

Negative Work as an Information Credit System: How Bias, Correlation, and Geometry Finance Low-Dissipation Computation

Ian Todd

Sydney Medical School

University of Sydney

`itod2305@uni.sydney.edu.au`

Abstract

Landauer’s principle is routinely misquoted as “erasing a bit costs $k_{\text{B}}T \ln 2$.” The correct statement is entropic: minimal work scales with *entropy removed*, not bit-count. This distinction matters because real substrates—biological, nanoscale, or engineered—are not featureless thermal baths. They contain *bias* (nonuniform priors) and *correlations* (side information) that reduce the entropy cost of information processing, plus *geometric structure* (optimal control landscapes) that reduce finite-time dissipation.

We formalize the first two as **state credit**—thermodynamic resources that can be accrued and spent—and distinguish them from **protocol efficiency**, which reduces irreversible losses but is not itself a conserved quantity. Episodes of apparent “negative Landauer work” correspond to spending stored state credit, not violating thermodynamics. We derive quantitative predictions testable in colloidal and biological systems.

Keywords: Landauer principle; nonequilibrium free energy; mutual information; thermodynamic length; information geometry

1 Introduction

The physics of information processing is often summarized by Landauer’s limit: erasing one bit of information requires dissipating at least $k_{\text{B}}T \ln 2$ of heat [Landauer, 1961]. This statement is correct only for a maximally uncertain bit in contact with a featureless thermal reservoir. The general bound is

$$W_{\text{rev}} \geq k_{\text{B}}T \Delta S_{\text{register}}, \quad (1)$$

where $\Delta S_{\text{register}}$ is the entropy *removed* from the information-bearing degrees of freedom [Bennett, 1982, Parrondo et al., 2015]. A biased bit has less entropy; erasing it costs less.

This entropic framing resolves apparent paradoxes. Experiments have demonstrated erasure below the “ $k_{\text{B}}T \ln 2$ limit” using tilted potentials [Bérut et al., 2012] and work extraction from information using feedback [Toyabe et al., 2010]. These results are not violations—they are consequences of structure in the initial state or correlations with a measurement apparatus.

The purpose of this paper is to provide a clear accounting scheme for these phenomena. We distinguish two mechanisms:

1. **State credit:** Bias and correlations reduce the *reversible* work bound by reducing the entropy that must be removed. This is a true thermodynamic resource—creating it costs work; consuming it recovers work.
2. **Protocol efficiency:** Geometric structure in the control landscape reduces *irreversible* dissipation during finite-time operations. This is not a conserved resource but a property of the transformation path.

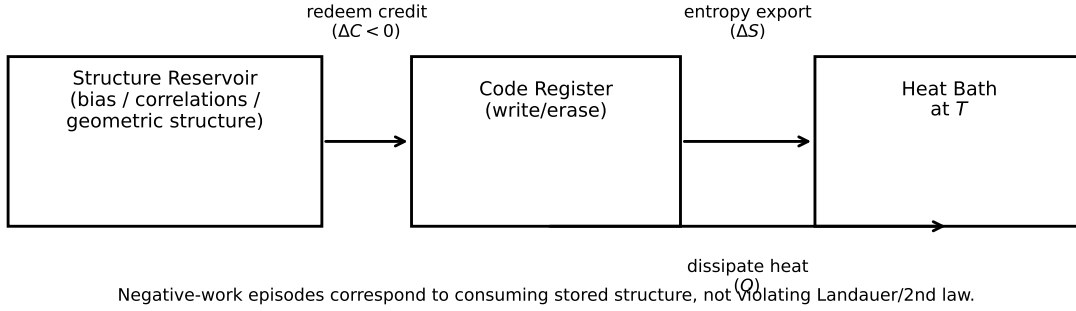


Figure 1: **The credit ledger.** State credit (bias, correlations) reduces the reversible work bound. Protocol efficiency (optimal paths through control space) reduces irreversible dissipation. Both mechanisms allow sub-naïve-Landauer operation; only state credit can produce negative reversible work.

