

From Shannon to Gibbs

Economics of Information: Part 2A

Steph Macurdy
Institute of Thermoeconomics

Dr. Karl Kreder
Dominant Strategies

October 26, 2026

Abstract

Shannon and Gibbs entropy share a mathematical form, but connecting information to thermodynamic cost requires specifying the physical states and operations involved. We develop this connection and apply it to proof of work. The law of large numbers gives expected surprisal an empirical interpretation, while Jaynes's maximum-entropy construction yields equilibrium Gibbs ensembles from physical constraints. Boltzmann's constant expresses their entropy in thermodynamic units. An exact decomposition of memory entropy into logical and internal contributions then identifies when logical erasure reduces physical entropy and entails a Landauer lower bound on mean heat transfer. For an ideal classical proof of work search with independent trials, we assume that every failed trial undergoes a reset with a specified minimum mean heat cost, regardless of the preceding trial history. Multiplying this minimum cost by the expected number of failures yields a lower bound on mean failed-reset heat per successful proof. Under repeated operation assumptions, a required sustained proof rate converts this heat bound into a lower bound on average failed reset power. These results separate the physical reset assumptions from the probabilistic search model and distinguish expected and long-run costs from the realized cost of an individual proof.

1 Introduction

Information theory assigns entropy to a distribution over messages; statistical mechanics assigns entropy to a distribution over physical microstates. Their shared mathematical form suggests a connection, but the form alone does not identify a message with a physical state or establish an energy cost for processing it. Those conclusions require a model of physical representation and a specified operation on that representation.

Shannon entropy is the expected surprisal of an outcome [1]. The law of large numbers explains how this expectation appears in repeated observations. Jaynes's maximum-entropy formulation supplies a second connection: when the outcomes are physical microstates and the constraints are physical observables, maximizing this same functional produces equilibrium ensembles [2]. The factor k_B converts the dimensionless functional into thermodynamic entropy.

A memory register makes this connection operational. A logical value generally corresponds to many physical microstates. Decomposing the register's physical entropy into uncertainty about the logical value and uncertainty within each logical state identifies when erasure reduces physical entropy. Landauer's principle then constrains the mean heat transferred to a reservoir during reset [7, 6].

Proof of work (PoW) is a useful application because it separates a probabilistic search requirement from the physical operations used to realize it. A target determines the success frequency of idealized trials. It does not, by itself, determine how failed outputs are represented, reused, uncomputed, or

reset. Connecting search difficulty to heat therefore requires both a search model and a physical reset model.

The contribution here is an expository synthesis and a conditional application of established results. Sections 2–4 develop the Shannon–Gibbs connection and its implications for memory. Sections 5 and 6 derive the resulting PoW bounds, including the role of a required computation rate. Throughout, *computational work* denotes a count of trials, whereas thermodynamic work and heat have units of energy.

2 Expected surprisal and the law of large numbers

Let X take values in a finite alphabet with probabilities p_i . Using natural logarithms, define the surprisal and Shannon entropy by

$$I(i) = -\ln p_i, \quad H(p) = -\sum_i p_i \ln p_i = \mathbb{E}[I(X)]. \quad (1)$$

Terms with $p_i = 0$ contribute zero. Entropy in bits is $H_2(p) = H(p)/\ln 2$. Entropy is a property of the distribution, while surprisal is assigned to an outcome relative to that distribution.

For independent, identically distributed samples $X_1, X_2, \dots$, the strong law of large numbers gives

$$\frac{1}{n} \sum_{j=1}^n I(X_j) \xrightarrow[n \rightarrow \infty]{\text{a.s.}} H(p). \quad (2)$$

Independence also gives $\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \prod_{j=1}^n p_{x_j}$. Consequently,

$$-\frac{1}{n} \ln \mathbb{P}(X_1, \dots, X_n) \xrightarrow[n \rightarrow \infty]{\text{a.s.}} H(p), \quad (3)$$

where the probability on the left is the probability mass of the realized sequence. This is the finite-alphabet, independent-sample form of the asymptotic equipartition property [3]. A set of sequences carrying nearly all the probability has cardinality with exponential rate $H(p)$; Appendix A states the precise bounds.

The assumptions matter. Uniform marginals do not establish independence. For a uniform alphabet of size M , each outcome already has surprisal $\ln M$, so its sample average equals $\ln M$ at every sample size. The law of large numbers is essential for general nonuniform distributions, but a uniform distribution does not need an asymptotic argument to compute its entropy.

