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

2023-Aug-05 21:00:30.388927

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/rand4_short.bin`
SHA-256 hash value of the acquisition data [hex]	`a9e2169c b1accc78 cd23892d 793a232b 84b0cd13 ccc39235 26e0b207 62bd77ac`

- 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++ 1.82.0
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	10000
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	3.79004	see 3.1	0.979189	see 4.1
The Collision Estimate	—	—	0.89818	see 4.2
The Markov Estimate	—	—	0.990617	see 4.3
The Compression Estimate	—	—	0.803872	see 4.4
The t-Tuple Estimate	3.56747	see 3.2	0.898777	see 4.5
The Longest Repeated Substring (LRS) Estimate	3.83353	see 3.3	0.932969	see 4.6
Multi Most Common in Window Prediction Estimate	3.86695	see 3.4	0.986561	see 4.7
The Lag Prediction Estimate	3.78365	see 3.5	0.982642	see 4.8
The MultiMMC Prediction Estimate	3.88466	see 3.6	0.977697	see 4.9
The LZ78Y Prediction Estimate	3.8825	see 3.7	0.980145	see 4.10

The intial entropy source estimate [bit / 4 - bit] $H_I = \min(H_{\text{original}}, 4 \times H_{\text{bitstring}})$	3.21549

[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