2 State Credit I: Bias

Consider a binary register $X \in \{0, 1\}$ with probability $p = P(X = 1)$. Resetting to a standard state removes entropy

$$H(X) = -p \log_2 p - (1 - p) \log_2 (1 - p) \quad (2)$$

bits. The minimal (reversible) work is

$$W_{\text{rev}} = k_B T \ln 2 \cdot H(X). \quad (3)$$

For $p = 0.5$, this gives the familiar $k_B T \ln 2$. For $p = 0.1$, the entropy is $H \approx 0.47$ bits and the cost drops to $0.47 k_B T \ln 2$. The register’s bias is not separate from the entropy—it *is* the reduced entropy. No additional “bias credit” term is needed; the effect is already in $H(X)$.

The negentropy $1 - H(X)$ quantifies how much cheaper erasure is compared to the unbiased case. This negentropy was “paid for” when the bias was created—by measurement, asymmetric initialization, or prior computation.

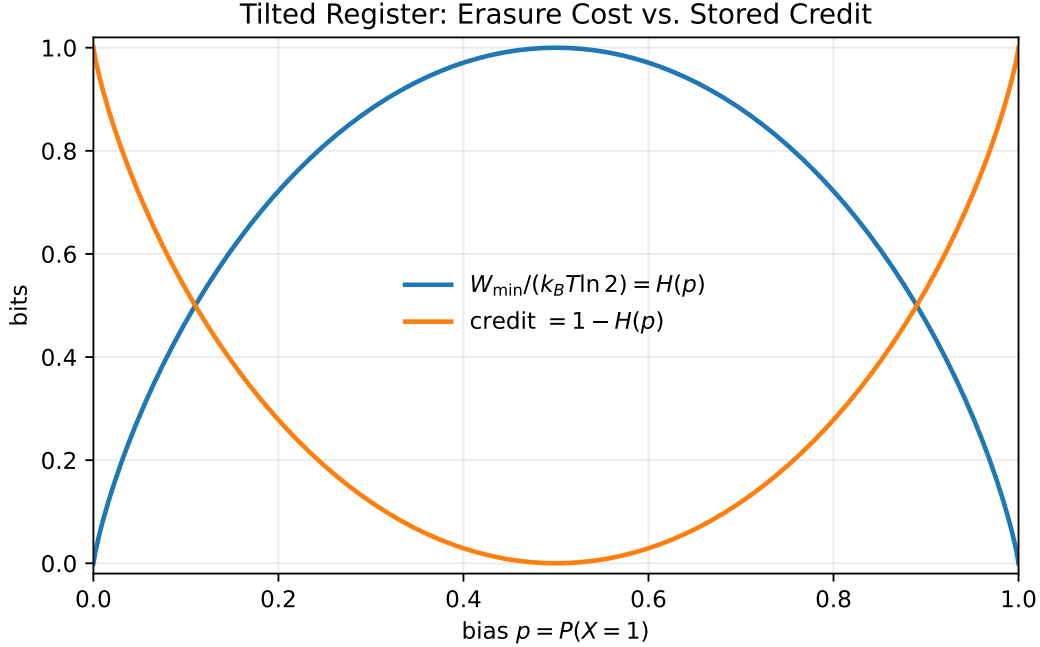


Figure 2: **Bias reduces erasure cost.** Erasure cost (blue) scales with Shannon entropy, not bit-count. The complement (orange) is negentropy—the “discount” relative to an unbiased bit.

