

“Capacity” Is a Number Attainable Only at Infinite Length

— At Blocklength 100 Only 41.68% of the Capacity Is Usable, and Reaching 99.99% Takes 3.4×10^9 — C Is a
Supremum, Not a Maximum — [Paper 304]

Yuuki YAMAGISHI (山岸 勇輝)

Yamagishi Complexity Torus Lab, Japan

August 2026

— *From thinking to syncing.* —

Abstract

Shannon’s coding theorem says that below the channel capacity C the error can be made arbitrarily small. This paper asks whether C itself is attainable—the answer is not at any finite blocklength. No new mathematical theorem and no new law is claimed. First, capacity is fixed by the error rate. For the binary symmetric channel $C = 1 - h(p)$, and $C = 0$ at $p = 0.5$ (Section 2). Second, this is the core of the paper. At blocklength $n = 100$, only 41.68% of the capacity is usable (Section 3). Third, reaching 99% takes $n = 3.4 \times 10^5$. Reaching 99.99% takes 3.4×10^9 (Section 3). Fourth, the loss falls only as $1/\sqrt{n}$. Multiply n by 100 and the loss is one tenth (Section 4). Fifth, C is not a maximum. $R = C$ is attained at no n , and only as $n \rightarrow \infty$ does $R \rightarrow C$ (Section 5). Sixth, the separator is an attained maximum against an unattained supremum (Section 6).

Scope of this paper (scope note): No new mathematical theorem and no new law is claimed—Shannon’s coding theorem, the capacity of the binary symmetric and Gaussian channels, and the finite-blocklength normal approximation are all standard. We do not build information theory—all we use is one binary entropy and one square root. We do not prove the coding theorem—achievability and the converse are merely quoted. We construct no codes—which codes approach capacity is not treated at all. Paper 163’s existence-versus-construction distinction is alive here too. We claim no accuracy for the normal approximation— $R(n, \varepsilon) \approx C - \sqrt{V/n} Q^{-1}(\varepsilon)$ is a second-order approximation and errs appreciably at small n . The 41.68% at $n = 100$ is an estimate, not an exact achievable rate. We do not discuss implementation—decoding effort, latency, and the performance of real codes are not treated. We do not criticise C —being a supremum is not a defect. Unattainable and meaningless are different. We fix one channel—Sections 3 and 4 are numbers for the single channel BSC($p = 0.11$). Other channels have other V and need other n . Relation to earlier papers: Paper 252 counted four quantities called “information,” needing different things—the C here is one of them (a supremum of mutual information), and its attainment is questioned. Paper 294 showed that separation improves only as the square root of length—the $1/\sqrt{n}$ here comes from the same root (the additivity of variance). Paper 163 separated “a good code exists” from “here is a good code”—this paper treats a third: how long it must be. Paper 302 treated how existence does not give quantity—here quantity can be answered, and the answer was infinity. What is added is computing that only 41.68% of capacity is usable at $n = 100$, giving the n required for 99% and 99.99% as 3.4×10^5 and 3.4×10^9 , confirming that the loss falls as $1/\sqrt{n}$, and placing as the separator that C is a supremum and not a maximum.

1. Notation and the Question

Setting: On the binary symmetric channel BSC(p) a transmitted bit flips with probability p . The capacity is

$$C = 1 - h(p), \quad h(p) = -p \log_2 p - (1 - p) \log_2 (1 - p)$$

Shannon’s coding theorem: If $R < C$, codes exist making the error arbitrarily small.

Question: Can $R = C$ be attained? And at what blocklength can $R < C$ be attained?

Observation 1.1

The theorem says “can be made arbitrarily small.” Not “becomes small”—a statement about a limit, not about any finite case.

2. Capacity Is Fixed by the Error Rate

Check 2.1:

Error rate p	$h(p)$	$C = 1 - h(p)$
0.000	0.000000	1.000000
0.010	0.080793	0.919207
0.050	0.286397	0.713603
0.110	0.499916	0.500084
0.250	0.811278	0.188722
0.500	1.000000	0.000000

Observation 2.2

At $p = 0.5$, $C = 0$. A channel that flips half the time carries nothing—because the output becomes independent of the input.