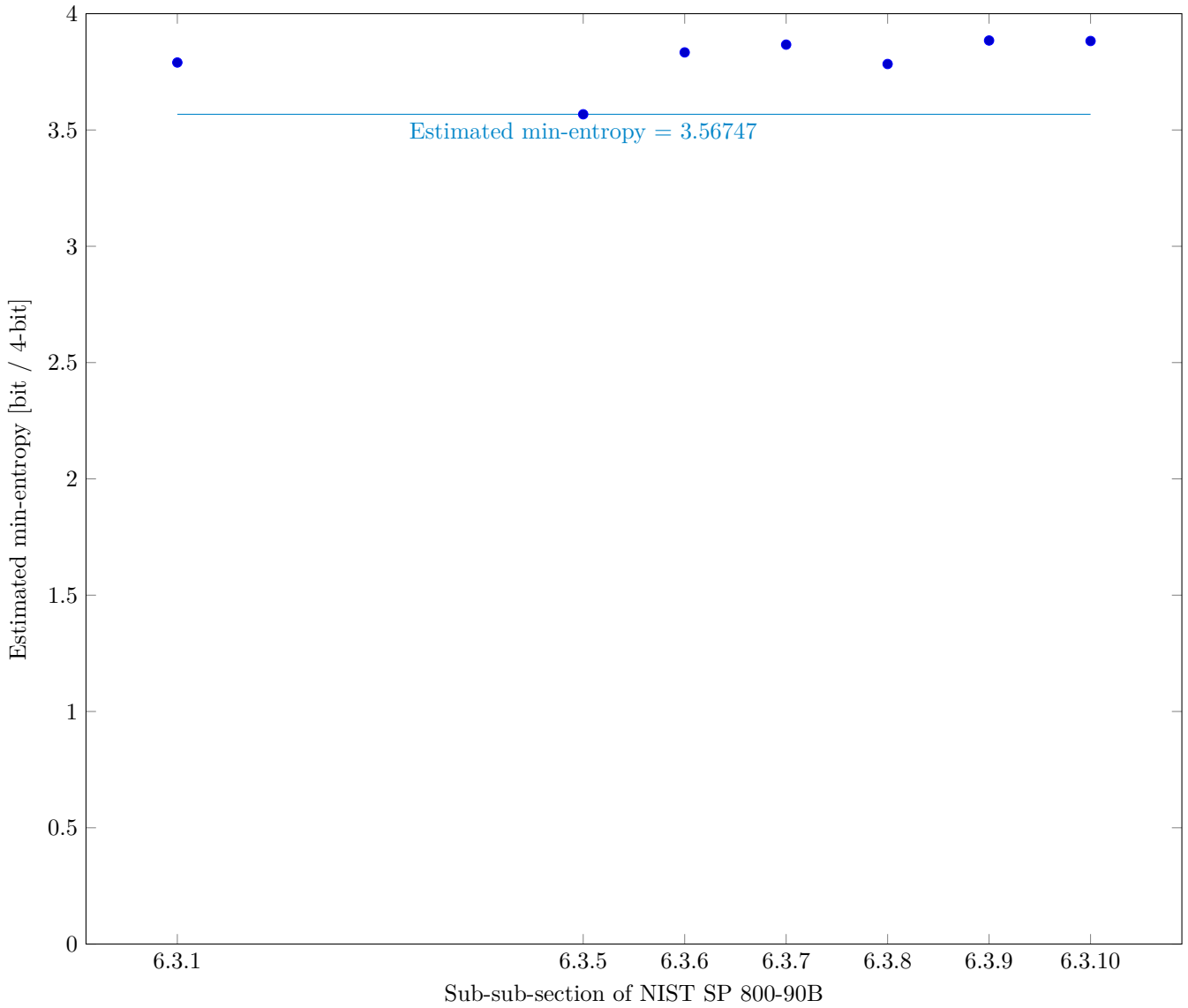

Fig. 1 Estimated Min-Entropy using §6.3 of NIST SP 800-90B

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

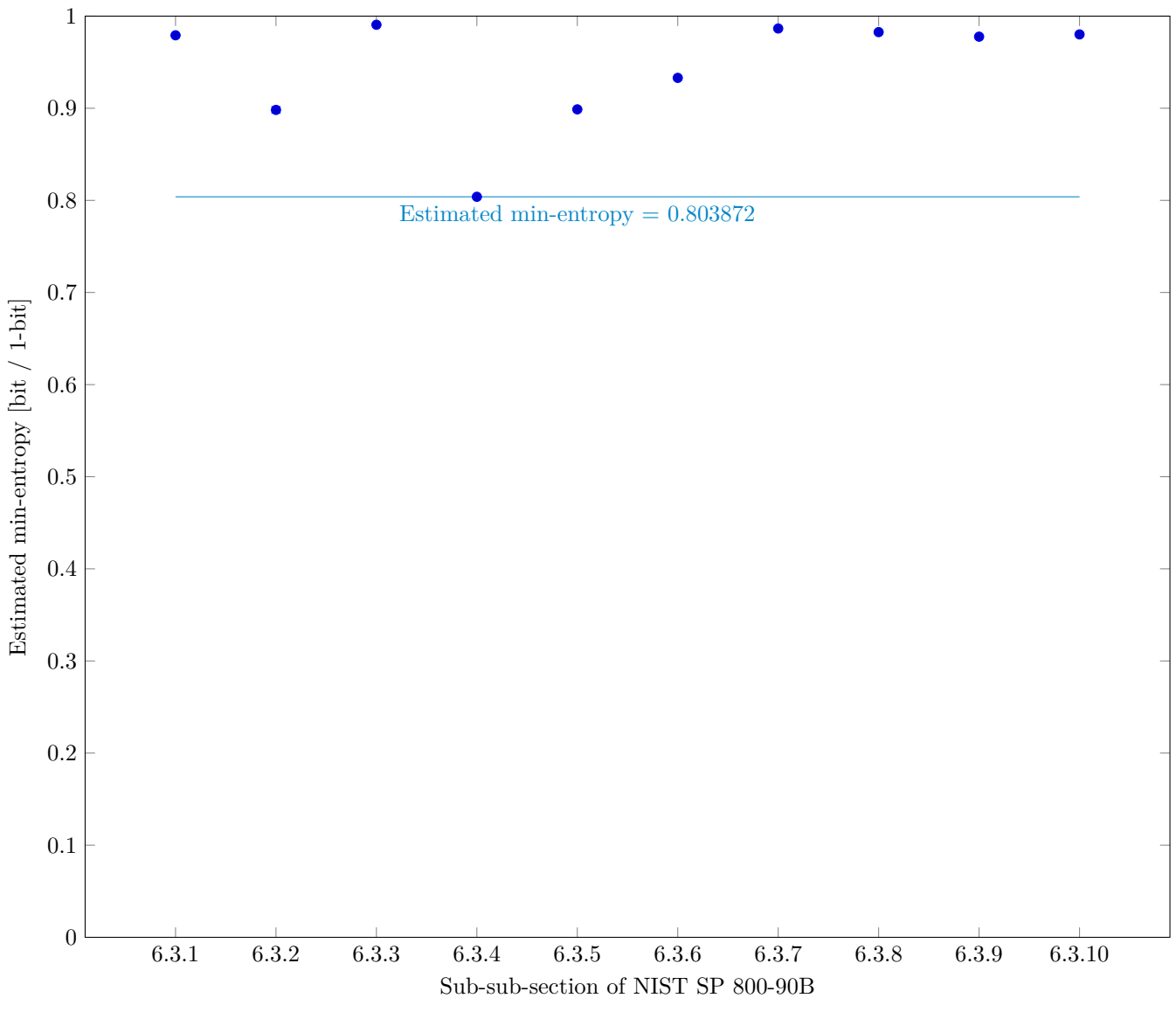

Fig. 2 Estimated Min-Entropy using §6.3 of NIST SP 800-90B

3 Detailed results of analysis from original samples

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

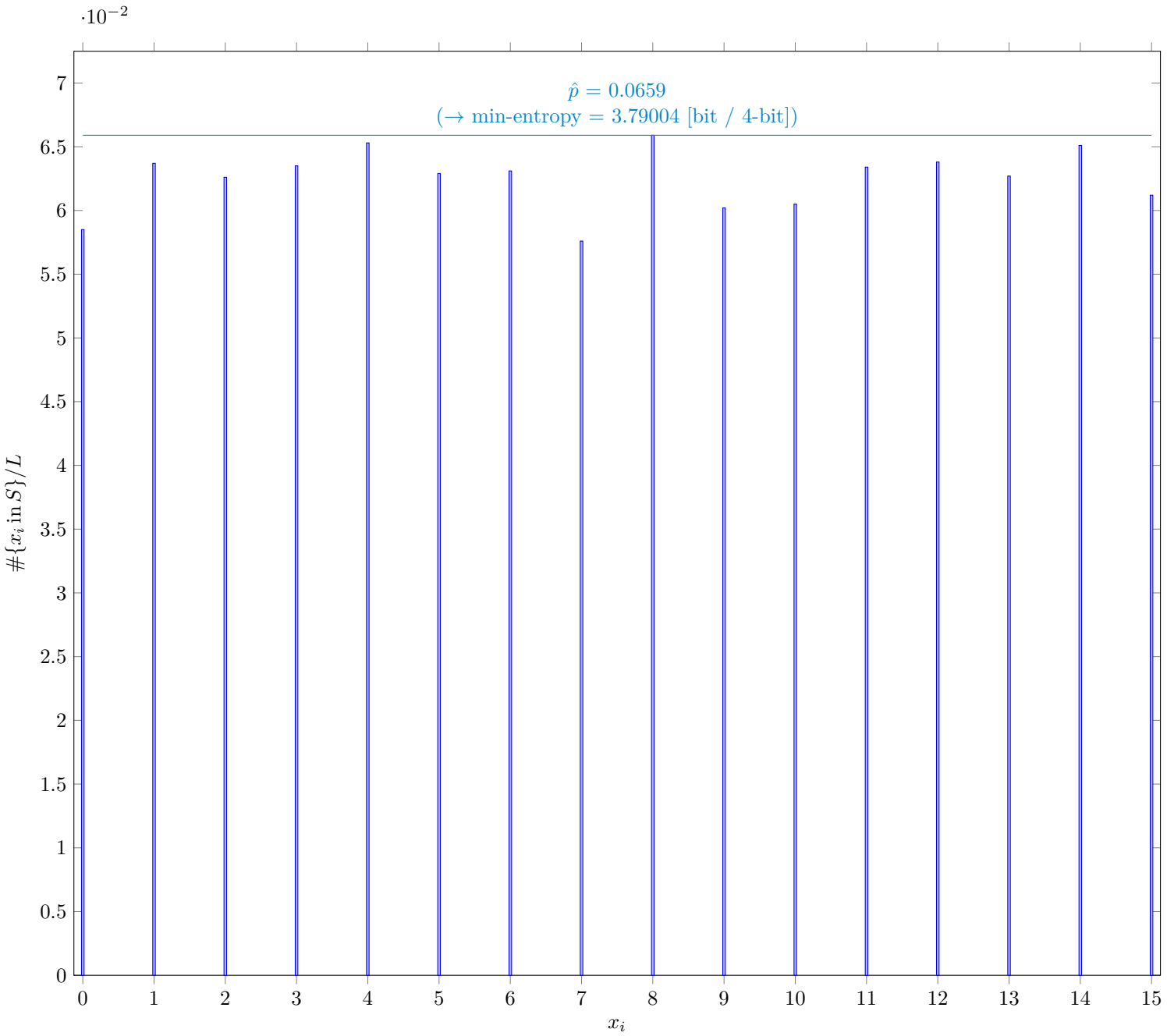

Fig. 3 Distribution of x_i

3.1.1 Supplemental information for traceability

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

Symbol	Value
mode	659
$\hat{p}$	0.0659
p_u	0.0722911

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

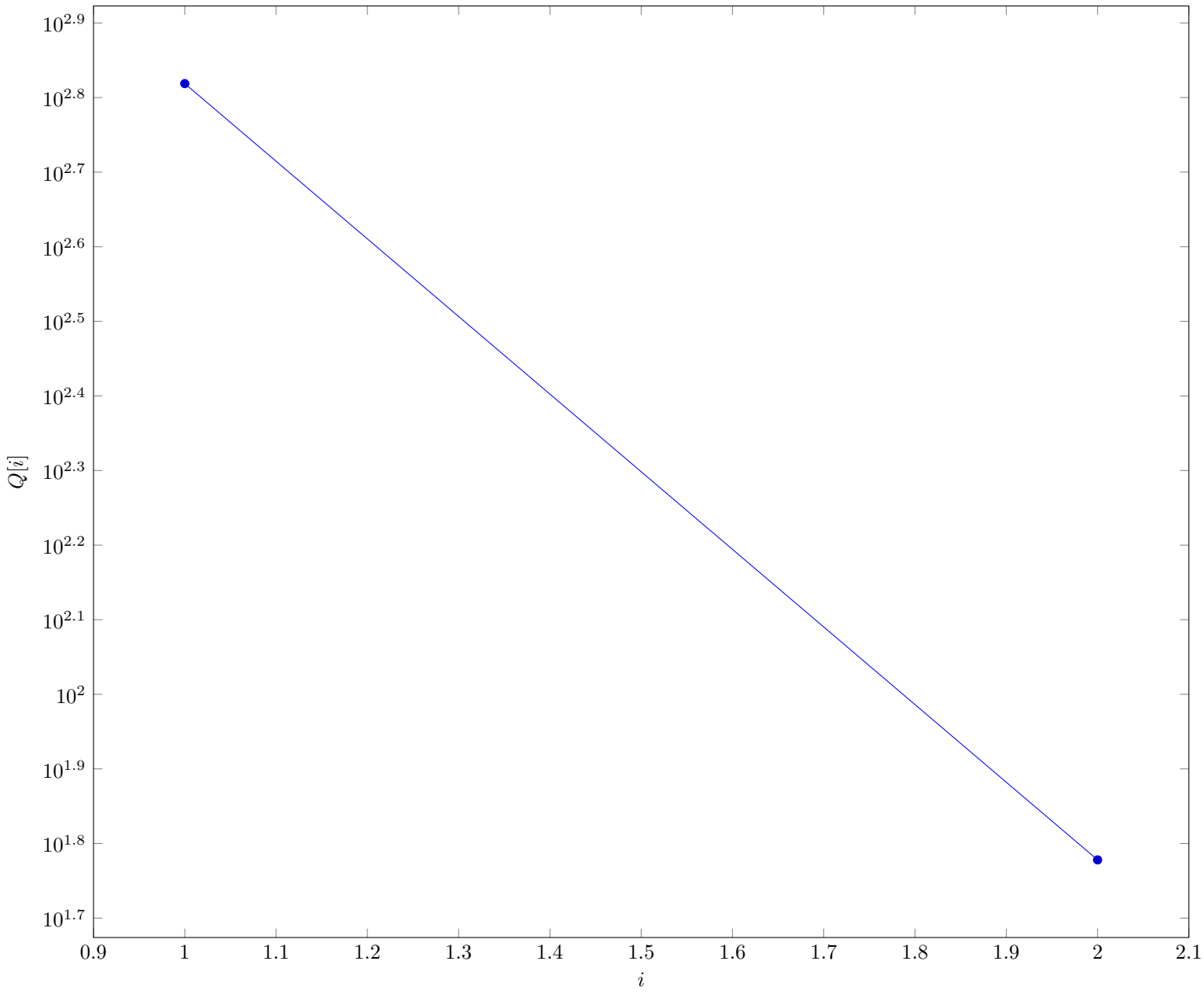

Fig. 4 Intermediate value $Q[i]$ in §6.3.5 of NIST SP 800-90B

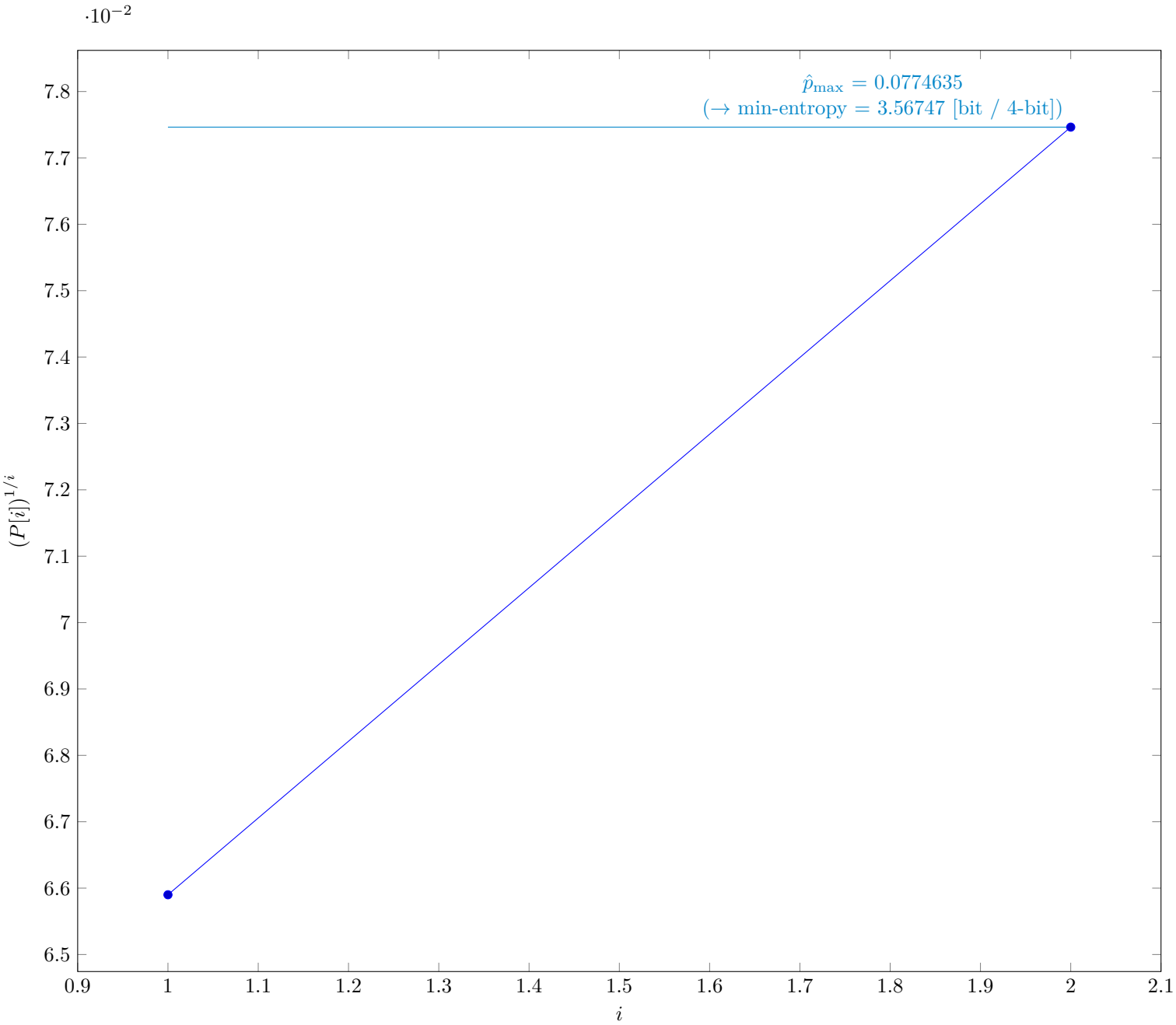

Fig. 5 $P[i]^{1/i}$ in §6.3.5 of NIST SP 800-90B

3.2.1 Supplemental information for traceability

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

Symbol	Value
t	2
$\hat{p}_{\max}$	0.0774635
p_u	0.0843497

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

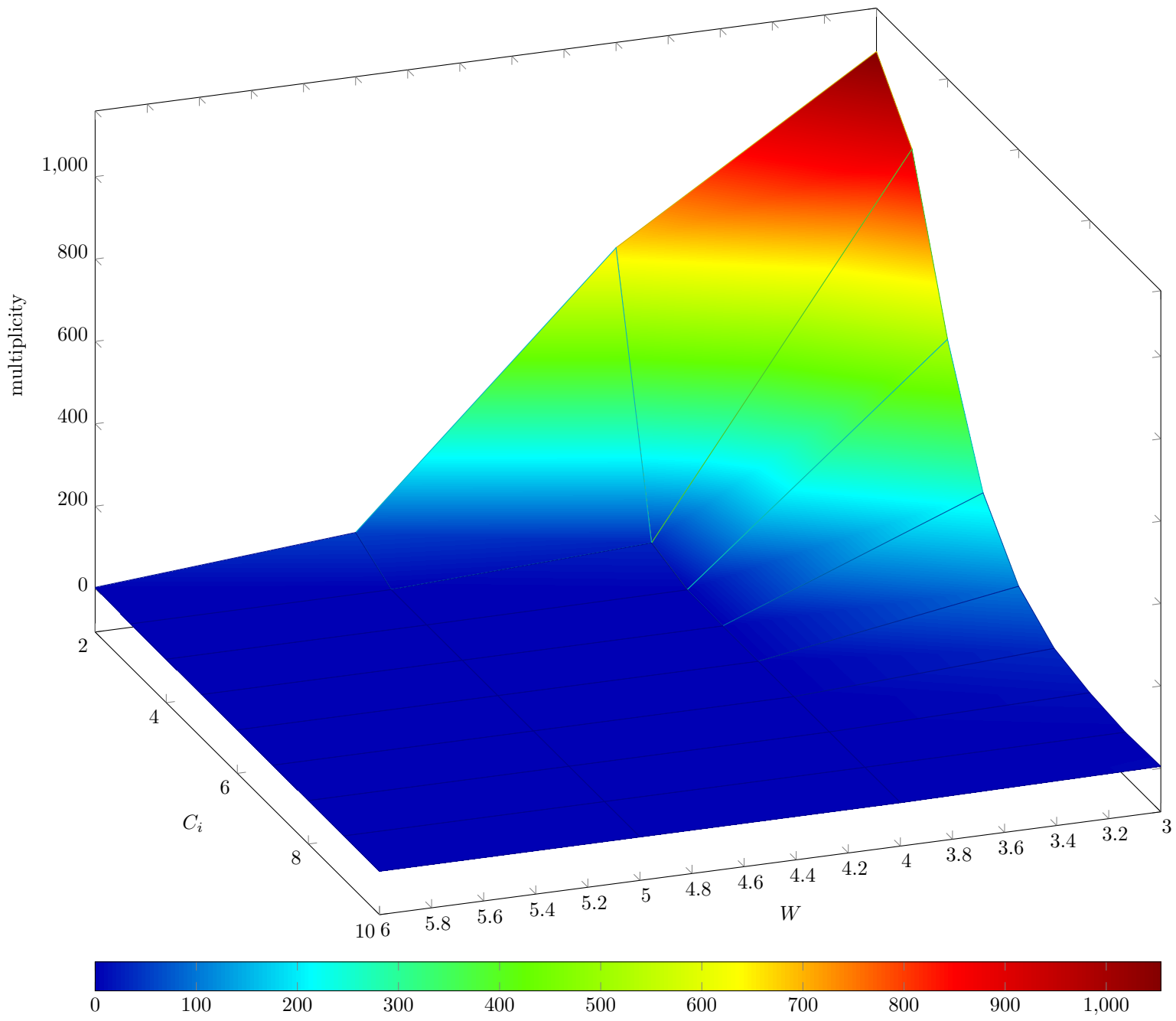

Fig. 6 Estimated W-tuple collision probability in Step 3 of §6.3.6 of NIST SP 800-90B