3 State Credit II: Correlations

If the register X is correlated with an auxiliary system Y , the erasure bound tightens further. Given access to Y , the minimal work to reset X is [del Rio et al., 2011, Sagawa and Ueda, 2010]

$$W_{\text{rev}} = k_{\text{B}}T \ln 2 \cdot H(X|Y), \quad (4)$$

where $H(X|Y) = H(X) - I(X;Y)$ is the conditional entropy. The mutual information $I(X;Y)$ is a true thermodynamic resource:

- Creating $I(X;Y)$ bits of correlation costs at least $k_{\text{B}}T \ln 2 \cdot I(X;Y)$ in work.
- Consuming it (via conditional erasure or feedback) recovers at most the same amount.

The Sagawa-Ueda equality makes this precise [Sagawa and Ueda, 2010, Toyabe et al., 2010]:

$$W_{\text{extract}} \leq k_{\text{B}}T \ln 2 \cdot I(X;Y). \quad (5)$$

State credit is thus the sum of negentropy and accessible mutual information:

$$C_{\text{state}} = [H_{\text{max}} - H(X)] + I(X; Y), \quad (6)$$

where $H_{\text{max}} = \log_2 |\mathcal{X}|$ is the register's capacity (1 bit for a binary register). The reversible work bound becomes

$$W_{\text{rev}} = k_{\text{B}} T \ln 2 \cdot [H_{\text{max}} - C_{\text{state}}]. \quad (7)$$

Note that classically $I(X; Y) \leq H(X)$, so $C_{\text{state}} \leq H_{\text{max}}$ and $W_{\text{rev}} \geq 0$ always. Work extraction (Eq. 5) comes from *feedback operations* that exploit correlations, not from negative erasure work. The credit perspective clarifies that such extraction is financed by previously established mutual information.

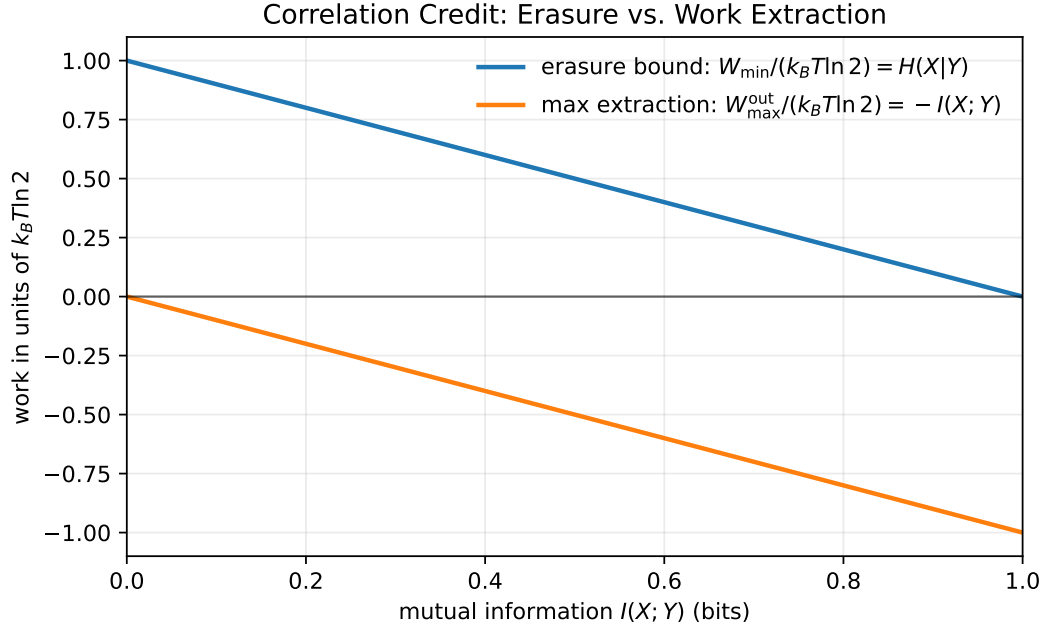


Figure 3: **Correlation credit.** For a uniform bit ($H(X) = 1$), side information $I(X; Y)$ reduces the erasure bound (blue) and sets the extraction ceiling (orange). At $I = 1$, erasure is free and full bit-to-work conversion is possible.