Observation 2.3

At $p = 0.11$, $C = 0.500084$, almost exactly a half. We use this channel below—its C is near a round number, which makes comparison easy.

Observation 2.4

At $p = 0$ and at $p = 1$ alike, $C = 1$. A channel that flips everything has the same capacity as an error-free one—the receiver need only know it flips. Capacity is fixed not by how many errors there are but by how unpredictable they are.

3. The Core: At Blocklength 100, 41.68% of Capacity

Setting: With a finite blocklength n and a permitted error ε , the achievable rate is, in the normal approximation,

$$R(n, \varepsilon) \approx C - \sqrt{\frac{V}{n}} Q^{-1}(\varepsilon)$$

where V is the channel dispersion.

Check 3.1: BSC($p = 0.11$) with $\varepsilon = 10^{-3}$.

Quantity	Value
Capacity C	0.500084
Channel dispersion V	0.890702
$Q^{-1}(10^{-3})$	3.090232

Check 3.2:

Blocklength n	Achievable R	% of C	Loss $C - R$
100	0.208437	41.68%	0.291647
1000	0.407857	81.56%	0.092227
10^4	0.470919	94.17%	0.029165
10^5	0.490861	98.16%	0.009223
10^6	0.497168	99.42%	0.002916
10^8	0.499792	99.94%	0.000292

Observation 3.3

This is the core of the paper. At blocklength 100, only **41.68%** of the capacity is usable—the number C means almost nothing to a short code.

Check 3.4: Conversely, the n required to reach a given fraction of C .

Fraction of C	Blocklength n required
99.00%	3.401×10^5
99.90%	3.401×10^7
99.99%	3.401×10^9

Observation 3.5

Reaching 99.99% takes a codeword of **3.4 billion bits**. One codeword, that long——each further ninety-nine costs a factor of 100 in length.

Observation 3.6

A point to write honestly——the normal approximation errs appreciably at small n . The 41.68% at $n = 100$ is an estimate, not an exact achievable rate——what can be shown is the order: a short code cannot use even half the capacity.

4. The Loss Falls Only as $1/\sqrt{n}$

Check 4.1:

Change in n	Loss before	Loss after	Factor
$100 \rightarrow 10^4$	0.291647	0.029165	0.1000
$10^4 \rightarrow 10^6$	0.029165	0.002916	0.1000

Observation 4.2

Multiply n by **100** and the loss falls to exactly one tenth. It is $1/\sqrt{n}$ ——from the same root (the additivity of variance) as Paper 294’s “separation improves only as the square root of length.”

Observation 4.3

So the price is the square of the improvement sought. Cutting the loss tenfold costs a hundredfold in blocklength—the same numbers as Paper 294’s table, standing in a wholly different field.

Observation 4.4

And it never reaches 0. $\sqrt{V/n}Q^{-1}(\varepsilon)$ is a positive quantity and always remains at finite n —which leads to the core of Section 5.

5. C Is a Supremum, Not a Maximum

Observation 5.1

The loss $\sqrt{V/n}Q^{-1}(\varepsilon)$ is positive at every finite n . So $R < C$ is always a strict inequality— $R = C$ is not attained.

Check 5.2: We write the two standings apart.

	Definition	Attained?
Maximum	the greatest element of the set	yes
Supremum	the least upper bound	not necessarily

Observation 5.3

C is the supremum of the set of achievable rates, not its maximum. It does not belong to the set—it has the standing of the right endpoint of the half-open interval $[0, C)$.

Observation 5.4

The distinction is not verbal. When one says “a channel of capacity C ,” no device sending C bits exists——only that devices arbitrarily close to C can be built.

Observation 5.5

The Gaussian channel is the same.

S/N [dB]	$C = \frac{1}{2} \log_2(1 + S/N)$
0	0.500000
3	0.791341
10	1.729716
20	3.329106
30	4.983613

Observation 5.6

Here too attainment requires infinitely long Gaussian codewords. Change the channel and the standing does not change—— C is a supremum on every channel.

6. Separator: An Attained Maximum against an Unattained Supremum

Observation 6.1

Separator——what must be moved for C to become attainable? Making the blocklength infinite. While it is finite, no code reaches it.