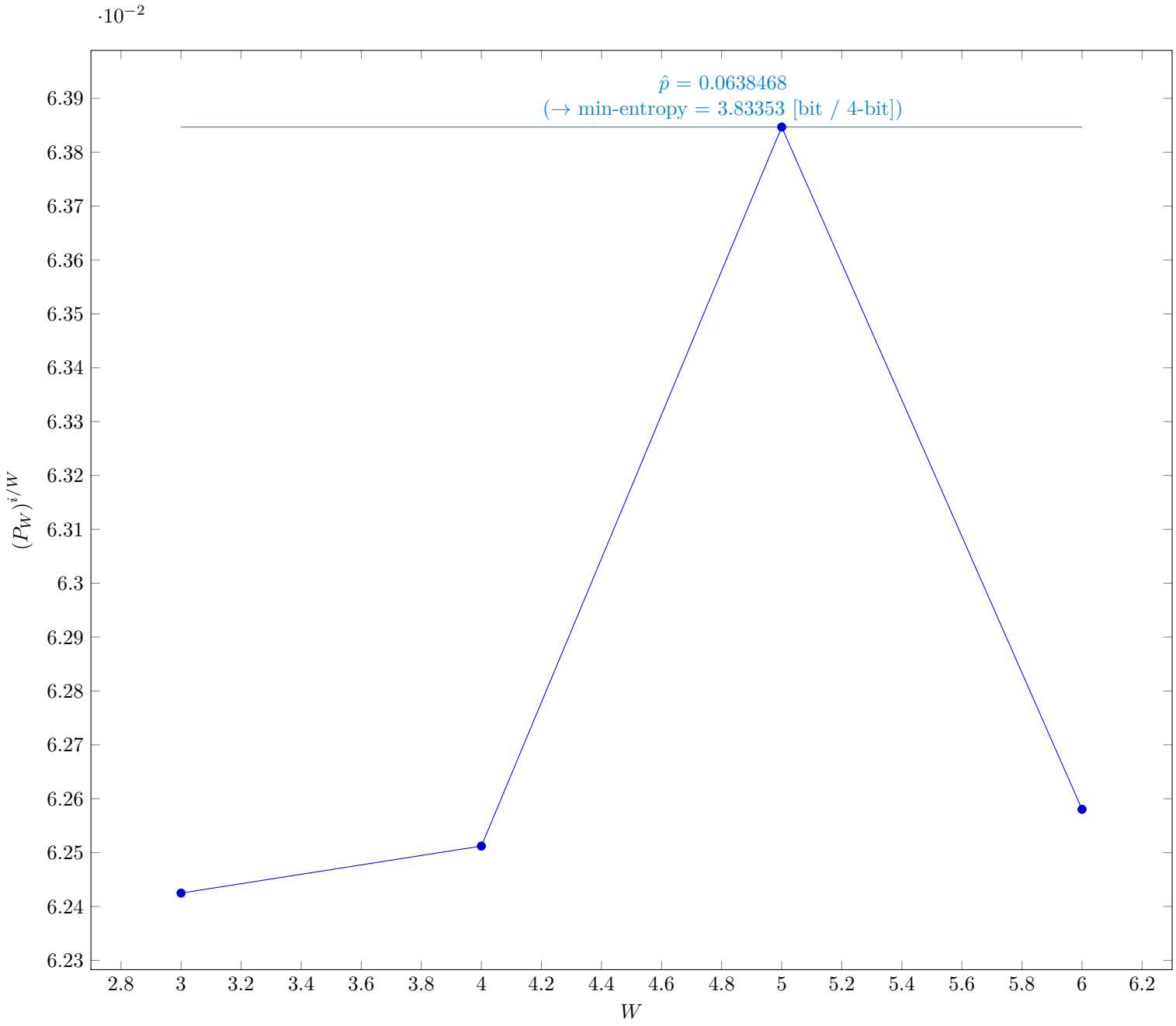

Fig. 7 Estimated average collision probability per string symbol in Step 3 of §6.3.6 of NIST SP 800-90B

3.3.1 Supplemental information for traceability

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

Symbol	Value
u	3
v	6
$\hat{p}$	0.0638468
p_u	0.0701445

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

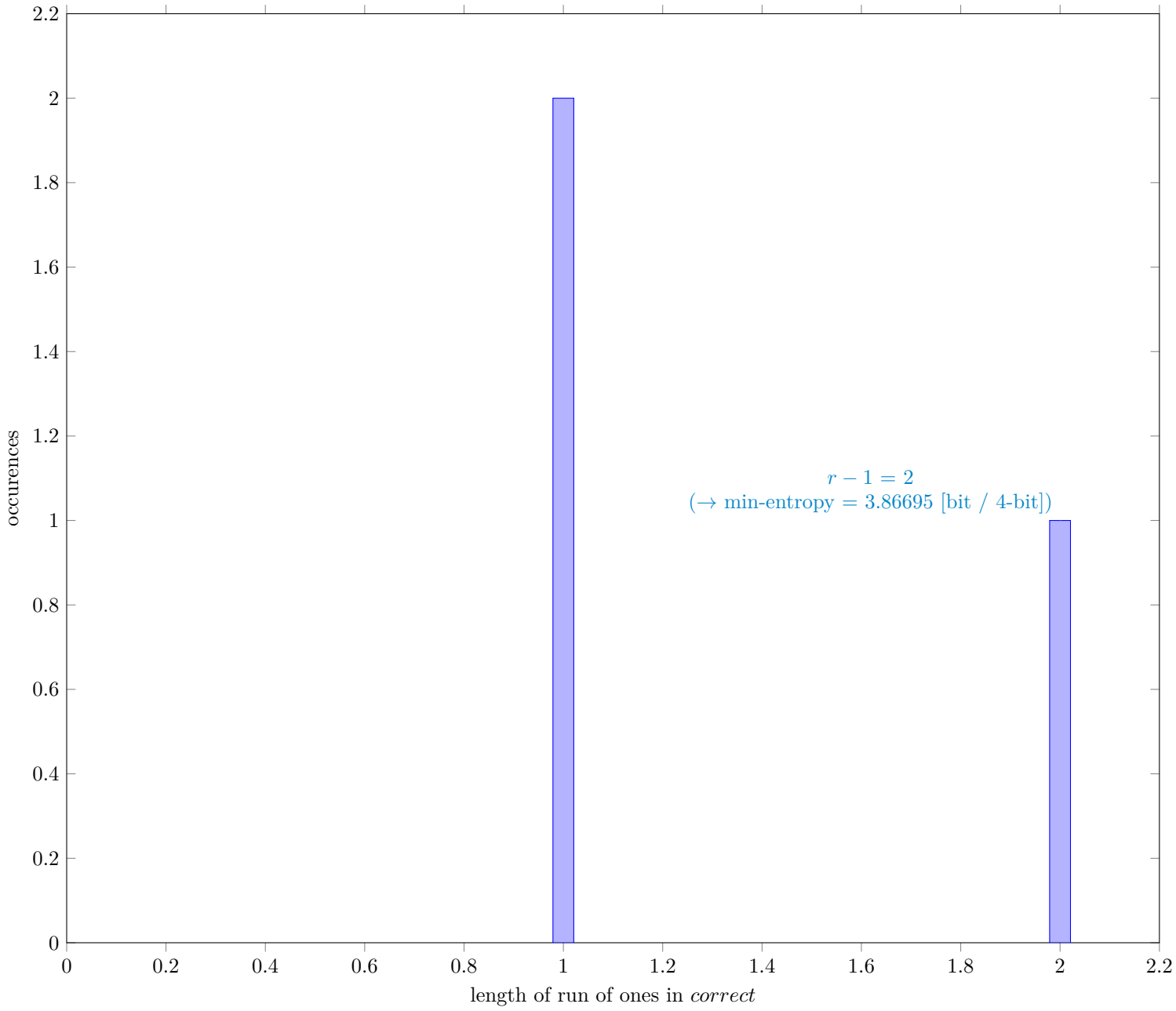

Fig. 8 Distribution of *correct*

3.4.1 Supplemental information for traceability

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

Symbol	Value
N	9937
C	619
P_{global}	0.0622924
P'_{global}	0.0685379
r	3
P_{local}	0.0100725

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

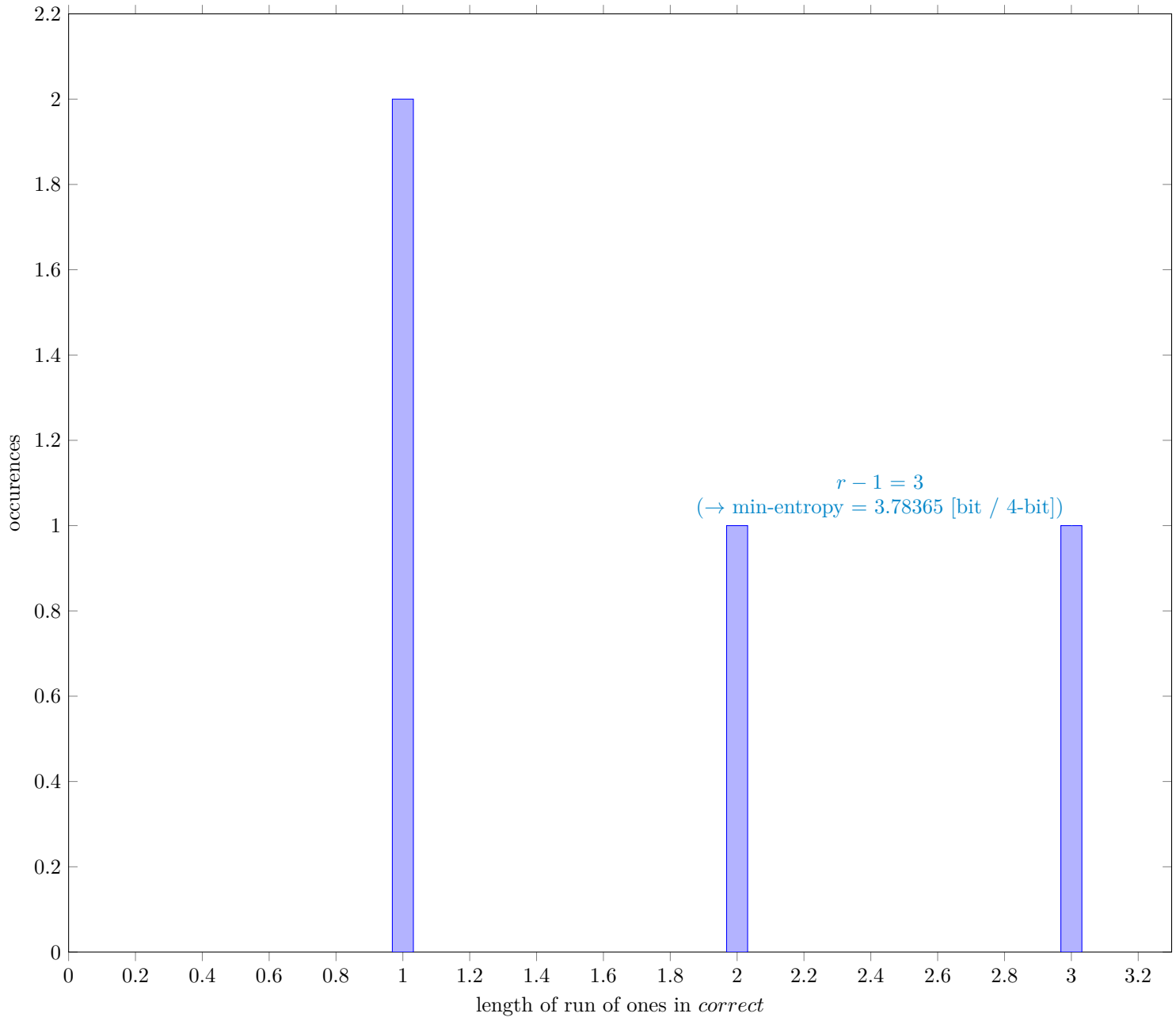

Fig. 9 Distribution of *correct*

3.5.1 Supplemental information for traceability

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

Symbol	Value
N	9999
C	662
P_{global}	0.0662066
P'_{global}	0.0726119
r	4
P_{local}	0.0319235

