PRX LIFE **3**, 037003 (2025)

Roadmap

Physics of Life: Exploring Information as a Distinctive Feature of Living Systems

Stuart Bartlett,[1,2] Andrew W. Eckford,[3] Matthew Egbert,[4] Manasvi Lingam,[5,6,7,8] Artemy Kolchinsky,[9,10,*]
Adam Frank,[11,†] and Gourab Ghoshal[11,12,‡]
[1]*Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California, USA*
[2]*SETI Institute, Mountain View, California, USA*
[3]*Department of Electrical Engineering and Computer Science, York University, Toronto, Ontario, Canada*
[4]*School of Computer Science, University of Auckland, Auckland, New Zealand*
[5]*Department of Aerospace, Physics and Space Sciences, Florida Institute of Technology, Melbourne, Florida, USA*
[6]*Department of Chemistry and Chemical Engineering, Florida Institute of Technology, Melbourne, Florida, USA*
[7]*Department of Mathematics and Systems Engineering, Florida Institute of Technology, Melbourne, Florida, USA*
[8]*Department of Physics, The University of Texas at Austin, Austin, Texas, USA*
[9]*ICREA-Complex Systems Lab, Universitat Pompeu Fabra, Barcelona, Spain*
[10]*Universal Biology Institute, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*
[11]*Department of Physics and Astronomy, University of Rochester, Rochester, New York, USA*
[12]*Department of Computer Science, University of Rochester, Rochester, New York, USA*

(Received 15 January 2025; accepted 11 August 2025; published 3 September 2025)

This paper explores the idea that information is an essential and distinctive feature of living systems. Unlike nonliving systems, living systems actively acquire, process, and use information about their environments to respond to changing conditions, sustain themselves, and achieve other intrinsic goals. We discuss relevant theoretical frameworks such as semantic information and fitness value of information. We also highlight the broader implications of our perspective for fields such as origin-of-life research and astrobiology. In particular, we touch on the transition to information-driven systems as a key step in abiogenesis, informational constraints as determinants of planetary habitability, and informational biosignatures for detecting life beyond Earth. We briefly discuss experimental platforms which offer opportunities to investigate these theoretical concepts in controlled environments. By integrating theoretical and experimental approaches, this perspective advances our understanding of life's informational dynamics and its universal principles across diverse scientific domains.

DOI: 10.1103/rsx4-8x5f

I. INTRODUCTION

Living systems actively sustain and renew themselves despite the natural tendency toward decay, a process sometimes termed *autopoiesis* in the literature [1–3]. Recent research has investigated this concept within broader frameworks of cognition and adaptive behavior [4–9]. Central to this research is the idea that living systems are agents that possess intrinsic goals [10], such as viability (maintenance of the living state), growth, and replication. In fact, the presence of goals that are intrinsic, rather than externally assigned, distinguish organisms from most nonliving systems considered in the natural sciences [11,12].

One important way that organisms achieve intrinsic goals is by changing their behavior in response to different environments [13,14]. In this way, organisms acquire, process, and use information about environmental states for functional purposes. This kind of information usage has been variously termed functional [15,16], meaningful [17], and semantic [14] information in the literature (with subtle differences).

In this paper, we argue that the use of semantic information (SI) is one of the most distinctive and important features of living and protoliving systems [18–23]. In other words, although many nonliving systems are usefully described by information-theoretic measures, organisms are distinguished by actively using information to sustain viability. We term this the "informational perspective".

Existing research in biology and biophysics has shown the power of information-theoretic ideas for studying modern organisms [22,24–28]. Among other achievements, this research has uncovered informational constraints on biomolecular processes at multiple scales, ranging from the genetic code [29] and regulation [30–32] to developmental signaling [33] and chemotactic navigation [34].

Our perspective builds on existing work while highlighting additional research directions. These include developing agnostic signatures for goal-directed information processing in physical systems, understanding its emergence in origin-of-life contexts, and identifying it in astrobiological environments. Our aim is not to propose a novel definition of life but to draw attention to a fundamental and pervasive dimension of living matter [21,26,35–37], complementary

---

[*] Contact author: artemyk@gmail.com
[†] Contact author: afrank@pas.rochester.edu
[‡] Contact author: gghoshal@pas.rochester.edu

to other well-studied features such as compartmentalization [38], natural selection [39], and thermodynamic organization, including metabolic cycles and nonequilibrium maintenance [40–44]. These dimensions are in principle distinguishable; for instance, certain physical systems, e.g., hurricanes, can sustain stable nonequilibrium dynamics without acquiring or using environmental information.

