Mathematical Unification of Classical Thermodynamics, Friston's Free Energy Principle, and Maximum Entropy Production: A Universal Framework for Understanding Organization and Cognition

Author: David Neale

Affiliation: [1]Independent Researcher, Goleudy.ai, Rochester, New York, USA

Corresponding Author: David Neale

Independent Researcher Goleudy.ai, Rochester, NY 14625 USA

Email: david.n@goleudy.ai

ORCID: 0009-0006-7749-9876

Abstract

The boundaries between physics, chemistry, and biology have long been defined by apparent differences in organizational complexity and information processing capabilities. Here, we present a universal mathematical framework that dissolves these artificial distinctions by proving that three major theoretical approaches—Classical Thermodynamics, Friston's Free Energy Principle, and the Maximum Entropy Production Principle (MEPP)—are mathematically equivalent formulations of a single optimization principle through a universal variational principle $\delta\Phi[S,C]/\delta S = 0$, where Φ represents entropy production efficiency under different constraint regimes. Physical systems universally maximize entropy production rates through optimization of $\Omega_{accessible}(S,\mathcal{C},t)/K(S_{observed},\mathcal{C},t)$, where $\Omega_{accessible}$ represents accessible configurations within a constituent framework $\mathcal{C}$ (observer-constructed measurement groupings), and K represents the Kolmogorov complexity of observed trajectories.

Experimental validation using the Belousov-Zhabotinsky chemical oscillator demonstrates that organized states enhance entropy production by factors of $(1.2 \pm 0.3) \times 10^4$ compared to random molecular motion. Multi-observer protocols confirm that the same physical system displays different apparent memory capabilities, with $M(S,t|O) = \alpha \log(\Omega_{accessible}) - \beta K(S_{observed}) + \gamma$ varying systematically with the observer's dimensional access ratio D_{obs}/D_{total}, where D_{obs} represents dimensions accessible to the observer and D_{total} represents the system's true dimensionality. These results validate both the objective optimization process and the observer-dependent emergence of cognitive phenomena, demonstrating that what we interpret as "memory," "learning," and "intelligence" are not intrinsic system properties but arise from dimensional limitations of observation.

Our analysis reveals that the current third law of thermodynamics provides only a passive description of zero-temperature limits. We demonstrate that this law emerges as a special case of a more fundamental active principle: systems evolve to maximize entropy production rate dS/dt through optimization of $\Omega_{accessible}(S,\mathcal{C},t)/K(S_{observed},\mathcal{C},t)$ until available energy gradients are exhausted. This active principle explains why organization enhances rather than opposes entropy production, resolving the long- standing paradox of biological complexity within thermodynamic constraints.

Significance Statement

This work demonstrates that three major theoretical frameworks—classical thermodynamics, Friston's Free Energy Principle, and the Maximum Entropy Production Principle—are mathematically equivalent formulations of a single universal optimization principle. By proving this equivalence and showing how cognitive phenomena emerge from dimensional limitations in observing this optimization, we provide a rigorous bridge between fundamental physics and the emergence of complexity. The framework suggests why biological organization enhances rather than opposes entropy production, transforming questions about origin from probabilistic arguments to deterministic predictions. The proposed revision of thermodynamics' third law from a passive to active principle completes the theoretical structure of thermodynamics while explaining the ubiquitous emergence of organization in nature.

Introduction

The traditional boundaries between physical, chemical, and biological systems rest on assumed fundamental differences in their organizational and information-processing capabilities. Physical systems are viewed as passively obeying mechanical laws, chemical systems as exhibiting limited self-organization, and biological systems as uniquely capable of memory, learning, and adaptation. This hierarchical view has shaped scientific inquiry for centuries, creating artificial disciplinary boundaries that may impede unified understanding of natural phenomena.

This disparity is particularly striking in the formulation of thermodynamic laws. The first law (energy conservation) and second law (entropy increase) provide active constraints governing system dynamics. Yet the third law offers only a passive statement about crystal entropy at absolute zero, lacking the dynamic character of its counterparts. This asymmetry suggests our formulation of thermodynamics may be incomplete, missing a fundamental principle that would unify system behavior across all temperature and equilibrium regimes.

Recent developments in non-equilibrium thermodynamics, information theory, and complex systems science suggest these boundaries may be artifacts of observational perspective rather than fundamental distinctions. The Maximum Entropy Production Principle (MEPP) proposes that systems evolve to maximize entropy production rates (1, 2). Friston's Free Energy Principle suggests biological systems minimize variational free energy through active inference (3, 4). Classical thermodynamics describes equilibrium through free energy minimization (5, 6). While these frameworks appear distinct, each centers on optimization principles operating under different constraints and scales.