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

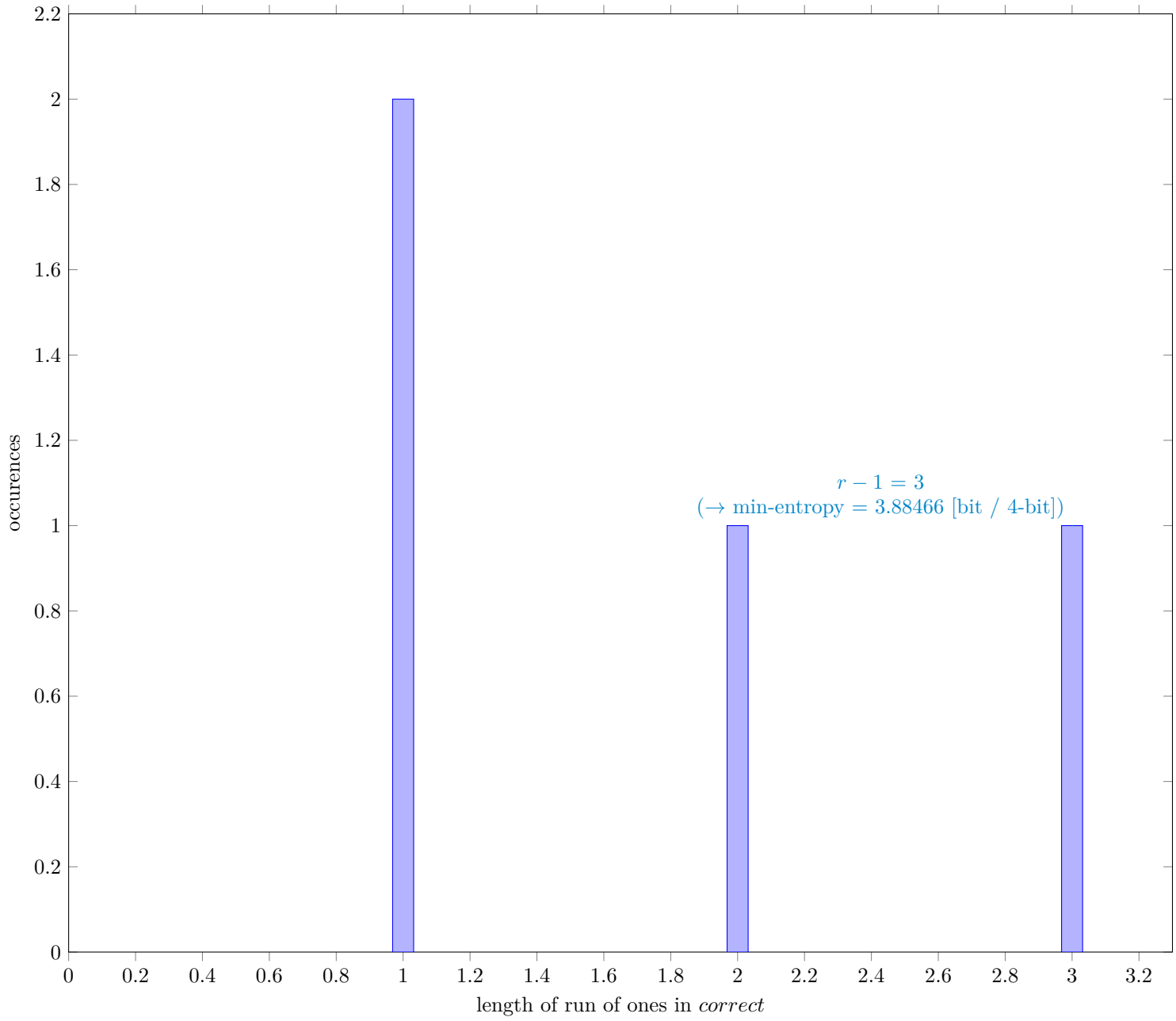

Fig. 10 Distribution of *correct*

3.6.1 Supplemental information for traceability

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

Symbol	Value
N	9998
C	615
P_{global}	0.0615123
P'_{global}	0.0677021
r	4
P_{local}	0.0319243

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

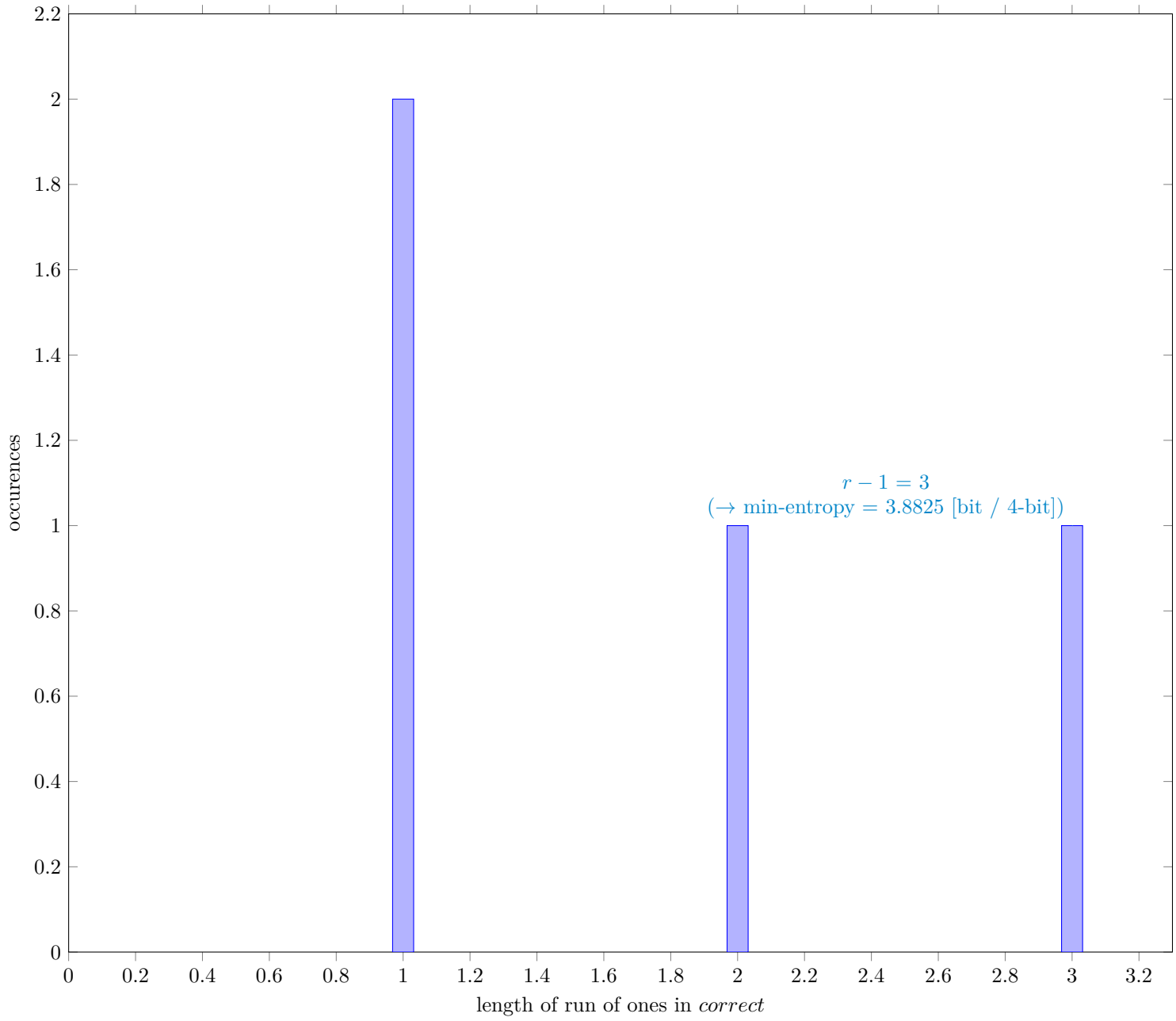

Fig. 11 Distribution of *correct*

3.7.1 Supplemental information for traceability

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

Symbol	Value
N	9983
C	615
P_{global}	0.0616047
P'_{global}	0.0678035
r	4
P_{local}	0.0319364

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)