3.1 Conservation of state credit

The second law guarantees that state credit cannot increase without external work:

$$\frac{dC_{\text{state}}}{dt} \leq \frac{\dot{W}_{\text{in}}}{k_{\text{B}}T \ln 2}. \quad (8)$$

Credit can be transferred between subsystems, converted between forms (measurement converts work into correlation), or dissipated—but not created from nothing. This is the bookkeeping constraint that prevents perpetual motion.

4 Protocol Efficiency: Geometry

Bias and correlation concern the *endpoints* of a transformation (what states are involved). A separate question is *how* you get there. Even between fixed initial and final distributions, different protocols incur different dissipation.

For slow driving near equilibrium, total work decomposes as

$$W = W_{\text{rev}} + W_{\text{irr}}, \quad (9)$$

where W_{rev} is the reversible (quasistatic) work from Eq. (7) and W_{irr} is irreversible dissipation. The latter is bounded by [Sivak and Crooks, 2012, Crooks, 2007]

$$W_{\text{irr}} \geq \frac{k_{\text{B}}T}{2\tau} \mathcal{L}^2, \quad (10)$$

where τ is the protocol duration and \mathcal{L} is the **thermodynamic length**—the path length through parameter space measured in a metric derived from fluctuations (typically the Fisher information metric).

Consider univariate Gaussian distributions with mean μ and standard deviation σ . The

Fisher metric is

$$ds^2 = \frac{d\mu^2}{\sigma^2} + \frac{2 d\sigma^2}{\sigma^2}. \quad (11)$$

This metric is *anisotropic*: changes in μ are expensive when σ is small (the distribution is narrow and “notices” shifts), while changes in σ are uniformly weighted. Geodesics—paths minimizing \mathcal{L} —curve through this landscape to exploit cheap directions.

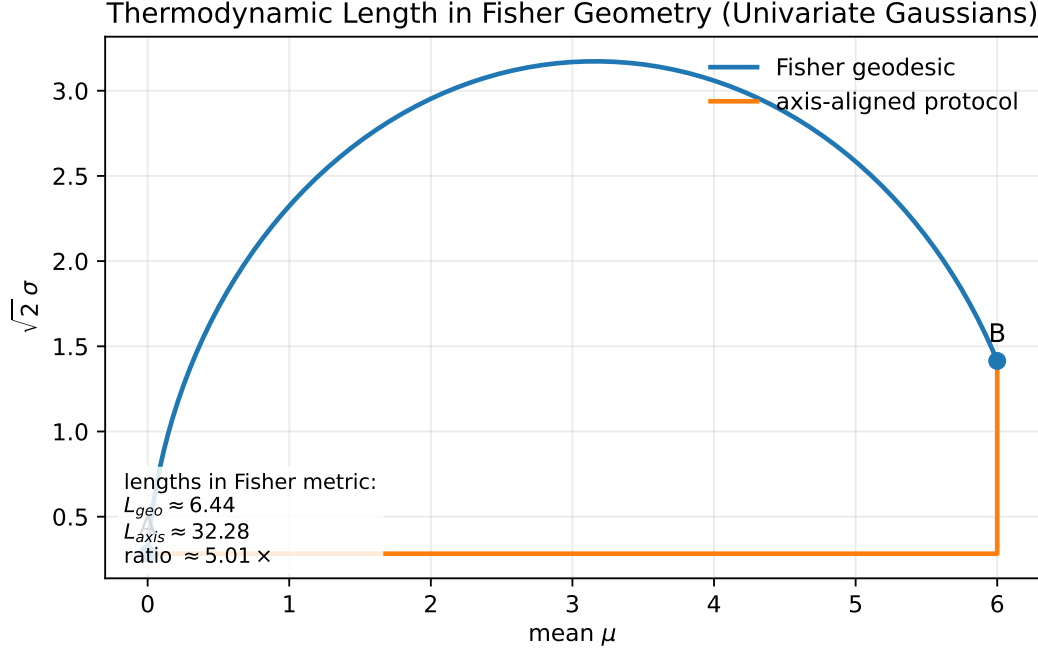


Figure 4: **Protocol efficiency.** In the Fisher geometry of Gaussians, geodesics (curved) are shorter than axis-aligned protocols. For duration τ , excess dissipation scales as \mathcal{L}^2/τ , so the geodesic reduces irreversible work by the squared-length ratio.