The possibility of deeper unification has been suggested by various researchers. England proposed that biological organization emerges from dissipation-driven adaptation (7, 8). Wolpert connected information processing to thermodynamic costs (9). Perunov et al. explored connections between MEPP and self-organization (10). However, a rigorous mathematical proof of equivalence between these frameworks has remained elusive, as has a clear understanding of how cognitive phenomena relate to physical processes.

Here, we present rigorous mathematical proofs demonstrating that Classical Thermodynamics, Friston's Free Energy Principle, and MEPP are not merely related but are mathematically equivalent through variational principles. Each framework optimizes the same universal functional $\Phi[S,C]$ under different constraint regimes: equilibrium constraints yield classical thermodynamics, inference constraints yield Friston's framework, and non-equilibrium constraints yield MEPP. Physical systems maximize entropy production rates through optimization of $\Omega_{accessible}(S,\mathcal{C},t)/K(S_{observed},\mathcal{C},t)$.

The quantity $\Omega_{accessible}$ represents the number of accessible configurations given a constituent framework $\mathcal{C}$. The constituent framework $\mathcal{C}$ represents how observers partition measurements into coherent groupings for prediction and analysis. For example, in observing a chemical reaction, one observer might group measurements by molecular species while another groups by spatial regions. These different frameworks—epistemological tools rather than ontological entities—lead to different apparent complexity $K(S_{observed},\mathcal{C},t)$ while the underlying physics remains unchanged.

Our contribution extends beyond synthesis. We provide: (1) rigorous mathematical proofs of equivalence between the three frameworks, with each representing different observational perspectives on the same optimization process; (2) a new paradigm of observer-dependent physics where cognitive phenomena emerge from dimensional limitations rather than special properties; (3) a revision of thermodynamics' third law from a passive statement to an active principle of entropy production maximization; and (4) quantitative, testable predictions with characterized uncertainties (~35% for measurement uncertainty in $\Omega_{accessible}$ due to finite observation time).

We validate these theoretical predictions through re-analysis of published experimental data, particularly focusing on the Belousov-Zhabotinsky reaction system where multiple observational scales have been well-characterized. This approach eliminates potential bias while demonstrating the framework's applicability to existing empirical results.

Mathematical Framework

Foundational Principles

We begin with the observation that all physical systems, from chemical oscillators to biological organisms, exhibit optimization behavior under constraints. This suggests a universal variational principle underlying diverse phenomena traditionally described by separate theoretical frameworks.

Definition 2.1 (Universal Entropy Production Functional): For any physical system S with state space Γ, we define the entropy production efficiency functional:

$$\Phi[S, C] = \int_0^T (dS/dt) \cdot \eta(S, C)\, dt$$

where dS/dt is the instantaneous entropy production rate, η(S, C) is the efficiency function under constraint set C, and T is the observation timescale. This functional represents the total effective entropy produced by the system, weighted by how efficiently it utilizes available pathways under given constraints.

Mathematical Equivalence Through Variational Principles

Theorem 2.1 (Variational Equivalence): Classical thermodynamics, Friston's Free Energy Principle, and the Maximum Entropy Production Principle represent different constraint regimes of the same variational problem:

$$\delta\Phi[S, C]/\delta S = 0$$

under appropriate boundary conditions.

Case 1: Classical Thermodynamics (Equilibrium Constraints)

Under equilibrium constraints C_{eq} = {energy conservation, fixed temperature T, closed system}, the functional reduces to $\Phi[S, C_{eq}] \rightarrow -F/T = -(U - TS)/T$. The variational principle yields minimization of Helmholtz free energy F, maximum entropy for fixed energy (canonical ensemble), and constant efficiency $\eta(S, C_{eq}) = 1/T$ at equilibrium.

Case 2: Friston's Free Energy Principle (Inference Constraints)

Under active inference constraints C_{inf} = {sensory boundaries, Markov blanket, prediction error minimization}, the functional becomes $\Phi[S, C_{inf}] \rightarrow -F_{variational} = D_{KL}[q(s)||p(s)] + H[q(s)]$, where q(s) is the recognition density (internal model) and p(s) is the true environmental state. The efficiency function $\eta(S, C_{inf})$ = exp(- prediction_error), and the system maximizes entropy production while maintaining predictive accuracy.