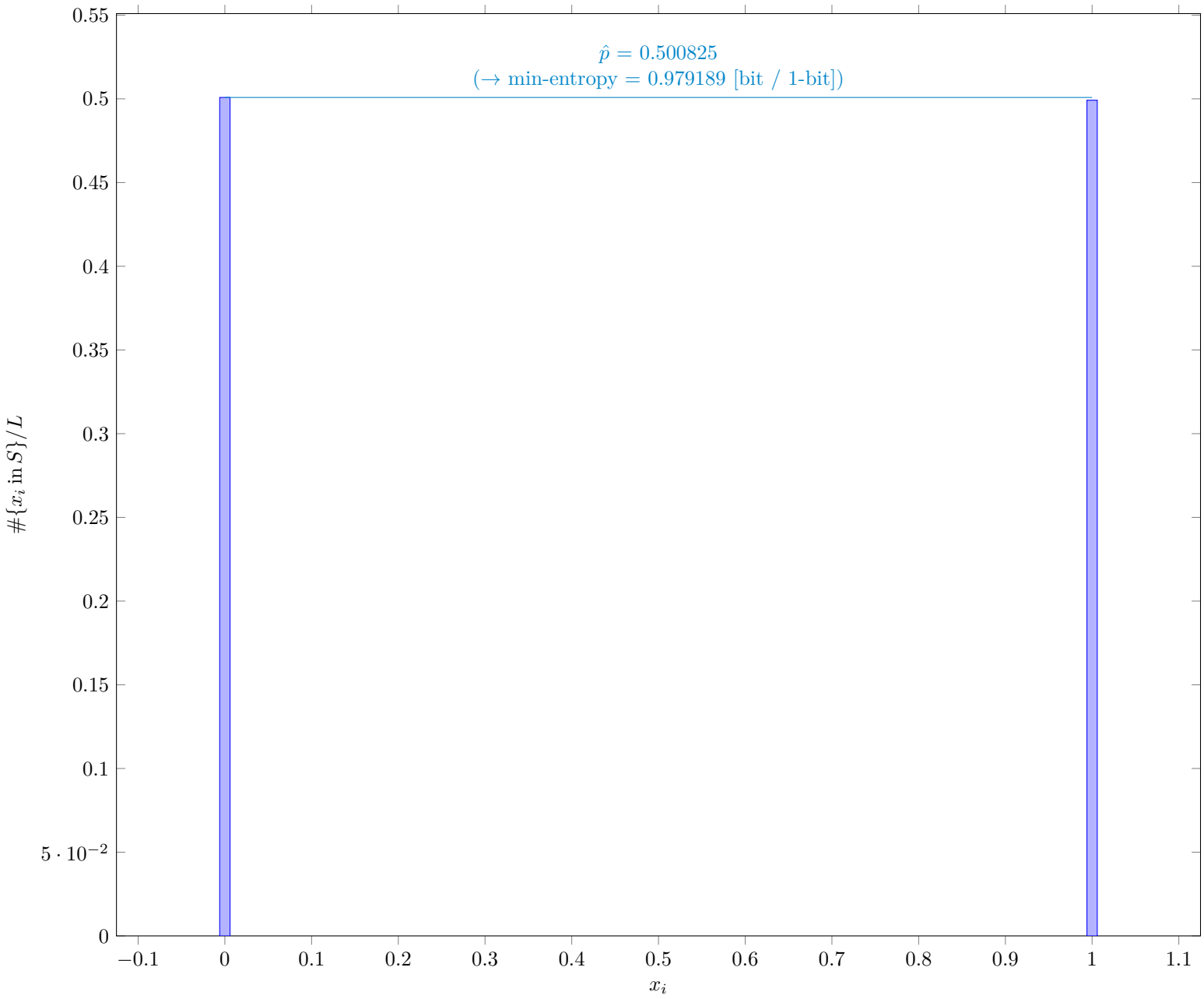

Fig. 12 Distribution of x_i

4.1.1 Supplemental information for traceability

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

Symbol	Value
mode	20033
$\hat{p}$	0.500825
p_u	0.507265

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

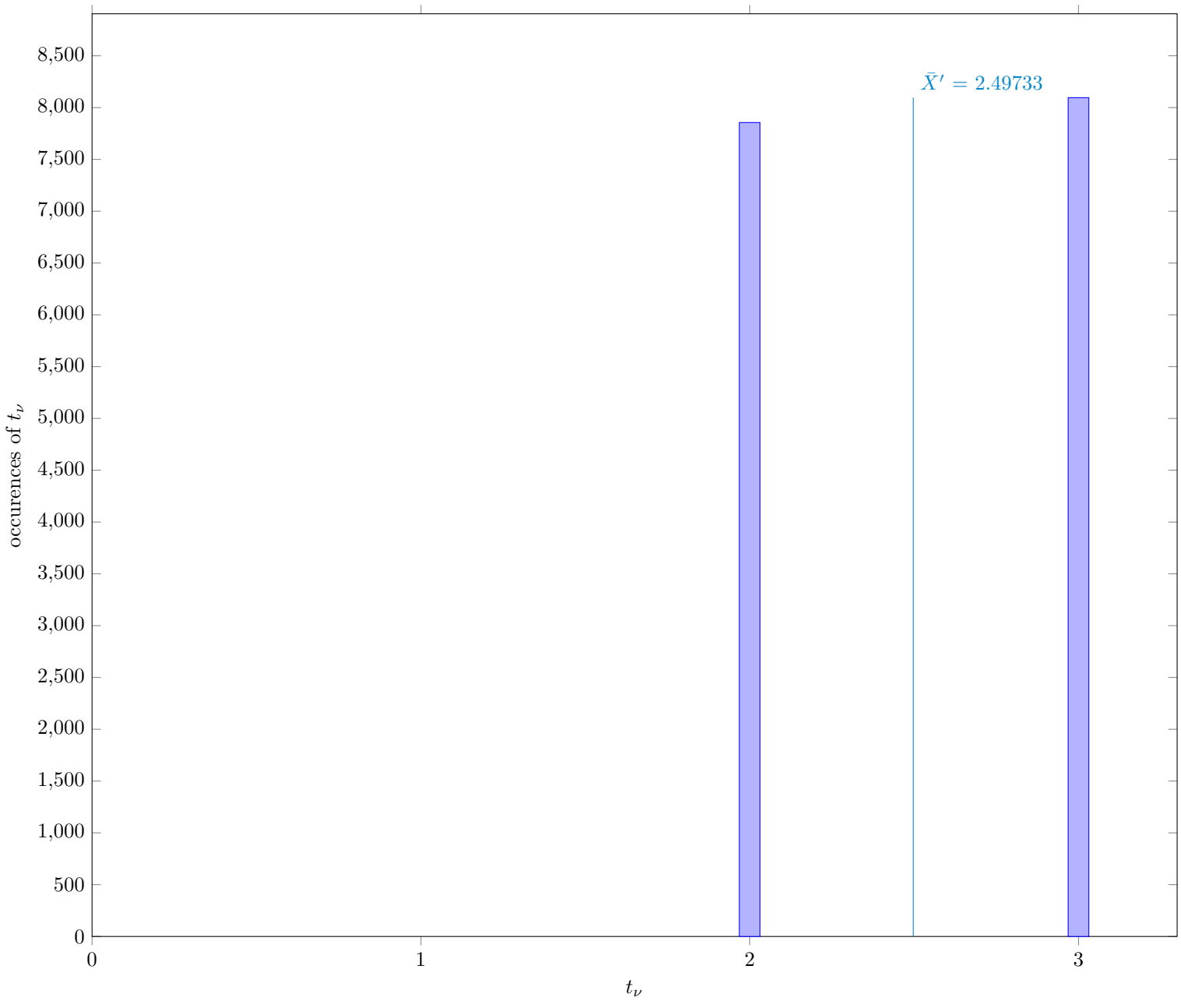

Fig. 13 Distribution of intermediate value t_ν

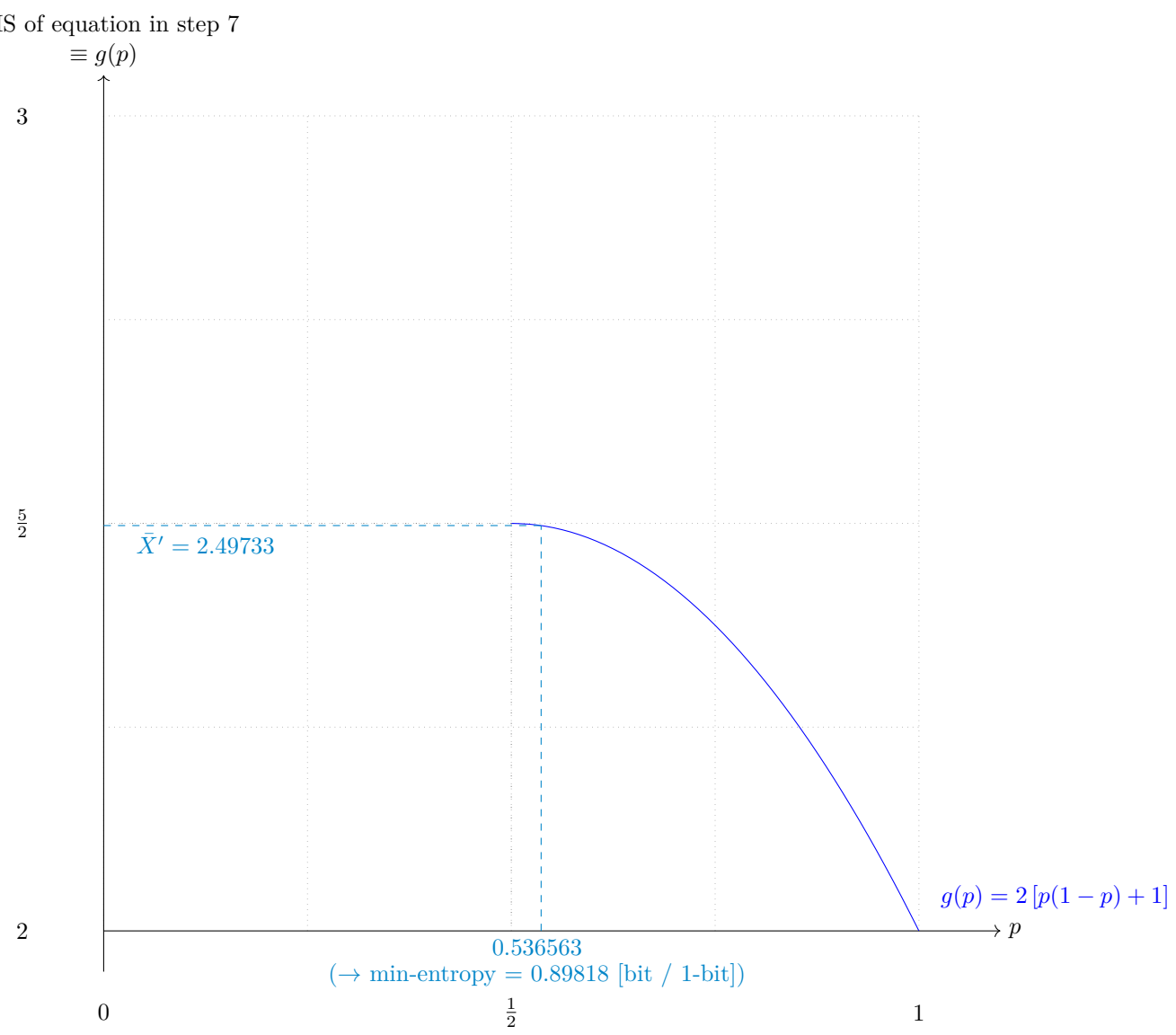

Fig. 14 Solution to the equation in step 7

4.2.1 Supplemental information for traceability

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

Symbol	Value
p	0.536563
$\bar{X}$	2.50752
$\bar{X}'$	2.49733
$\hat{\sigma}$	0.499959