Energy is another expectation when outcomes have assigned energies E_i :

$$U = \sum_i p_i E_i. \quad (4)$$

The shared expectation structure does not equate energy with surprisal. A relationship between them arises only after imposing a physical ensemble, as in the canonical construction below.

3 From Shannon entropy to Gibbs entropy

3.1 The probability distribution must describe physical states

Let ρ_a be a distribution over discrete physical microstates a . The Gibbs entropy is

$$S_G(\rho) = -k_B \sum_a \rho_a \ln \rho_a = k_B H(\rho). \quad (5)$$

This is an exact identity between functionals of the *same* distribution. It does not identify the Shannon entropy of an arbitrary message with the thermodynamic entropy of the device carrying it.

Independent samples of this physical ensemble satisfy $n^{-1} \sum_{j=1}^n [-k_B \ln \rho_{a_j}] \rightarrow S_G$. This gives an empirical interpretation of Gibbs entropy, but it is not a derivation of the factor k_B , and it does not establish that samples along a physical trajectory are independent. For a uniform distribution over M accessible microstates, (5) reduces to Boltzmann's expression $S = k_B \ln M$.

The state space is part of the physical model. For continuous phase space, an invariant reference measure and a normalization convention must be specified; a coordinate-dependent differential entropy alone is insufficient. We use discrete states for the entropy identities and entropy differences for the classical calibration below.

3.2 Jaynes's maximum-entropy construction

Suppose the microstate energies E_a are known, and the only statistical information about an equilibrium system is normalization and mean energy:

$$\sum_a \rho_a = 1, \quad \sum_a \rho_a E_a = U. \quad (6)$$

Jaynes's prescription is to maximize $H(\rho)$ subject to these constraints [2]. At an interior solution, Lagrange multipliers give

$$\rho_a^* = \frac{e^{-\beta E_a}}{Z(\beta)}, \quad Z(\beta) = \sum_a e^{-\beta E_a}, \quad U = -\frac{\partial \ln Z}{\partial \beta}. \quad (7)$$

For a finite state space, strict concavity of entropy makes the maximizing distribution unique whenever the constraints admit an interior solution. More general state spaces require the relevant sums or integrals to exist.

Substitution into the Shannon functional gives

$$H(\rho^*) = \ln Z + \beta U, \quad \left(\frac{\partial H(\rho^*)}{\partial U} \right)_{\text{ext}} = \beta, \quad (8)$$

where external parameters determining the energy levels are held fixed. If equilibrium thermodynamic entropy is calibrated as $S = k_B H(\rho^*)$, the thermodynamic relation $(\partial S / \partial U)_{\text{ext}} = 1/T$ identifies

$$\beta = \frac{1}{k_B T}, \quad F = U - TS = -k_B T \ln Z. \quad (9)$$

Thus the exponential Gibbs distribution follows from maximizing Shannon entropy with an energy constraint, and thermodynamics identifies its multiplier with inverse temperature in energy units.

The physical content resides in the microstates, energies, constraints, and calibration. Maximum entropy supplies a rule for assigning an ensemble from that information; it does not by itself provide a dynamical explanation of equilibration. Jaynes also related Gibbs entropy to the volume of high-probability regions in appropriate thermodynamic limits [4]. Such concentration arguments complement the construction, but a general interacting system cannot be treated as independent particle samples without further assumptions.

3.3 Where Boltzmann's constant enters

The Shannon functional is dimensionless. To express it as thermodynamic entropy, introduce a proportionality constant κ , with units of energy per temperature. Its relation to macroscopic thermodynamics can be seen through an ideal-gas expansion.

Consider N classical ideal-gas particles undergoing a reversible isothermal expansion from V_1 to V_2 . The position contribution to the accessible phase-space volume scales as V^N ; the momentum contribution is unchanged. Therefore the change in dimensionless entropy is

$$\Delta H = N \ln \frac{V_2}{V_1}. \quad (10)$$

On the thermodynamic side, $PV = n_{\text{mol}}RT$ gives

$$\Delta S = \int_{V_1}^{V_2} \frac{P dV}{T} = n_{\text{mol}} R \ln \frac{V_2}{V_1}. \quad (11)$$

Using $N = n_{\text{mol}}N_A$, matching $\Delta S = \kappa\Delta H$ identifies

$$\kappa = k_B = \frac{R}{N_A}. \quad (12)$$

Volume-independent factors in the classical state count cancel in this entropy difference. The argument calibrates the statistical expression using an established physical model and the ideal-gas law; it does not derive a dimensionful numerical constant from probability theory alone.