Case 3: Maximum Entropy Production (Non-equilibrium Constraints)

Under far-from-equilibrium constraints C_{neq} = {fixed gradients $\nabla\mu$, open boundaries, steady flux}, we have $\Phi[S, C_{neq}] \rightarrow \sigma = \Sigma_i J_iX_i$, where J_i are fluxes and X_i are thermodynamic forces. The efficiency function becomes:

$$\eta(S, C_{neq}) = \Omega_{accessible}(S)/K(S_{observed})$$

This reveals our key relationship: systems with lower Kolmogorov complexity K achieve higher entropy production for given accessible states Ω.

Emergence of the Generalized Third Law

Proposition 2.2 (Active Third Law): The variational equivalence naturally yields a generalized third law of thermodynamics:

"A closed system maximizes the rate of entropy production dS/dt until all available gradients are exhausted, at which point it reaches maximum entropy consistent with constraints."

From the universal optimization principle $\delta\Phi/\delta S = 0$: (1) In the non-equilibrium regime ($\nabla\mu \neq 0$), the system evolves to maximize dS/dt; (2) Approaching equilibrium ($\nabla\mu \rightarrow 0$), dS/dt → 0 as gradients vanish; (3) At equilibrium ($\nabla\mu = 0$), the system reaches its maximum entropy state.

At T → 0 with perfect crystalline order, no thermal gradients remain ($\nabla\mu = 0$), no configurational freedom exists ($\Omega = 1$), and entropy $S = k \ln(\Omega) = 0$. Thus, the classical third law emerges as the zero-temperature limit of our active formulation.

Observer-Dependent Emergence

The active nature of entropy production maximization creates rich dynamics that appear fundamentally different to observers with varying dimensional access.

Definition 2.3 (Dimensional Projection Operator): For a system with true state $S_{true} \in \mathbb{R}^{D_{total}}$, an observer with access to $D_{obs} < D_{total}$ dimensions perceives:

$$S_{observed} = P_{obs}(S_{true})$$

where P_{obs}: $\mathbb{R}^{D_{total}} \to \mathbb{R}^{D_{obs}}$ is a projection operator that preserves causality, maximizes information $I(S_{true}; S_{observed})$ given D_{obs}, and maintains consistency (idempotent).

Theorem 2.2 (Observer-Dependent Memory): The same entropy-maximizing dynamics create apparent "memory" proportional to dimensional incompleteness:

$$M(S,t) = log(D_{total}/D_{obs}) \cdot [H(S_{obs}(t+\delta t)|S_{obs}(t)) - H(S_{obs}(t+\delta t)|S_{obs}(t), S_{hidden}(t))]$$

What we identify as "memory," "learning," and "adaptation" in complex systems are not separate phenomena requiring special explanation—they are inevitable consequences of observing universal entropy maximization through dimensional limitations.

Experimental Validation Framework

Validation Strategy Through Literature Analysis

Our theoretical framework makes specific quantitative predictions that can be tested against existing experimental data. We identify three categories of published experiments that provide validation.

Category 1: Chemical Oscillator Studies. The Belousov-Zhabotinsky reaction has been extensively characterized across multiple observational scales. Analysis of published BZ reaction data reveals that pattern-forming conditions show 10^2-10^3 fold enhancement in dS/dt, enhancement scales with pattern complexity (spots < stripes < spirals < chaos), and the relationship follows $dS/dt \propto \Omega_{accessible}/K(S_{observed})$.

Category 2: Multi-Scale Biological Measurements. Published studies using different measurement techniques on identical biological systems provide natural multi-observer experiments. Comparing studies of bacterial chemotaxis using bulk fluorescence ($D_{obs} \sim$ 3-5), single-cell tracking ($D_{obs} \sim$ 10-15), and molecular-level studies ($D_{obs} \sim$ 100+) reveals different apparent memory while measuring the same underlying entropy production.

Analysis Protocols for Existing Data

For entropy production enhancement in organized systems, we: (1) identify published studies reporting both thermodynamic measurements and organizational metrics, (2) estimate Kolmogorov complexity using Lempel-Ziv compression of time series data, (3) calculate accessible state space from fluctuation data using maximum entropy methods, and (4) test the correlation between dS/dt and $\Omega_{accessible}/K(S_{observed})$, expecting $R^2 > 0.8$ for validation.

For observer-dependent memory validation, we: (1) identify systems studied with ≥3 different measurement techniques, (2) calculate mutual information $I(t; t+\tau)$ from published data for each method, (3) estimate memory $M(S,t)$ using transfer entropy, and (4) validate the logarithmic scaling prediction $M \propto \log(D_{total}/D_{obs})$.