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

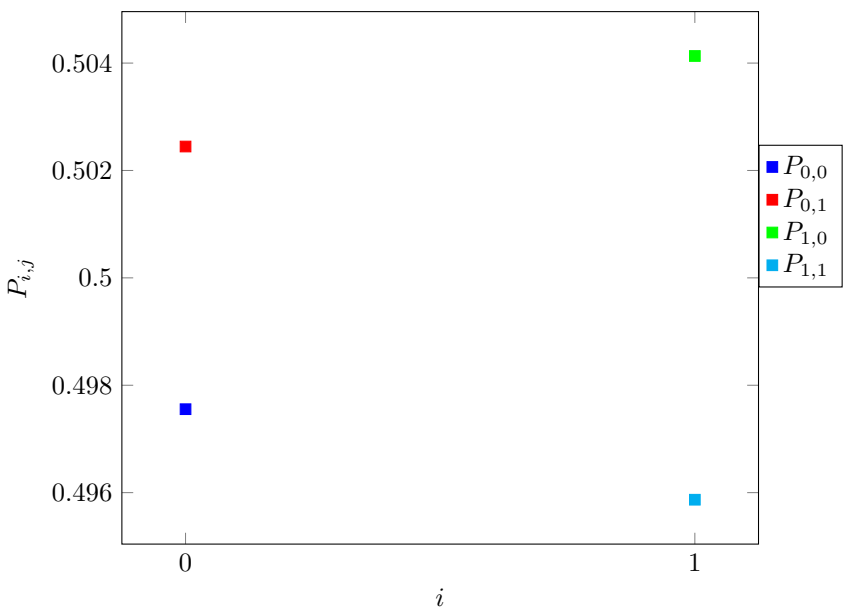

Fig. 15 Transition probability $P_{i,j}$ of §6.3.3 of NIST SP 800-90B

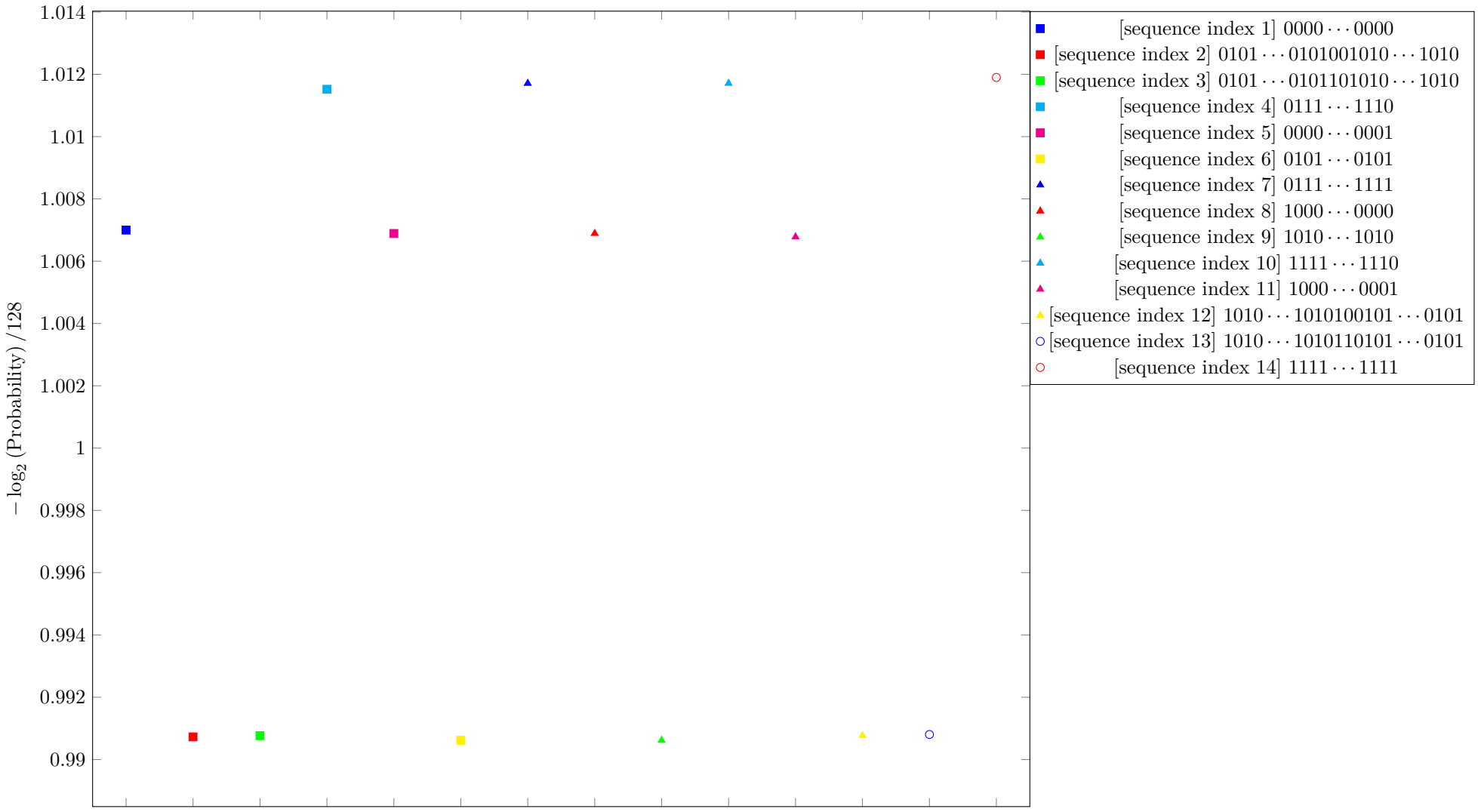

Fig. 16 Estimated Min-Entropy using §6.3.3 of NIST SP 800-90B

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

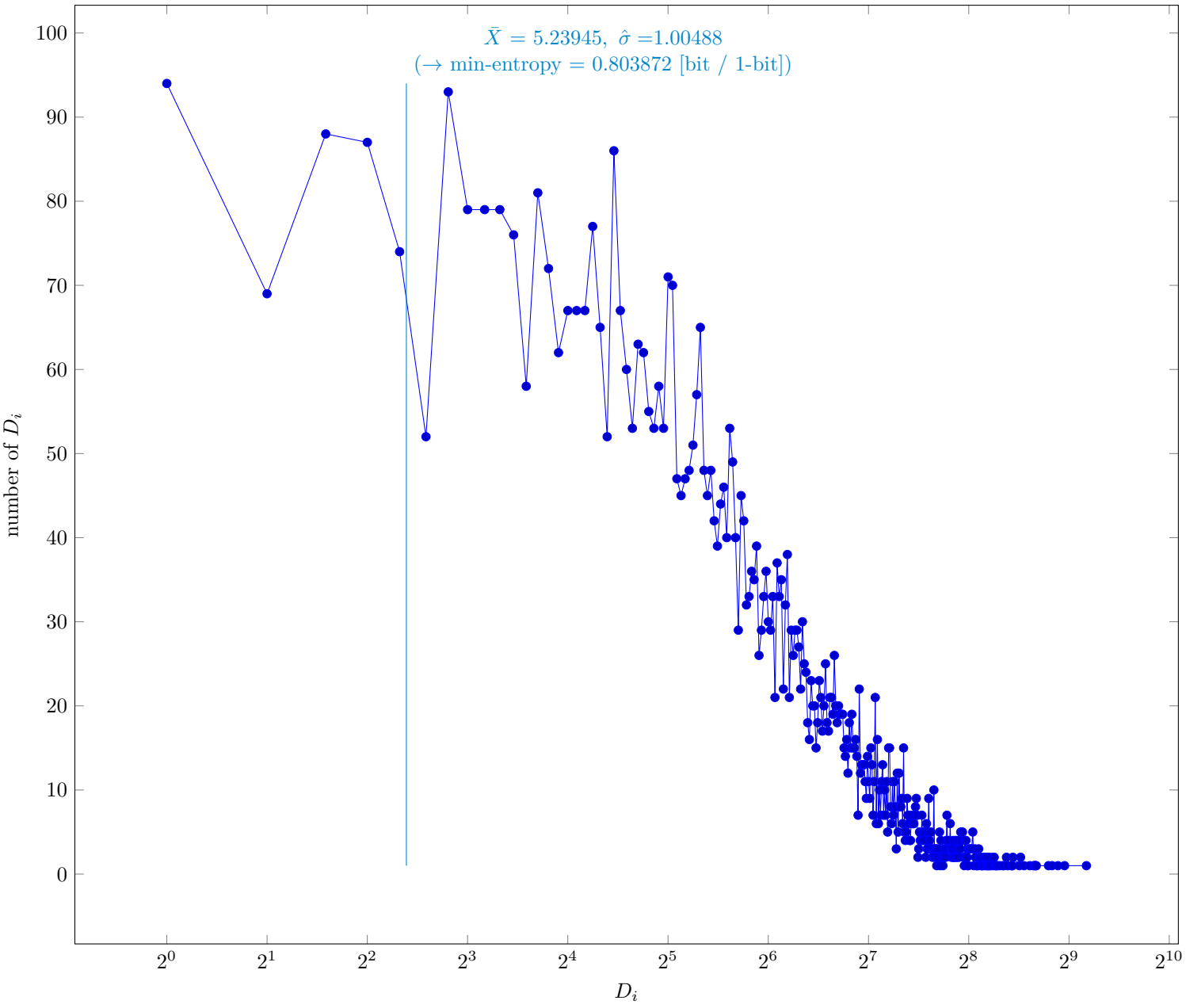

Fig. 17 Distribution of intermediate value D_i

4.4.1 Supplemental information for traceability

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

Symbol	Value
p	0.0353234
$\bar{X}$	5.23945
$\hat{\sigma}$	1.00488
$\bar{X}'$	5.20506