Key distinction: Protocol efficiency is *not* a conserved resource. Using an optimal protocol does not “spend” anything—it simply avoids waste. The geodesic is always available; the question is whether the controller knows it. This is why we separate protocol efficiency from state credit:

- State credit obeys a conservation law (Eq. 8).
- Protocol efficiency does not. It reduces W_{irr} but cannot make W_{rev} negative.

A system that has learned or evolved the optimal protocol dissipates less than one using

a naive protocol. But this advantage comes from *knowledge*, not from a thermodynamic reservoir that depletes upon use.

5 Combined Bound

Putting together state credit and protocol efficiency, the total work for a finite-time operation satisfies

$$W \geq k_{\text{B}}T \ln 2 \cdot [H_{\text{max}} - C_{\text{state}}] + \frac{k_{\text{B}}T}{2\tau} \mathcal{L}^2. \quad (12)$$

The first term is the reversible bound (reduced by state credit); the second is the irreversible floor (reduced by protocol efficiency). Both must be paid, but they are paid in different currencies:

- The first term approaches zero as $C_{\text{state}} \rightarrow H_{\text{max}}$ (erasure becomes free).
- The second term is always non-negative (dissipation cannot be negative).

Episodes of ultra-low dissipation occur when state credit is high (reducing the reversible term) *and* the protocol is efficient (reducing the irreversible term). Work extraction requires feedback operations (Eq. 5), not negative erasure.

6 Testable Predictions

6.1 Colloidal erasure with variable bias

Bérut et al. [Bérut et al., 2012] measured heat dissipation during erasure of a colloidal particle in a double-well potential. Extending their protocol:

- Vary the initial trap asymmetry to create bias $p \neq 0.5$.
- Prediction: measured heat $Q = k_{\text{B}}T \ln 2 \cdot H(p) + W_{\text{irr}}$, where $H(p)$ is the binary entropy.
- At $p = 0.1$, the reversible contribution should be $\sim 47\%$ of the symmetric case.

6.2 Feedback engine with partial information

Toyabe et al. [Toyabe et al., 2010] demonstrated work extraction using feedback. A more stringent test:

- Introduce controlled noise into the measurement channel, reducing $I(X;Y)$ below 1 bit.
- Prediction: extractable work $W_{\text{out}} \leq k_{\text{B}}T \ln 2 \cdot I(X;Y)$, saturating the Sagawa-Ueda bound.
- The mutual information can be independently estimated from measurement statistics.

6.3 Protocol optimization in molecular machines

For molecular motors or enzymes operating in finite time:

- Compare dissipation under naive (e.g., linear ramp) vs. optimized protocols.
- Prediction: irreversible dissipation ratio scales as $(\mathcal{L}_{\text{naive}}/\mathcal{L}_{\text{opt}})^2$.
- Biological motors may have evolved near-geodesic conformational pathways; synthetic motors typically have not.

6.4 Credit depletion in finite reservoirs

State credit conservation (Eq. 8) predicts depletion effects when the credit source is finite:

- Consider a register coupled to a finite biasing reservoir (e.g., a small thermal gradient or a finite-capacity battery).
- First erasure cycle: low dissipation (credit available from reservoir).
- Subsequent cycles without reservoir replenishment: dissipation increases as reservoir equilibrates.

- Prediction: exponential approach to unbiased-Landauer cost with time constant set by reservoir relaxation.