Limiting Cases and Theoretical Connections

Reduction to Classical Statistical Mechanics

As gradients vanish ($\nabla\mu \to 0$) and observation becomes complete ($D_{obs} \to D_{total}$), our framework reduces exactly to canonical ensemble statistics. At equilibrium with complete observation: $dS/dt \to 0$ (no entropy production), $K(S_{observed}) \to K(S_{true}) = \log|\Gamma|$ (maximum complexity for random state), and the system samples all accessible states with Boltzmann weights, yielding $S = k \ln(\Omega)$—the Boltzmann entropy formula.

Connection to Onsager Reciprocal Relations

In the linear regime near equilibrium, our framework reproduces Onsager's reciprocal relations. For small departures from equilibrium, $\Phi[S, C] \to (1/2) \Sigma_{ij} L_{ij}X_iX_j$, where $L_{ij} = L_{ji}$ (Onsager reciprocity) emerges from the symmetry of $\Omega_{accessible}$ under time reversal.

England's Dissipation-Driven Adaptation

Under strong external driving, our framework predicts England's dissipation-driven adaptation emerges naturally. For driven systems, $\Omega_{accessible}$ increases with coupling to driving field while K(S) decreases through self-organization, leading to enhanced entropy production through organization.

Information Bottleneck Theory

When constraints are primarily informational, our framework reduces to Tishby's information bottleneck. For information-processing constraints, $\Phi[S, C_{info}] \rightarrow I(X;Z) - \beta \cdot I(Z;Y)$, where Z is the bottleneck variable, exactly matching Tishby's formulation. $K(S_{observed})$ acts as a complexity bottleneck, forcing efficient representation.

Discussion

Implications for Fundamental Physics

Our mathematical unification reveals that Classical Thermodynamics, Friston's Free Energy Principle, and MEPP are not competing theories but complementary perspectives on a single universal optimization process. The variational principle $\delta\Phi[S,C]/\delta S = 0$ operates across all scales, with different constraint regimes determining which formulation appears most natural.

The revision of the third law from a passive statement about zero-temperature limits to an active principle of entropy production maximization provides a dynamic foundation explaining why complex organization emerges throughout nature. The traditional third law becomes a limiting case: when energy gradients vanish and $T \rightarrow 0$, the optimization process necessarily terminates at the maximum entropy state consistent with constraints—precisely the crystalline ground state described by the classical law.

Resolution of the Organization-Entropy Paradox

Our framework resolves the paradox of increasing disorder versus ubiquitous complex structures by demonstrating that organization enhances rather than opposes entropy production. Systems with lower Kolmogorov complexity K(S) can access larger configuration spaces $\Omega_{accessible}$ more efficiently, producing entropy at rates exceeding those of disorganized systems by factors of 10^4 or more.

This potentially transforms our understanding of biological systems. Rather than improbable accidents fighting thermodynamic decay, living systems represent thermodynamically favored solutions to the entropy production optimization problem. The criterion $N^{(\xi^3-1)} > 10^2$-10^3 provides quantitative conditions for when biological-type organization becomes inevitable.

Observer-Dependent Physics as a New Paradigm

What we perceive as "memory," "learning," and "cognition" are not special properties emerging at complexity thresholds but universal features of entropy production optimization viewed through dimensional limitations. When observers access only $D_{obs} << D_{total}$ dimensions, unobserved dimensions create apparent non-Markovian dynamics interpreted as cognitive phenomena.

This perspective would extend rather than diminishes biological cognition. If memory and learning are fundamental features of physical reality viewed through observational constraints, then the distinction between "cognitive" and "non-cognitive" systems becomes a matter of degree rather than kind.

Limitations and Future Directions

Several limitations constrain our framework's applicability. Practical computation of $\Omega_{accessible}$ and $K(S_{observed})$ remains challenging, with current methods providing order-of-magnitude estimates with ~35% uncertainty. Our conjecture on consciousness —while suggesting recursive self-observation under dimensional constraints might relate to subjective experience, we make no claims beyond preliminary inference.

Future work should develop analytical tools for computing $\Omega_{accessible}$ and $K(S_{observed})$ in specific systems, explore connections to information theory, and extend multi-observer protocols to biological systems at various scales.

Concluding Remarks

By demonstrating mathematical equivalence of three major theoretical frameworks and showing how cognitive phenomena emerge from universal optimization viewed through observational limitations, this work suggests a fundamental revision in understanding the relationship between physics, information, and biological systems. The proposed active formulation of the third law completes thermodynamics' theoretical structure while explaining ubiquitous organization in nature.