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

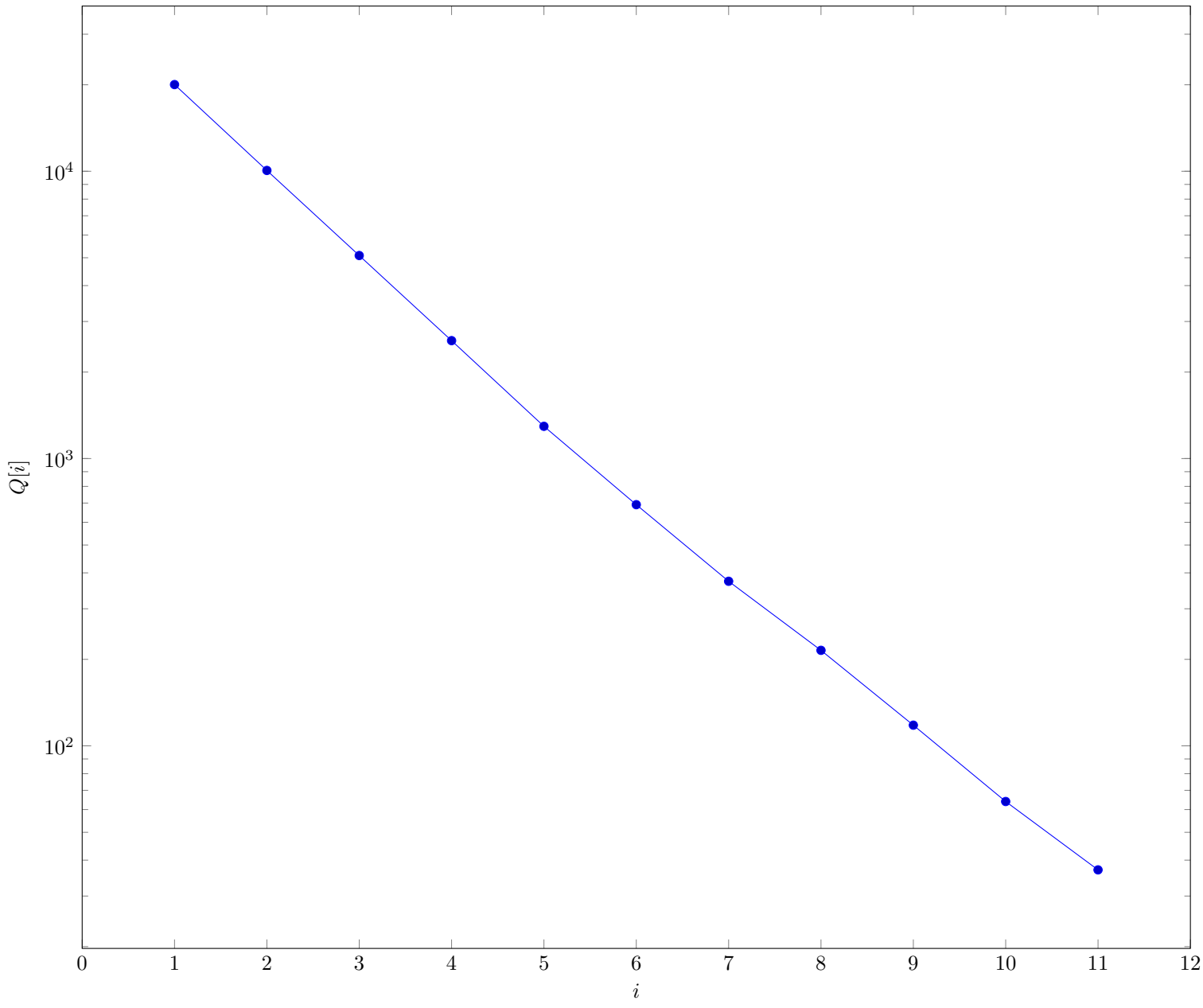

Fig. 18 Intermediate value $Q[i]$ in §6.3.5 of NIST SP 800-90B

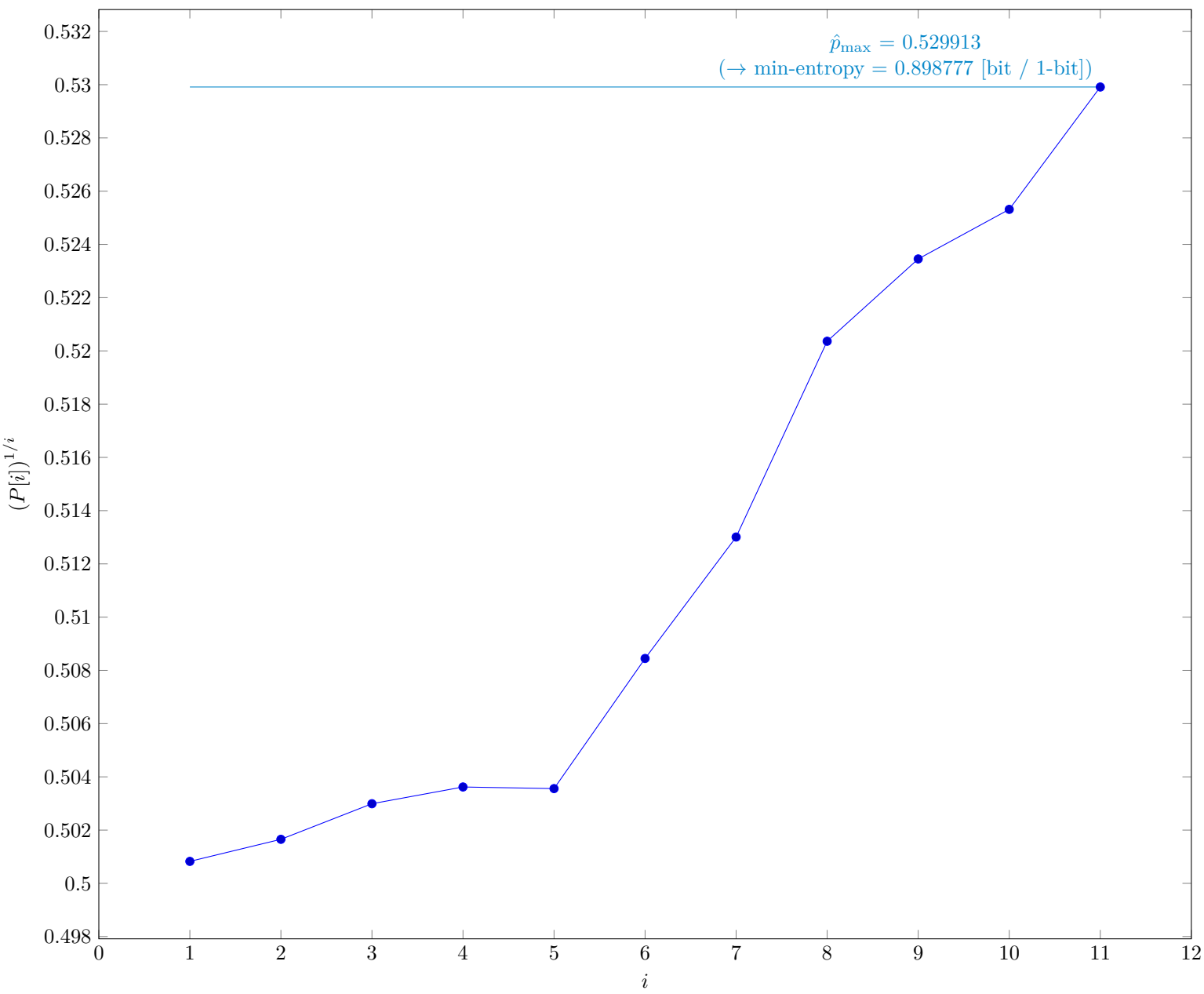

Fig. 19 $P[i]^{1/i}$ in §6.3.5 of NIST SP 800-90B

4.5.1 Supplemental information for traceability

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

Symbol	Value
t	11
$\hat{p}_{\max}$	0.529913
p_u	0.536341

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

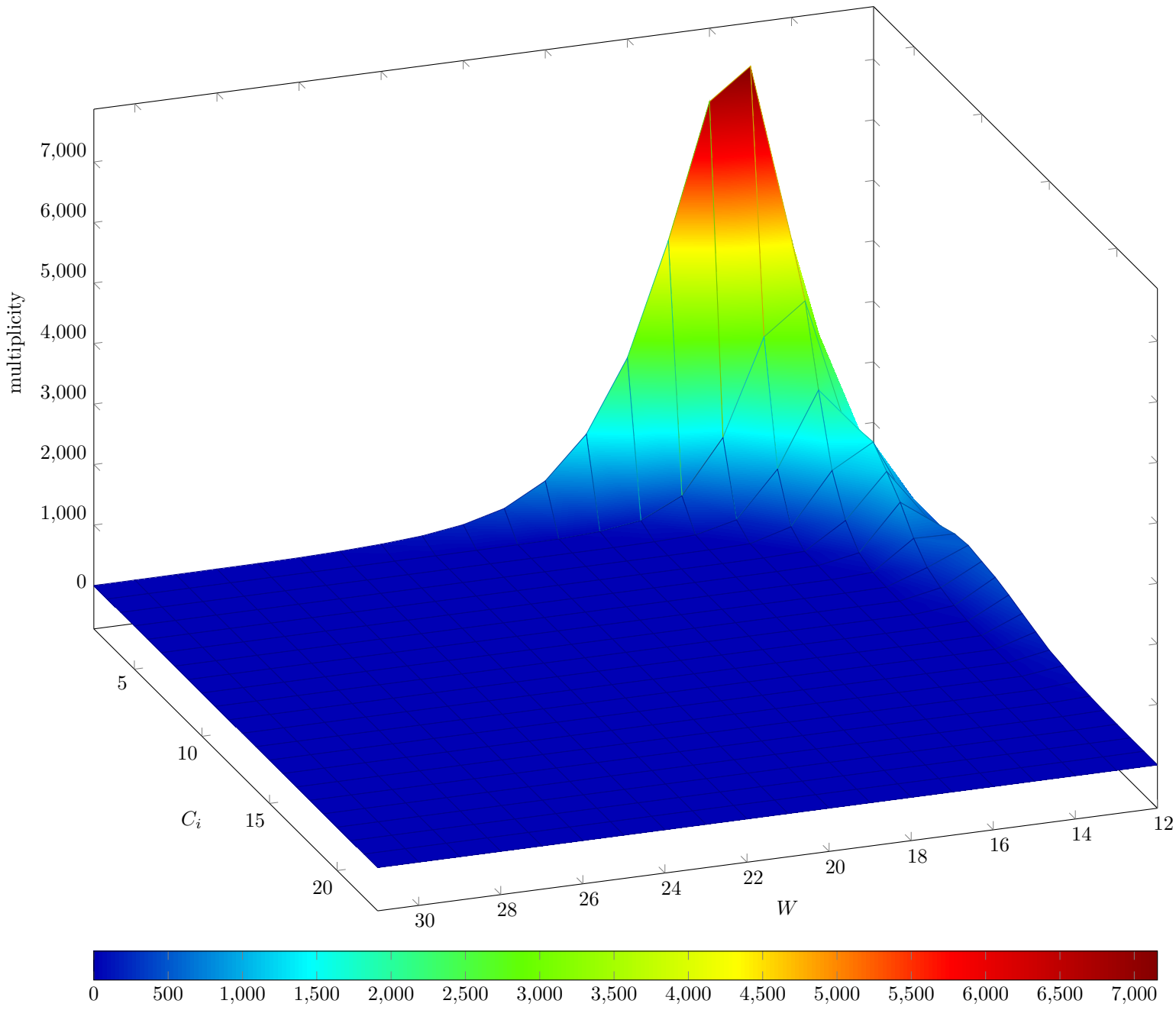

Fig. 20 Estimated W-tuple collision probability in Step 3 of §6.3.6 of NIST SP 800-90B

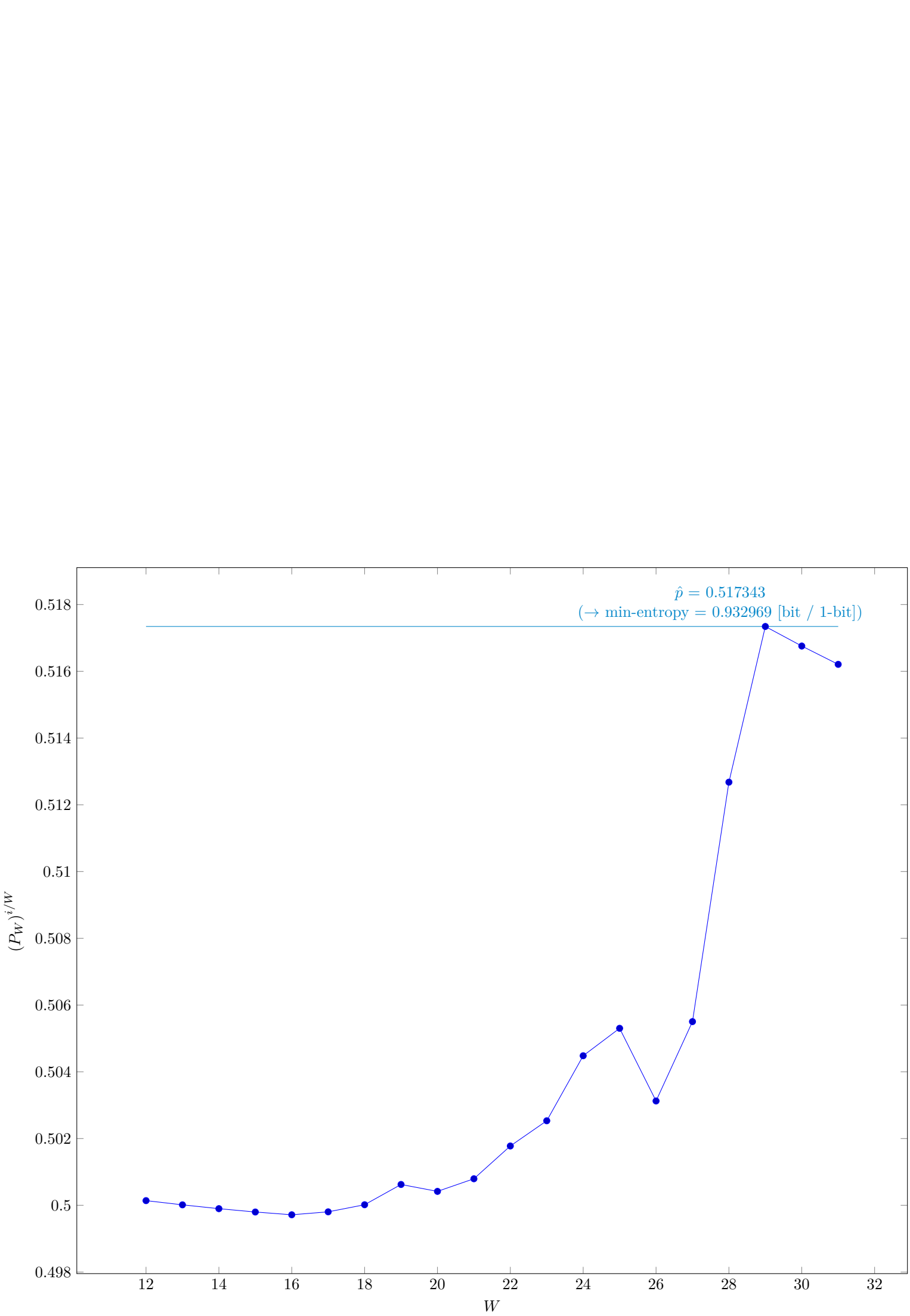

Fig. 21 Estimated average collision probability per string symbol in Step 3 of §6.3.6 of NIST SP 800-90B

4.6.1 Supplemental information for traceability

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

Symbol	Value
u	12
v	31
$\hat{p}$	0.517343
p_u	0.523779

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

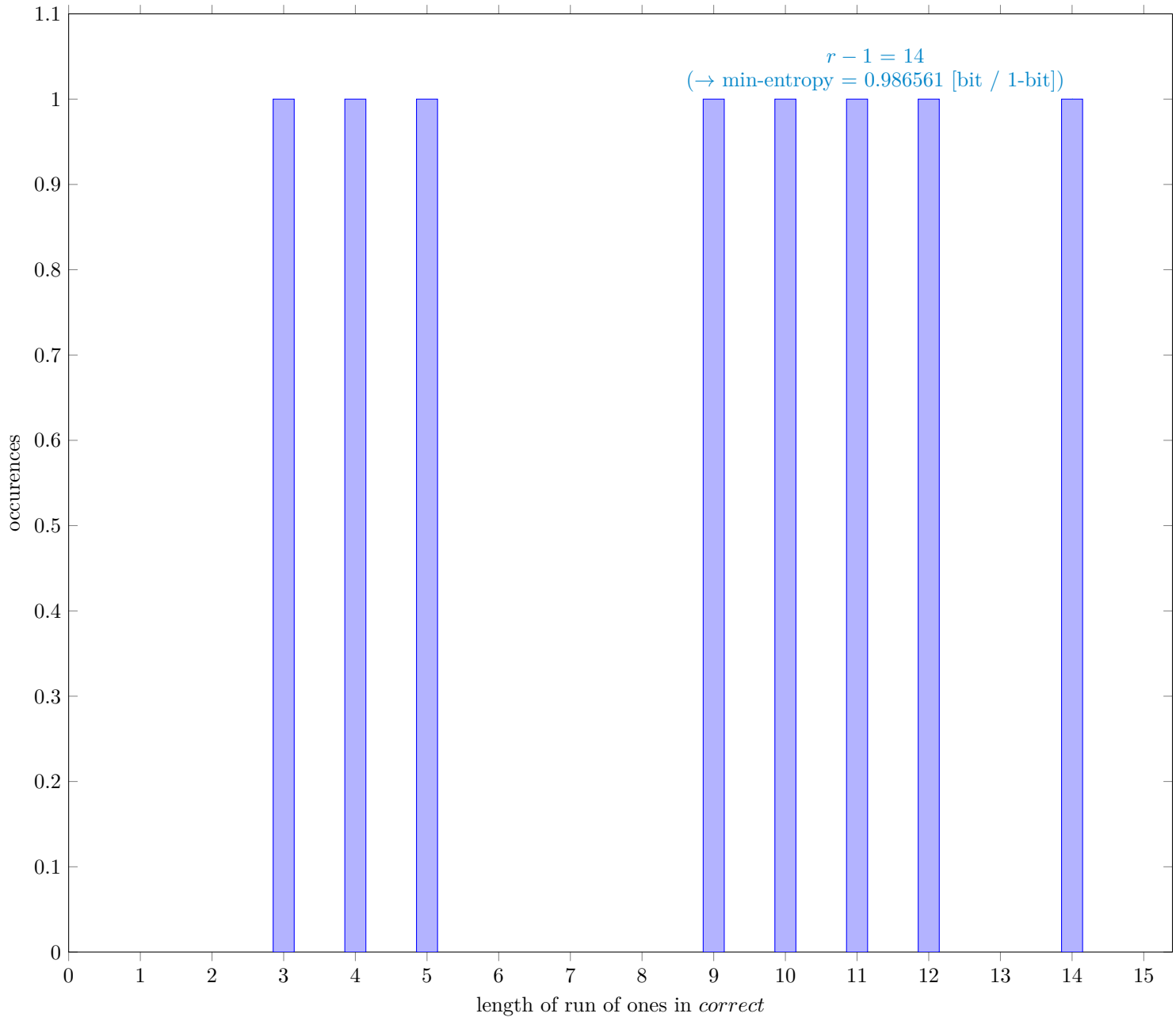

Fig. 22 Distribution of *correct*

4.7.1 Supplemental information for traceability

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

Symbol	Value
N	39937
C	19898
P_{global}	0.498235
P'_{global}	0.504679
r	15
P_{local}	0.374676