In the modern SI, the kelvin is defined by fixing $k_B = 1.380649 \times 10^{-23} \text{ J K}^{-1}$ exactly [5]. Historically, its value connected measurements on microscopic and macroscopic scales. In either presentation, the physical calibration is indispensable. Changing the logarithm base adds only a mathematical conversion:

$$S_G = k_B H = k_B \ln 2 H_2. \quad (13)$$

4 Physical memory and erasure

4.1 Logical and internal entropy

Partition a memory's physical microstates into disjoint sets Ω_i , one for each logical value i . Define

$$p_i = \sum_{a \in \Omega_i} \rho_a, \quad r_{a|i} = \frac{\rho_a}{p_i}, \quad S_i = -k_B \sum_{a \in \Omega_i} r_{a|i} \ln r_{a|i}. \quad (14)$$

Terms with $p_i = 0$ are omitted. Expanding $\ln \rho_a = \ln p_i + \ln r_{a|i}$ gives the exact decomposition

$$\boxed{S_G(\rho) = k_B H(p) + \sum_i p_i S_i.} \quad (15)$$

This is the entropy chain rule applied to a physical memory [6]. The first term represents uncertainty about the logical value; the second represents physical entropy within the logical states.

Equation (15) supplies the missing condition in a direct identification of logical Shannon entropy with physical entropy. For a symmetric memory whose conditional internal entropies all equal S_{int} , with that contribution unchanged between initial and final states,

$$S_G = k_B H(p) + S_{\text{int}}, \quad \Delta S_G = k_B \Delta H(p). \quad (16)$$

If the internal contribution changes, it must remain in the accounting. Logical entropy reduction alone then does not determine physical entropy reduction.

4.2 The mean heat cost of reset

Consider a register coupled to a reservoir initially in thermal equilibrium at temperature $T > 0$. Assume an initially uncorrelated register and reservoir, and include or restore any auxiliary systems used as entropy sinks. With Q_{bath} positive for heat transferred to the reservoir, the entropy balance gives

$$\Delta S_{\text{G}} + \frac{\mathbb{E}[Q_{\text{bath}}]}{T} \geq 0. \quad (17)$$

A finite-reservoir treatment retains nonnegative corrections to this familiar Landauer inequality [8].

An exact reset maps every supported logical value to one fixed value. For the symmetric memory in (16), the final logical entropy is zero, so

$$\boxed{\mathbb{E}[Q_{\text{bath}}] \geq k_{\text{B}} T H(p) = k_{\text{B}} T \ln 2 H_2(p).} \quad (18)$$

An unbiased bit has the familiar bound $k_{\text{B}} T \ln 2$. A biased bit has a smaller entropy and a correspondingly smaller bound. Register width equals logical entropy only under the appropriate distribution.

This is a mean heat bound, not a lower bound on every fluctuating trajectory. Logical irreversibility also differs from thermodynamic irreversibility: in an ideal quasistatic reset, the bath's entropy increase can exactly compensate the memory's entropy decrease, with vanishing total entropy production [9].

Retained correlations matter. If h is a fixed deterministic hash and X its input, then $H_2(h(X) | X) = 0$. Apparent unpredictability of the hash does not remove this mathematical dependence. A reversible implementation can use retained inputs to uncompute intermediate values without inverting the hash [10]. A reset acting only on the output register and ignoring correlated information may still incur the marginal-entropy cost in (18); an implementation that uses that information requires joint accounting [8, 6]. The scope of the reset operation must therefore be stated before assigning it a heat cost.

5 The information in a proof-of-work search

5.1 An ideal classical trial model

Fix a challenge and a target set $A \subset \{0, 1\}^w$, with

$$p = \frac{|A|}{2^w}, \quad 0 < p < 1, \quad d = -\log_2 p. \quad (19)$$

Model fresh trial outputs $Y_1, Y_2, \dots$ as independent uniform w -bit strings, and set $B_j = \mathbf{1}_{\{Y_j \in A\}}$. Then B_j are independent Bernoulli variables with success probability $p = 2^{-d}$. The random-oracle abstraction motivates this model for distinct, previously unqueried inputs [11]. It is an idealization of a search process, not a claim that deterministic hashing generates independent physical randomness.

This analysis counts target-valid proofs for a fixed challenge. Network acceptance, stale blocks, changing targets, and reuse across challenges require additional modeling.