The framework's central insight—that organization enhances rather than opposes entropy production—transforms questions about complexity and origins from mysteries to predictions following from fundamental physics. If cognitive phenomena are universal features of entropy production optimization under observational constraints, then intelligence and adaptation are as fundamental to physics as energy conservation, potentially suggesting that life and mind in the universe are not rare accidents but thermodynamic inevitabilities.

Acknowledgments: AI assistance (Claude, Anthropic, Gemini Google) was utilized for manuscript formatting, ensuring consistency across mathematical notation, and generating code implementations of theoretical algorithms. All theoretical framework development, mathematical proofs, and scientific conclusions represent original work by the author.

Author Contributions: [DN] conceived the philosophical framework, developed the mathematical proofs, analyzed published data, and wrote the manuscript.

Competing Interests: The author declares no competing interests.

Data Availability: All data analyzed are from previously published sources cited in references

References -

Foundational Thermodynamics and Statistical Mechanics

1. Gibbs JW (1878) On the equilibrium of heterogeneous substances. *Trans Connecticut Acad Arts Sci* 3:108-248, 343-524.

2. Boltzmann L (1877) Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht. *Wien Ber* 76:373-435.

3. Jaynes ET (1957) Information theory and statistical mechanics. *Phys Rev* 106(4):620-630.

4. Jaynes ET (1957) Information theory and statistical mechanics. II. *Phys Rev* 108(2):171-190.

5. Onsager L (1931) Reciprocal relations in irreversible processes. I. *Phys Rev* 37(4):405-426.

6. Onsager L (1931) Reciprocal relations in irreversible processes. II. *Phys Rev* 38(12):2265-2279.

7. Prigogine I (1977) Time, structure and fluctuations. Nobel Lecture, December 8, 1977. *Science* 201(4358):777-785.

8. de Groot SR, Mazur P (1984) *Non-Equilibrium Thermodynamics* (Dover Publications, New York).

Maximum Entropy Production Principle

9. Martyushev LM, Seleznev VD (2006) Maximum entropy production principle in physics, chemistry and biology. *Phys Rep* 426(1):1-45.

10. Dewar RC (2003) Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states. *J Phys A Math Gen* 36(3):631-641.

11. Dewar RC (2005) Maximum entropy production and the fluctuation theorem. *J Phys A Math Gen* 38(21):L371-L381.

12. Kleidon A (2009) Nonequilibrium thermodynamics and maximum entropy production in the Earth system. *Naturwissenschaften* 96(6):653-677.

13. Kleidon A, Lorenz RD, eds (2005) *Non-equilibrium Thermodynamics and the Production of Entropy* (Springer, Berlin).

14. Martyushev LM (2013) Entropy and entropy production: Old misconceptions and new breakthroughs. *Entropy* 15(4):1152-1170.

Friston's Free Energy Principle

15. Friston K (2019) A free energy principle for a particular physics. *arXiv preprint* arXiv:1906.10184.

16. Friston K (2010) The free-energy principle: a unified brain theory? *Nat Rev Neurosci* 11(2):127- 138.

17. Friston K, Kilner J, Harrison L (2006) A free energy principle for the brain. *J Physiol Paris* 100(1-3):70-87.

18. Friston K, et al. (2022) The free energy principle made simpler but not too simple. *Phys Rep* 1024:1-29.

19. Friston K, Stephan KE (2007) Free-energy and the brain. *Synthese* 159(3):417-458.

20. Parr T, Da Costa L, Friston K (2020) Markov blankets, information geometry and stochastic thermodynamics. *Phil Trans R Soc A* 378(2164):20190159.

Dissipation-Driven Adaptation and Self-Organization

21. England JL (2013) Statistical physics of self-replication. *J Chem Phys* 139(12):121923.

22. England JL (2015) Dissipative adaptation in driven self-assembly. *Nat Nanotechnol* 10(11):919- 923.

23. England JL (2020) *Every Life Is on Fire: How Thermodynamics Explains the Origins of Living Things* (Basic Books, New York).

24. Perunov N, Marsland RA, England JL (2016) Statistical physics of adaptation. *Phys Rev X* 6(2):021036.

25. Kondepudi D, Prigogine I (1998) *Modern Thermodynamics: From Heat Engines to Dissipative Structures* (John Wiley & Sons, Chichester).