This distinguishes “cheap because structured” from “cheap because slow” (the latter shows no depletion).

7 Discussion

7.1 Relation to existing frameworks

The state-credit/protocol-efficiency distinction reorganizes established results [Parrondo et al., 2015, Seifert, 2012] into a cleaner accounting:

- **“Beating Landauer”**: No system beats the entropic bound. Systems with low $H(X|Y)$ simply have less entropy to remove.
- **Maxwell’s demon**: The demon’s memory is a correlation reservoir. Erasing it dissipates exactly what was “saved” during sorting [Bennett, 1982].
- **Feedback engines**: Work extraction is financed by mutual information, which was created by a prior measurement that cost at least as much.

7.2 Biological implications

Living systems maintain structure at multiple scales. The credit perspective connects to several active research programs:

1. **Cheap perception.** If internal states are already biased toward expected inputs, updating them costs less than updating on surprises. Metabolic cost of perception should scale with surprise—a thermodynamic Bayesian brain.
2. **Development as credit accumulation.** Embryogenesis creates correlated, low-entropy structures. These constitute a credit reserve for later adaptation.

3. **Aging as credit depletion.** Loss of molecular correlations and regulatory coherence depletes the state-credit reserve, making homeostatic operations progressively more costly. This aligns with complex-systems approaches to aging that emphasize declining resilience and network integrity [Cohen et al., 2022].
4. **Evolved protocol efficiency.** Natural selection can optimize not just what states organisms visit but how they traverse state space. Near-geodesic developmental trajectories would minimize dissipation.

7.3 Dimensional prerequisites

The framework presented here assumes a fixed state space. But all information-theoretic quantities—entropy, mutual information, surprise—are implicitly parameterized by the embedding dimension of that space. A “bit” only makes sense relative to a dimensional frame: the same physical process might register as one bit in a two-dimensional projection, zero bits if the apparent distinction collapses to a single equivalence class, or multiple bits in a higher-dimensional representation that resolves additional structure.

This is not merely a measurement issue. If a system is constrained to a low-dimensional embedding, distinctions that would exist in higher dimensions become *unrepresentable*, not merely expensive. For deterministic recurrent processes—those that revisit states without self-intersection—phase-preserving embeddings require at least three dimensions; in two dimensions, temporally distinct states are forced to collide [Todd, 2025]. (This is analogous to classical results on embedding dimension for attractors; cf. Takens’ theorem.) No amount of credit can purchase distinctions that the embedding cannot represent.

The credit framework thus operates downstream of a prior constraint: the system must already have access to a state space of sufficient dimension to represent the relevant distinctions. Dimensional constraints on representability are logically prior to thermodynamic costs of manipulation.