Several different information quantities appear in this search:

Quantity	Value in bits
Entropy of an unrestricted trial output	w
Surprisal of the success event	d
Entropy of the success/failure indicator	$h_2(p)$
Output entropy conditioned on success	$w - d$
Output entropy conditioned on failure	$w + \log_2(1 - p)$

Here $h_2(p) = -p \log_2 p - (1 - p) \log_2(1 - p)$. The conditional entropies follow because a uniform output remains uniform on A or its complement after conditioning. In particular, increasing difficulty increases the surprisal of success while leaving the unrestricted output entropy equal to w . Difficulty bits are not automatically erased bits.

5.2 Expected trials and long-run frequencies

Let $N = \min\{j \geq 1 : B_j = 1\}$ and let $F = N - 1$ count failures before the first success. The geometric distribution gives

$$\mathbb{P}(N = n) = (1 - p)^{n-1}p, \quad \mathbb{E}[N] = \frac{1}{p}, \quad \mathbb{E}[F] = \frac{1 - p}{p}. \quad (20)$$

Moreover, $\text{Var}(N) = (1 - p)/p^2$, so the ratio of standard deviation to mean is $\sqrt{1 - p}$. At high difficulty this ratio is close to one: an individual search does not concentrate around its expected number of trials.

For $K_n = \sum_{j=1}^n B_j$, the strong law instead yields

$$\frac{K_n}{n} \xrightarrow{\text{a.s.}} p, \quad \frac{n - K_n}{K_n} \xrightarrow{\text{a.s.}} \frac{1 - p}{p}. \quad (21)$$

The second ratio is defined eventually almost surely. These are long-run statements about repeated searches. A valid proof alone does not certify that a particular number of unsuccessful trials actually occurred.

6 Conditional heat and power bounds

6.1 A specified reset model

The probabilistic model supplies a failure count. To turn that count into heat, we now add a physical assumption.

Assumption 1 (Failed-trial reset). Each unsuccessful trial is followed by a specified reset operation at reservoir temperature T . Under the entropy accounting of Section 4, the reset decreases the physical entropy of the specified memory by at least $k_B \ln 2 s_f$, with $s_f \geq 0$. Correlated side information is either included in that accounting or is unavailable to the reset operation. The resulting conditional mean heat per failed trial is at least

$$q_0 = k_B T \ln 2 s_f, \quad (22)$$

uniformly over the preceding trial history. Conditional absolute mean heats are uniformly bounded, so stopped sums below are integrable.

This is an implementation assumption. For a symmetric register containing an output physically distributed uniformly over the failed-output set, reset locally without using correlated input information, it gives

$$s_f = H_2(Y \mid B = 0) = \log_2(2^w - |A|) = w + \log_2(1 - p). \quad (23)$$

The physical ensemble and reset conditions in this example are additional to the search abstraction. Statistical unpredictability to a verifier does not by itself establish them. In particular, a random-oracle ensemble over possible functions must not silently be substituted for the physical state distribution of a particular machine.

Proposition 2 (Mean heat before a successful proof). *Under the independent trial model and Assumption 1, the total heat Q_{failed} associated with failed-trial resets before the first success satisfies*

$$\boxed{\mathbb{E}[Q_{\text{failed}}] \geq k_B T \ln 2 s_f \frac{1-p}{p}}. \quad (24)$$

Proof. The probability that trial j is reached and fails is $(1-p)^j$. Its contribution to expected reset heat is therefore at least $q_0(1-p)^j$. Summing over $j \geq 1$, justified by the integrability assumption, gives $\mathbb{E}[Q_{\text{failed}}] \geq q_0 \sum_{j \geq 1} (1-p)^j = q_0(1-p)/p$. $\square$

For the full-output reset model (23), this becomes

$$\mathbb{E}[Q_{\text{failed}}] \geq k_B T \ln 2 [w + \log_2(1-p)] \frac{1-p}{p}. \quad (25)$$

For $p \ll 1$, the expression on the right has leading behavior $w k_B T \ln 2 2^d$. The factor 2^d comes from expected search length, while the factor close to w comes from the specified register's entropy. If an implementation does not materialize and reset that register, its appropriate s_f must be determined separately.

6.2 From repeated trials to sustained power

For a long-run statement, suppose the pairs (B_j, D_j) are independent and identically distributed, where D_j is the integrable failed-reset heat at trial j , set to zero on success. Set $q_f = \mathbb{E}[D_j \mid B_j = 0] \geq q_0$ and write $C_n = \sum_{j=1}^n D_j$. The strong law gives

$$\frac{C_n}{n} \xrightarrow{\text{a.s.}} (1-p)q_f, \quad \frac{C_n}{K_n} \xrightarrow{\text{a.s.}} \frac{(1-p)q_f}{p}. \quad (26)$$

Thus the mean per-proof expression also bounds the limiting heat per successful proof under these repeated-operation assumptions.