Observation 6.2

This is the same figure as Paper 266. The sampling theorem is correct, and its premise (band limitation) is never met—here the theorem is correct, and its limit (infinite length) is never realised.

	Theorem	What is never met
Paper 266 (sampling)	correct	the premise (band limitation)
This paper (coding)	correct	the limit (infinite length)

Observation 6.3

The two fall short in different places. A premise is at the entrance, a limit at the exit—the single phrase “the theorem is correct” cannot tell them apart.

Observation 6.4

Paper 163 separated “a good code exists” from “here is a good code.” This paper sets down a third sentence—“how long a good code must be.” The answer is 3.4×10^9 (for 99.99% of C on BSC(0.11) at $\epsilon = 10^{-3}$).

Observation 6.5

A point to write honestly—being a supremum is not a defect. C gives a line that cannot be crossed, which is useful in itself—what this paper says is only that reading C as “an attainable value” gets the numbers wrong.

7. Roll Call and Closing

1. $C = 0$ at $p = 0.5$ (Check 2.1): Output independent of input.
2. $C = 1$ at both $p = 0$ and $p = 1$ (Observation 2.4): Unpredictability acts.
3. 41.68% at $n = 100$ (Check 3.2): Not even half usable.
4. 3.401×10^5 for 99% (Check 3.4): 3.4×10^9 for 99.99%.

5. A factor of 100 per ninety-nine (Observation 3.5): **3.4 billion bits.**
6. $100 \times n$ gives loss $1/10$ (Check 4.1): $1/\sqrt{n}$.
7. The same root as Paper 294 (Observation 4.2): **Additivity of variance.**
8. The loss is positive at every finite n (Observation 4.4): **Never zero.**
9. C is a supremum, not a maximum (Observation 5.3): **Not in the set.**
10. Different from Paper 266's failure (Observation 6.3): **Premise or limit.**

Closing: Shannon's coding theorem says that below C the error can be made arbitrarily small. But C itself is attained at no finite blocklength. Counting on BSC ($p = 0.11$) at error 10^{-3} : at blocklength 100 only **41.68%** of the capacity is usable—99% takes 3.401×10^5 , and 99.99% takes **3.4×10^9** , a codeword of 3.4 billion bits. The loss falls only as $1/\sqrt{n}$ —multiplying n by **100** divides the loss by 10, from the same root (the additivity of variance) by which Paper 294 measured separation. And at every finite n the loss is positive— C is not the maximum of the set of achievable rates but its supremum. One thing separates them—whether the number belongs to the set or not. Say “a channel of capacity C ” and still no device sending C bits exists. What exists is only the fact that devices arbitrarily close to C can be built.

References

1. Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3-4).
 2. Polyanskiy, Y., Poor, H. V. and Verdu, S. (2010). Channel Coding Rate in the Finite Blocklength Regime. *IEEE Transactions on Information Theory*, 56(5).
 3. Strassen, V. (1962). Asymptotische Abschätzungen in Shannons Informationstheorie. In *Transactions of the Third Prague Conference on Information Theory*.
 4. Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory*. Hoboken: Wiley.
 5. Gallager, R. G. (1968). *Information Theory and Reliable Communication*. New York: Wiley.
 6. Yamagishi, Y. (2026). “A Good Code Exists” and “Here Is a Good Code” Are Different Sentences [Paper 163]. Zenodo. 10.5281/zenodo.22092677
 7. Yamagishi, Y. (2026). *Information Is Four Quantities, and They Need Different Things* [Paper 252]. Zenodo. 10.5281/zenodo.22161148
 8. Yamagishi, Y. (2026). *The Sampling Theorem Is Correct and Its Premise Is Never Met* [Paper 266]. Zenodo. 10.5281/zenodo.22169765
 9. Yamagishi, Y. (2026). *Separation Improves Only as the Square Root of the Length* [Paper 294]. Zenodo. 10.5281/zenodo.22171603
-

Yuuki YAMAGISHI (山岸 勇輝)

© 2026 Yamagishi Complexity Torus Lab, Japan · Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)