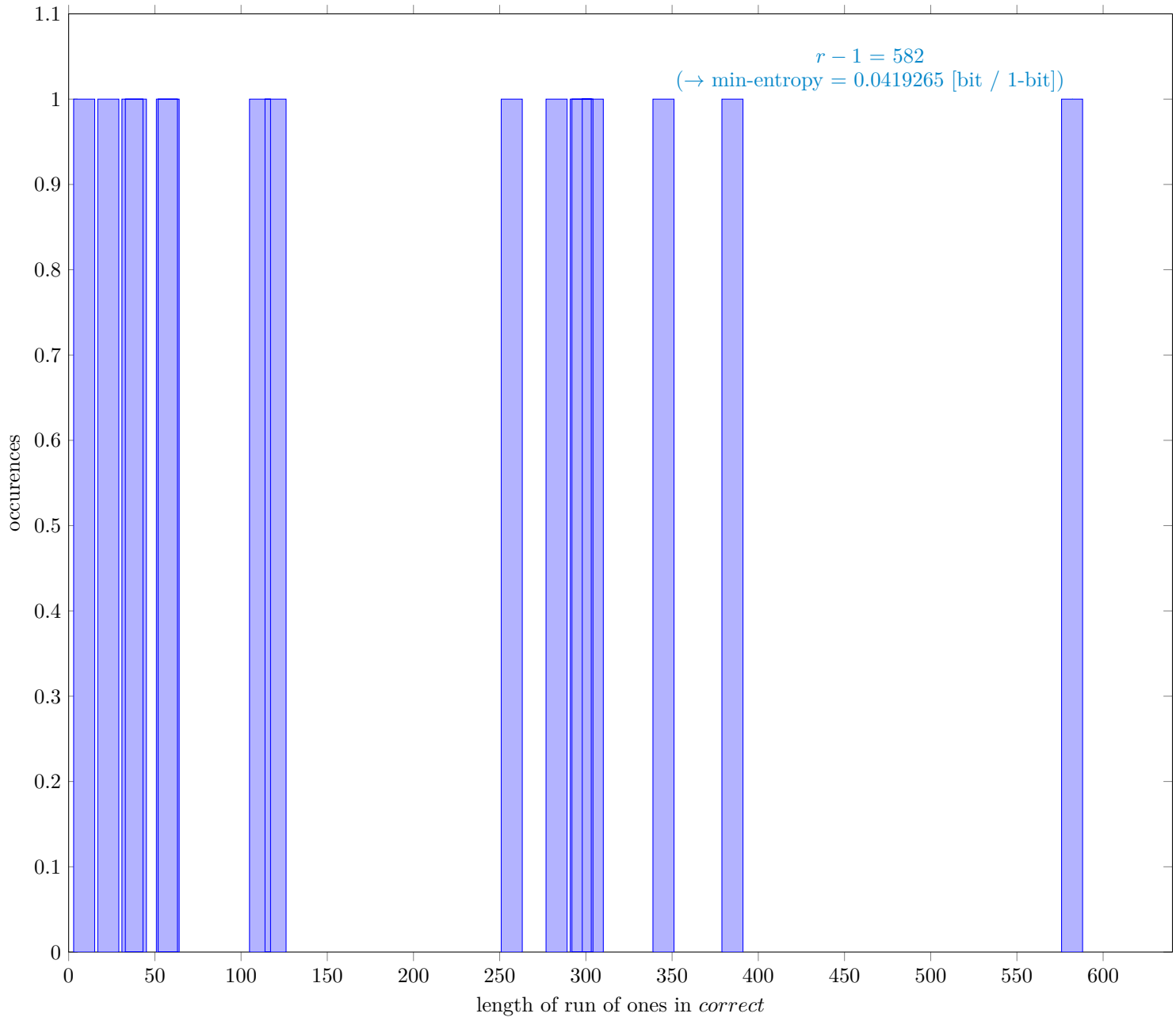

Fig. 25 Distribution of *correct*

4.10.1 Supplemental information for traceability

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

Symbol	Value
N	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 `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`