Let $n(t)$ count completed trials, including completion of their required failed resets, and assume $n(t)/t \rightarrow R > 0$. The rate R may be the aggregate throughput of many devices. Then

$$\lambda = \lim_{t \rightarrow \infty} \frac{K_{n(t)}}{t} = Rp, \quad \bar{P}_{\text{failed}} = \lim_{t \rightarrow \infty} \frac{C_{n(t)}}{t} = R(1-p)q_f \geq R(1-p)q_0. \quad (27)$$

If useful operation requires a sustained successful-proof rate $\lambda \geq 1/\tau$, it follows that

$$\boxed{\bar{P}_{\text{failed}} \geq \frac{k_B T \ln 2 s_f}{\tau} \frac{1-p}{p}}. \quad (28)$$

The time requirement therefore converts a conditional mean heat per proof into a conditional power requirement. It does not remove the physical assumptions needed for the heat bound.

A target block interval, as in difficulty adjustment, describes an aggregate average rate [12]; it is not a guarantee that every block arrives by a deadline. These requirements should be kept distinct. If success by a deadline is required with probability at least $\alpha \in (0, 1)$, the ideal search must have capacity for at least

$$q_{\min} = \left\lceil \frac{\ln(1 - \alpha)}{\ln(1 - p)} \right\rceil \quad (29)$$

trials within that time. Capacity is not the number actually executed by a search that stops at success. Appendix B derives the corresponding expected heat and states when it can be divided by the deadline duration to obtain a window-average power bound.

6.3 Illustrative scale

For illustration, take $T = 300$ K, $w = 256$, and the full-output reset model. For small p , the failed-reset heat floor is approximately 7.35×10^{-19} J per failed trial. Table 1 evaluates (25) and (28) for an aggregate mean proof interval of 600 s.

$d = -\log_2 p$	Mean heat floor per proof (J)	Erase power floor (W)
32	3.16×10^{-9}	5.26×10^{-12}
64	1.36×10^1	2.26×10^{-2}
80	8.89×10^5	1.48×10^3

Table 1: Conditional failed-reset bounds for $T = 300$ K, $w = 256$, and $\tau = 600$ s. These are ideal model bounds, not measured mining costs.

This calculation excludes heat from successful trials, other circuit operations, communication, and supporting equipment. It does not establish a finite-time optimum or show that any architecture can attain the reset floor at the stated throughput. Comparing with hardware requires measuring or modeling its reset process and the other energy contributions.

7 Discussion and conclusion

The connection developed here has a definite sequence. A distribution over physical microstates gives a Shannon functional; physical calibration expresses it as Gibbs entropy. A register partition then relates logical entropy to physical entropy through (15). Under specified reset conditions, physical entropy reduction entails a mean heat transfer. A probabilistic search supplies a failure count, and a sustained timing requirement converts the resulting heat per proof into average power.

The law of large numbers supports this sequence in two roles. It connects surprisal with empirical average information, and it connects repeated search and reset events with stable long-run costs. It neither supplies the units carried by k_B nor establishes a particular device’s entropy reduction. Jaynes’s construction supplies the statistical-mechanical connection through physical constraints; the memory decomposition supplies the connection needed for erasure.

The PoW result is correspondingly conditional. Hash outputs designed to behave unpredictably make the independent-trial search model useful, but deterministic dependence on retained inputs remains relevant to physical implementation. Reversible computation, algorithms using correlated information, quantum search, and changing-target protocols fall outside the stated model.

Finally, Q_{failed} is a specified heat contribution. A lower bound on it is not automatically a lower bound on an arbitrary apparatus's net electrical consumption or total thermodynamic work; those quantities need a complete energy and resource balance. Within the model, however, the distinction between search statistics and reset physics makes the result precise: the required rate of statistically defined work determines the rate at which a specified entropy-reducing operation must be performed, and thereby its minimum mean heat load.