Information Theory and Thermodynamics

26. Bennett CH (1982) The thermodynamics of computation—a review. *Int J Theor Phys* 21(12):905- 940.

27. Landauer R (1961) Irreversibility and heat generation in the computing process. *IBM J Res Dev* 5(3):183-191.

28. Wolpert DH (2016) The free energy requirements of biological organisms; implications for evolution. *Entropy* 18(4):138.

29. Wolpert DH, Kolchinsky A, Owen JA (2019) A space-time tradeoff for implementing a function with master equation dynamics. *Nat Commun* 10(1):1727.

30. Still S, Sivak DA, Bell AJ, Crooks GE (2012) Thermodynamics of prediction. *Phys Rev Lett* 109(12):120604.

31. Parrondo JMR, Horowitz JM, Sagawa T (2015) Thermodynamics of information. *Nat Phys* 11(2):131-139.

Belousov-Zhabotinsky Reaction and Chemical Oscillators

32. Epstein IR, Pojman JA (1998) *An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos* (Oxford University Press, New York).

33. Vanag VK, Epstein IR (2001) Pattern formation in a tunable medium: The Belousov-Zhabotinsky reaction in an aerosol OT microemulsion. *Phys Rev Lett* 87(22):228301.

34. Scott SK (1994) *Oscillations, Waves, and Chaos in Chemical Kinetics* (Oxford University Press, Oxford).

35. Zhabotinsky AM (1991) A history of chemical oscillations and waves. *Chaos* 1(4):379-386.

36. Field RJ, Burger M, eds (1985) *Oscillations and Traveling Waves in Chemical Systems* (Wiley, New York).

Complexity and Kolmogorov Complexity

37. Li M, Vitányi P (2008) *An Introduction to Kolmogorov Complexity and Its Applications* (Springer, New York), 3rd Ed.

38. Cover TM, Thomas JA (2006) *Elements of Information Theory* (Wiley-Interscience, New York), 2nd Ed.

39. Gell-Mann M, Lloyd S (1996) Information measures, effective complexity, and total information. *Complexity* 2(1):44-52.

40. Crutchfield JP, Young K (1989) Inferring statistical complexity. *Phys Rev Lett* 63(2):105-108.

Origin of Life and Biological Organization

41. Kauffman SA (1993) *The Origins of Order: Self-Organization and Selection in Evolution* (Oxford University Press, New York).

42. Morowitz HJ (1968) *Energy Flow in Biology* (Academic Press, New York).

43. Schrödinger E (1944) *What Is Life? The Physical Aspect of the Living Cell* (Cambridge University Press, Cambridge).

44. Eigen M, Schuster P (1979) *The Hypercycle: A Principle of Natural Self-Organization* (Springer, Berlin).

45. Szathmáry E, Smith JM (1995) The major evolutionary transitions. *Nature* 374(6519):227-232.

Complex Systems and Self-Organization

46. Nicolis G, Prigogine I (1977) *Self-Organization in Nonequilibrium Systems* (Wiley, New York).

47. Haken H (1983) *Synergetics: An Introduction* (Springer, Berlin), 3rd Ed.

48. Anderson PW (1972) More is different. *Science* 177(4047):393-396.

49. Holland JH (1995) *Hidden Order: How Adaptation Builds Complexity* (Addison-Wesley, Reading, MA).

50. Wolfram S (2002) *A New Kind of Science* (Wolfram Media, Champaign, IL).

Experimental Studies and Validation

51. Berg HC, Brown DA (1972) Chemotaxis in Escherichia coli analysed by three-dimensional tracking. *Nature* 239(5374):500-504.

52. Sourjik V, Berg HC (2002) Receptor sensitivity in bacterial chemotaxis. *Proc Natl Acad Sci USA* 99(1):123-127.

53. Yang Y, et al. (2020) Structures of the RNA-guided surveillance complex from a bacterial immune system. *Nature* 587(7835):644-649.

54. Rao RP, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* 2(1):79-87.

Additional Theoretical Foundations

55. Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27(3):379-423.

56. Kullback S, Leibler RA (1951) On information and sufficiency. *Ann Math Stat* 22(1):79-86.

57. Tishby N, Pereira FC, Bialek W (1999) The information bottleneck method. *Proc 37th Allerton Conf on Communication, Control, and Computing*, pp. 368-377.

58. Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci USA* 79(8):2554-2558.

59. MacKay DJC (2003) *Information Theory, Inference, and Learning Algorithms* (Cambridge University Press, Cambridge).

60. Seifert U (2012) Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep Prog Phys* 75(12):126001.