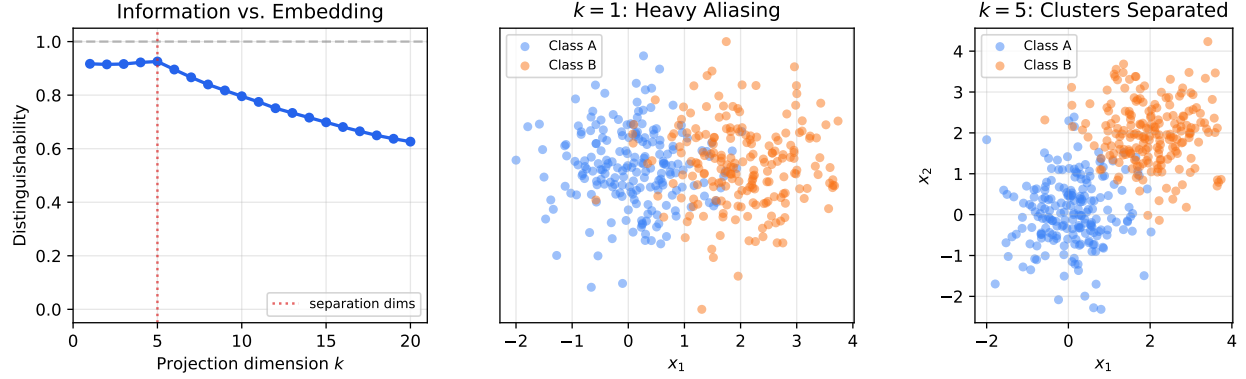


Figure 5: **Bits depend on embedding dimension.** Two synthetic clusters in \mathbb{R}^{20} , separated only in the first 5 dimensions. *Left*: Classification accuracy (linear SVM) vs. projection dimension k , where projection retains the first k coordinates. Distinguishability saturates once all separating dimensions are included. *Center*: At $k=1$, clusters overlap heavily—the distinction is *unrepresentable*. *Right*: At $k=5$, clusters separate cleanly. The “bit” distinguishing them exists only at sufficient k .

7.4 Limitations

This framework is accounting, not dynamics. It does not predict:

- How fast credit can be created or spent (depends on relaxation timescales).
- The optimal protocol for a given system (requires solving a control problem).
- How credit converts between forms (though conservation constrains the total).

The thermodynamic-length bound (Eq. 10) holds in linear response and may require corrections far from equilibrium.

8 Conclusion

Landauer’s bound is not a fixed tax on computation. It is a lower limit that depends on the entropy removed, which depends on the state credit available. Bias and correlations are thermodynamic resources obeying a conservation law; protocol efficiency reduces waste but is not conserved.

Episodes of ultra-low dissipation or work extraction are withdrawals from the state-credit account, not thermodynamic free lunches. The accounting clarifies what must be measured—entropy, mutual information, thermodynamic length—to predict dissipation in real systems. For biological and engineered information processors, the question is not “how many bits?” but “how much credit, and how good is the protocol?”

References

- Charles H. Bennett. The thermodynamics of computation—a review. *International Journal of Theoretical Physics*, 21(12):905–940, 1982.
- Antoine Bérut, Artak Arakelyan, Artyom Petrosyan, Sergio Ciliberto, Raoul Dillenschneider, and Eric Lutz. Experimental verification of Landauer’s principle linking information and thermodynamics. *Nature*, 483:187–189, 2012.
- Alan A. Cohen, Luigi Ferrucci, Tamás Fülöp, Dominique Gravel, Nan Hao, Andres Kriete, Morgan E. Levine, Lewis A. Lipsitz, Marcel G. M. Olde Rikkert, Andrew Rutenberg, Nicholas Stroustrup, and Ravi Varadhan. A complex systems approach to aging biology. *Nature Aging*, 2:580–591, 2022.
- Gavin E. Crooks. Measuring thermodynamic length. *Physical Review Letters*, 99:100602, 2007.
- L. del Rio, J. Åberg, R. Renner, O. Dahlsten, and V. Vedral. The thermodynamic meaning of negative entropy. *Nature*, 474:61–63, 2011.
- R. Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3):183–191, 1961.
- Juan M. R. Parrondo, Jordan M. Horowitz, and Takahiro Sagawa. Thermodynamics of information. *Nature Physics*, 11:131–139, 2015.

- Takahiro Sagawa and Masahito Ueda. Generalized Jarzynski equality under nonequilibrium feedback control. *Physical Review Letters*, 104:090602, 2010.
- Udo Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics*, 75:126001, 2012.
- David A. Sivak and Gavin E. Crooks. Thermodynamic metrics and optimal paths. *Physical Review Letters*, 108:190602, 2012.
- Ian Todd. Minimal embedding dimension for self-intersection-free recurrent processes. *Information Geometry*, 2025. Under review.
- Shoichi Toyabe, Takahiro Sagawa, Masahito Ueda, Eiro Muneyuki, and Masaki Sano. Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality. *Nature Physics*, 6:988–992, 2010.