A Typical sequences and entropy

For a finite-alphabet independent and identically distributed source with entropy H , define

$$T_{n,\epsilon} = \left\{ x^n : \left| -\frac{1}{n} \ln \mathbb{P}(x^n) - H \right| \leq \epsilon \right\}. \quad (30)$$

The strong law implies $\mathbb{P}(T_{n,\epsilon}) \rightarrow 1$. Each sequence in this set has probability between $e^{-n(H+\epsilon)}$ and $e^{-n(H-\epsilon)}$. Hence, for any $\delta > 0$ and all sufficiently large n ,

$$(1 - \delta)e^{n(H-\epsilon)} \leq |T_{n,\epsilon}| \leq e^{n(H+\epsilon)}. \quad (31)$$

Allowing arbitrarily small fixed ϵ identifies H as the exponential rate of typical-set cardinality [3]. The statement concerns sequences from the specified source. Identifying such sequences with microstates of a physical system is a further modeling step, not a consequence of the cardinality bound.

B A search stopped at success or a deadline

Suppose at most $q \geq 1$ trials fit within a duration τ . Execute trials until the first success or until the budget is exhausted. With N as in (20), the executed trial count is $N_q = \min(N, q)$. The success probability is

$$\alpha_q = \mathbb{P}(N \leq q) = 1 - (1 - p)^q, \quad (32)$$

which gives (29). The tail-sum formula yields

$$\mathbb{E}[N_q] = \sum_{j=1}^q (1 - p)^{j-1} = \frac{\alpha_q}{p}. \quad (33)$$

The number of executed failures is $F_q = \sum_{j=1}^q \mathbf{1}_{\{N > j\}}$, so

$$\mathbb{E}[F_q] = \sum_{j=1}^q (1 - p)^j = \frac{1 - p}{p} \alpha_q. \quad (34)$$

Under Assumption 1,

$$\mathbb{E}[Q_{\text{failed},q}] \geq q_0 \frac{1 - p}{p} \alpha_q. \quad (35)$$

This expectation includes both successful and unsuccessful runs. It is not conditioned on meeting the deadline.

If all the associated resets are also completed within the time window, the expected heat divided by τ bounds the mean reset power over that window. In particular, $\alpha_q \geq \alpha$ implies

$$\frac{\mathbb{E}[Q_{\text{failed},q}]}{\tau} \geq \frac{q_0}{\tau} \frac{1 - p}{p} \alpha. \quad (36)$$

If reset is deferred until after the deadline, (35) still bounds the eventual mean reset heat, but it does not locate that heat within the search window. A process that finishes early and idles also differs from the continuously repeated search used in (28).

References

- [1] C. E. Shannon, “A Mathematical Theory of Communication,” *Bell System Technical Journal* **27**, 379–423 and 623–656 (1948). doi:10.1002/j.1538-7305.1948.tb01338.x.
- [2] E. T. Jaynes, “Information Theory and Statistical Mechanics,” *Physical Review* **106**, 620–630 (1957). doi:10.1103/PhysRev.106.620.
- [3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley, 2006). doi:10.1002/047174882X.
- [4] E. T. Jaynes, “Gibbs vs Boltzmann Entropies,” *American Journal of Physics* **33**, 391–398 (1965). doi:10.1119/1.1971557.
- [5] Bureau International des Poids et Mesures, “SI Base Units: Kelvin,” <https://www.bipm.org/en/si-base-units/kelvin> (accessed October 8, 2026).
- [6] T. Sagawa, “Thermodynamic and Logical Reversibilities Revisited,” *Journal of Statistical Mechanics: Theory and Experiment*, P03025 (2014). doi:10.1088/1742-5468/2014/03/P03025.
- [7] R. Landauer, “Irreversibility and Heat Generation in the Computing Process,” *IBM Journal of Research and Development* **5**, 183–191 (1961). doi:10.1147/rd.53.0183.
- [8] D. Reeb and M. M. Wolf, “An Improved Landauer Principle with Finite-Size Corrections,” *New Journal of Physics* **16**, 103011 (2014). doi:10.1088/1367-2630/16/10/103011.
- [9] J. M. R. Parrondo, J. M. Horowitz, and T. Sagawa, “Thermodynamics of Information,” *Nature Physics* **11**, 131–139 (2015). doi:10.1038/nphys3230.
- [10] C. H. Bennett, “Logical Reversibility of Computation,” *IBM Journal of Research and Development* **17**, 525–532 (1973). doi:10.1147/rd.176.0525.
- [11] M. Bellare and P. Rogaway, “Random Oracles are Practical: A Paradigm for Designing Efficient Protocols,” in *Proceedings of the First ACM Conference on Computer and Communications Security*, 62–73 (1993). Authors’ publication page.
- [12] S. Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System” (2008). <https://bitcoin.org/bitcoin.pdf>.