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

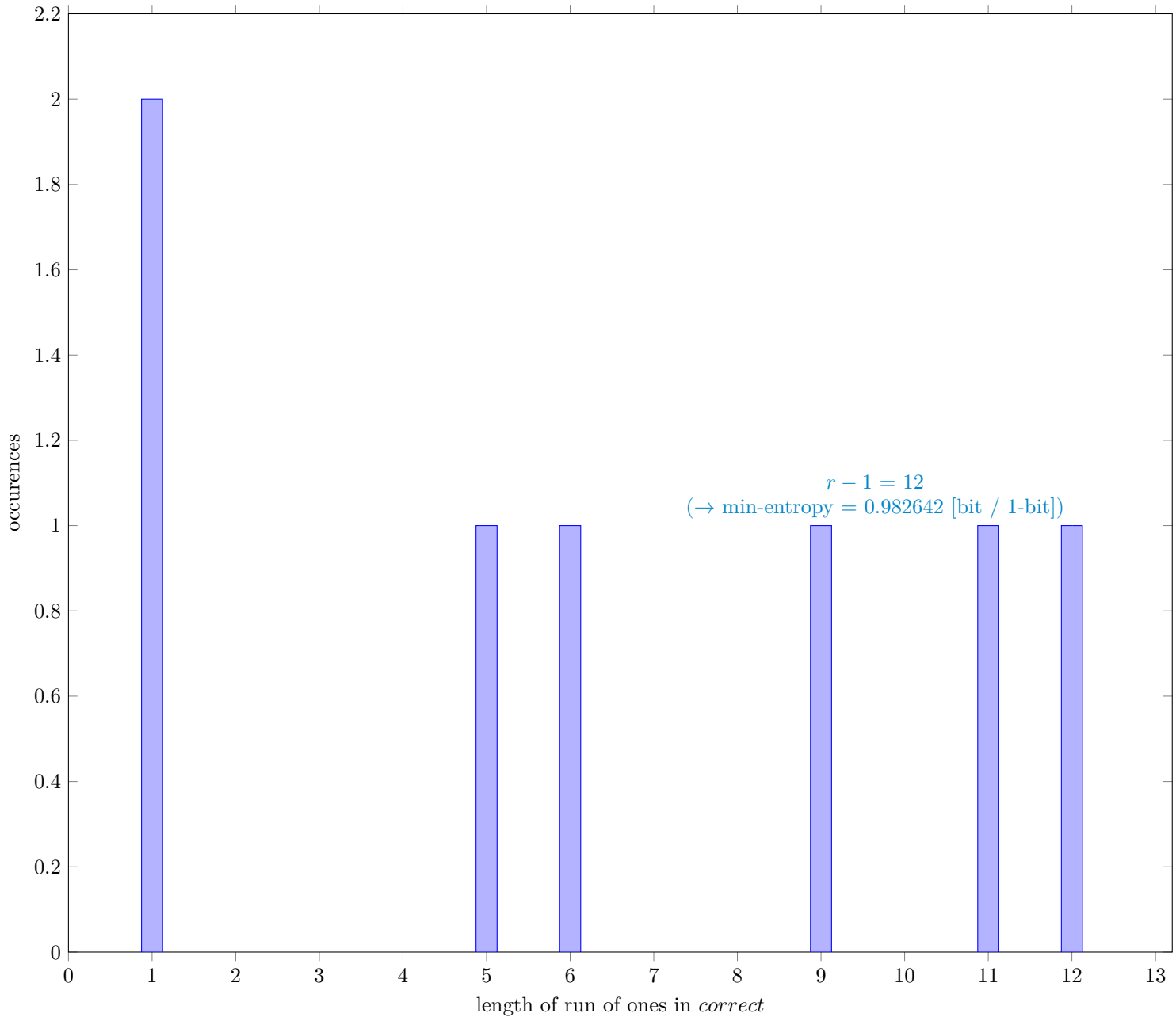

Fig. 23 Distribution of *correct*

4.8.1 Supplemental information for traceability

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

Symbol	Value
N	39999
C	19984
P_{global}	0.499612
P'_{global}	0.506052
r	13
P_{local}	0.320047

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

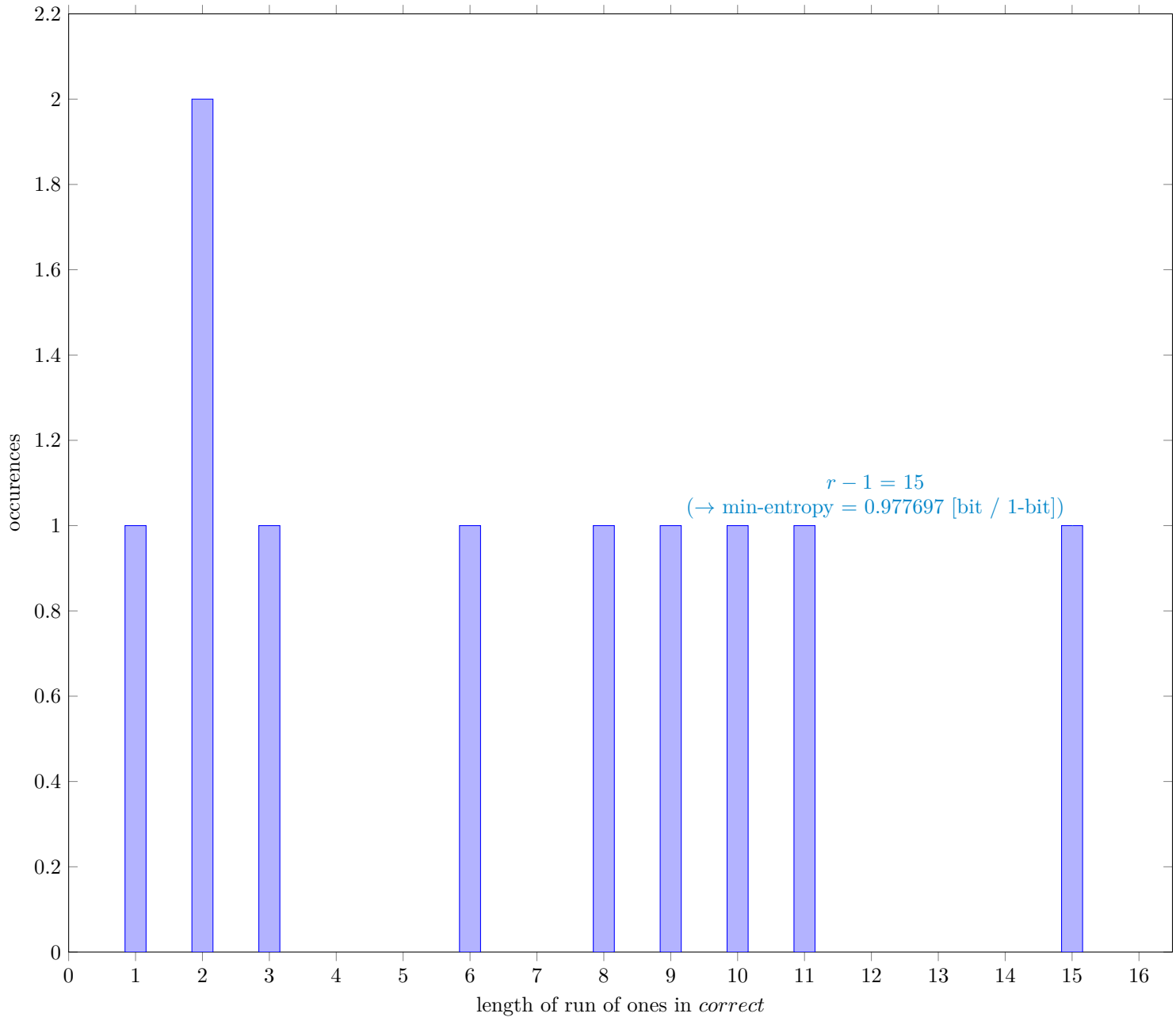

Fig. 24 Distribution of *correct*

4.9.1 Supplemental information for traceability

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

Symbol	Value
N	39998
C	20053
P_{global}	0.50135
P'_{global}	0.50779
r	16
P_{local}	0.399351

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

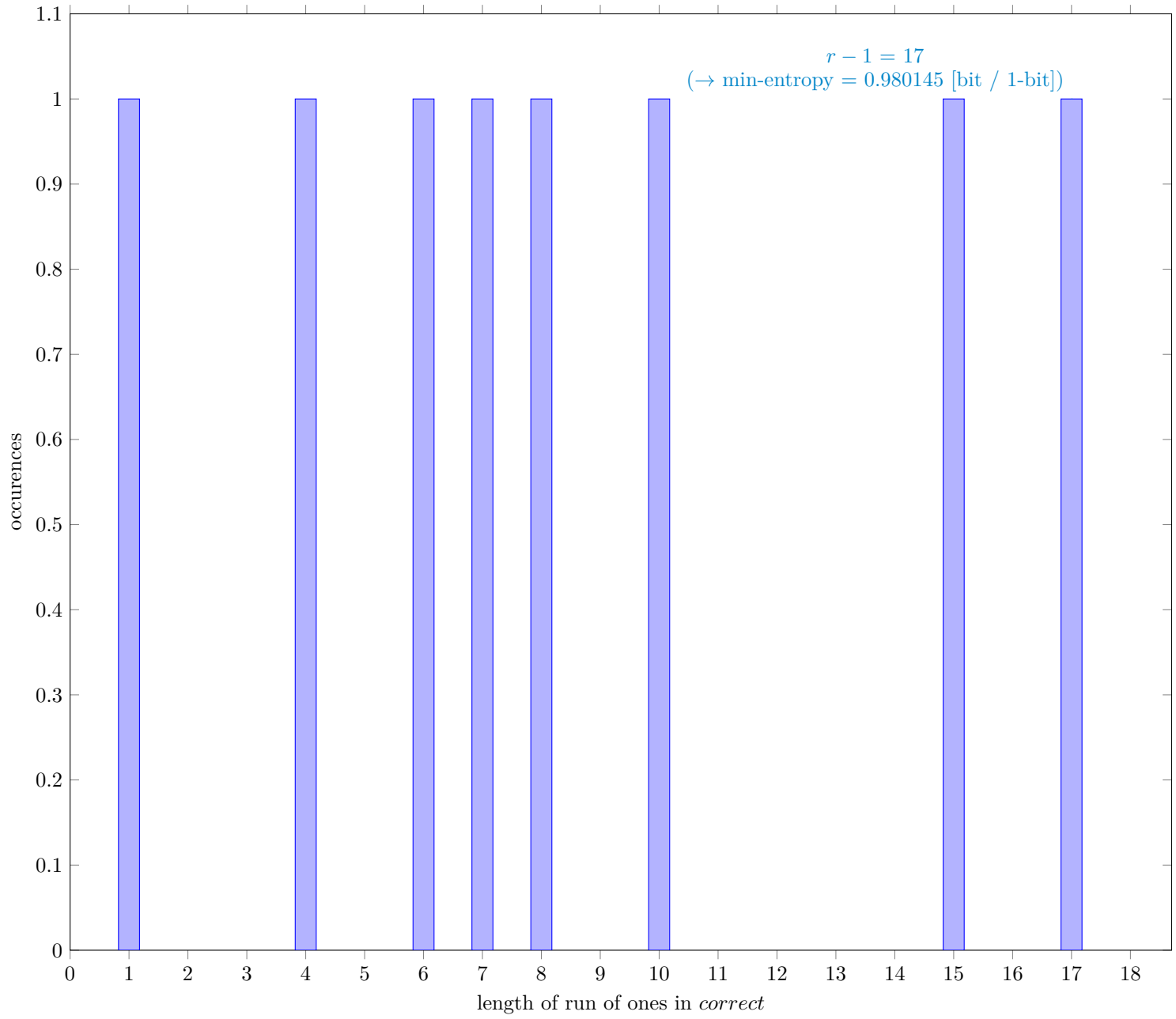

Fig. 25 Distribution of *correct*

4.10.1 Supplemental information for traceability

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

Symbol	Value
N	39983
C	20011
P_{global}	0.500488
P'_{global}	0.506929
r	18
P_{local}	0.444149

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`