Rather than serving as a definition of life, SI offers a metric for a particular kind of organization, one in which internal dynamics is tuned to extract and use meaningful environmental structure. This framing may help identify transitional regimes between nonlife and life and can guide experimental investigations in marginal systems, such as autocatalytic chemistries or protocellular assemblies.

The framework is empirically grounded: The viability-contributing role of environmental information is already evident across diverse biological domains. In bacterial chemotaxis, for example, cells modulate their movement in response to chemical gradients, directly influencing functional outcomes like survival [34,45]. Neural systems leverage sensory representations to guide decisions under uncertainty, with consequences for organismal fitness [46]. Genetic regulatory networks respond dynamically to environmental signals, modulating gene expression in ways that shape growth and adaptation [24,30]. Even molecular systems, such as ribozymes interacting with peptides, may display emergent information-processing that supports self-organization and persistence [47,48]. These examples demonstrate that the information-viability relationship central to SI is empirically accessible in well-characterized biological settings.

In what follows, we outline emerging information-theoretic approaches for investigating the informational perspective, along with their implications for origin-of-life research, astrobiology, and related fields. The ideas presented here build on discussions from Ref. [49]. The workshop brought together more than 20 researchers across disciplines, spanning artificial life, biophysics, prebiotic chemistry, and planetary science, to explore how the informational perspective might reveal unifying principles of living systems.

II. THEORETICAL FRAMEWORKS

A. Semantic information

The informational perspective calls for a mathematically precise and scientifically applicable theory of semantic information. One such theory was introduced in Ref. [14]. The approach begins by defining state spaces and probability distributions for an agent X and its environment Y, possibly at some coarse-grained level of description. Correlations between agent and environment may be quantified by the mutual information $I(X;Y)$, as usual in information-theoretic treatments. The agent's ability to maintain its own existence is quantified by a viability function. By scrambling some of the correlations between the agent and its environment, the viability function quantifies which part of the overall mutual information is essential for survival. Semantic information refers to that part of the overall correlated randomness that influences the agent's viability.

Importantly, the viability function is usually not an externally imposed utility function, but rather it is defined an emergent property of the intrinsic dynamics of an agent coupled to an environment [14,50]. For instance, viability may be defined in terms of the agent's ability to resist equilibration (maintain low entropy) or to avoid absorption into a deathlike dynamical attractor [51]. The functional correlations detected by SI may also be considered as intrinsic to the system under study, as long as the statistical ensemble over X and Y captures actual statistics of agent and environment states. Such ensembles may be constructed from trajectory measurements, e.g., from steady-state dynamics, repeated experiments, populations of agents, or models informed by domain knowledge [30,52,53].

Semantic information is inherently environment-dependent, since the same agent may exploit different correlations in different environments. Furthermore, unlike regular mutual information, SI is an asymmetric measure of coupling, because the correlations that are causally relevant to the maintenance of agent X are not necessarily relevant to the maintenance of environment Y. However, the distinction between agent and environment is not fixed, but defined operationally: an agent can be any subsystem whose long-term persistence (viability) can be meaningfully measured and perturbed.

This operational view accommodates nested structures: An agent could be a single cell, a population, or even a planetary-scale biosphere, so long as viability can be meaningfully defined. This flexibility is especially relevant in ecological and astrobiological settings, where clear-cut boundaries may be absent but informational dependencies remain measurable. An interesting direction for future research is to study whether natural agent-environment distinctions can be identified automatically from data.

An important example of SI is provided by the functional information stored in the genome. Here we may consider correlations in an ensemble of genetic sequences X and different environments Y, given some way of quantifying the viability of different genomes in different environments. In practice, viability in different environments is often measured using knockout experiments, which quantify the effect of genetic perturbations on growth and survival [54,55]. Other approaches exploit statistical measures, such as conserved sequences and genome-environment correlations, as proxies [26,56,57]; such observational studies exploit natural knockout experiments performed by mutation and selection in past populations.

At the same time, the framework of SI goes beyond functional genetic information. For one, it can be applied to systems that do not possess genomes and may not undergo replication or Darwinian evolution, e.g., minimal protocells. Moreover, by appropriate choices of statistical ensemble, SI can be used to study information encoded in the fluctuation of a single agent and environment over time. In this, it differs from genetic information, which is typically defined in relation to an (actual or possible) ensemble of different genomes in different organisms, i.e., a population of agents. Distinguishing these concepts has practical implications: While genetic information is defined over ensembles of replicating systems, semantic information can be evaluated for the trajectory-level dynamics of single agents as they maintain

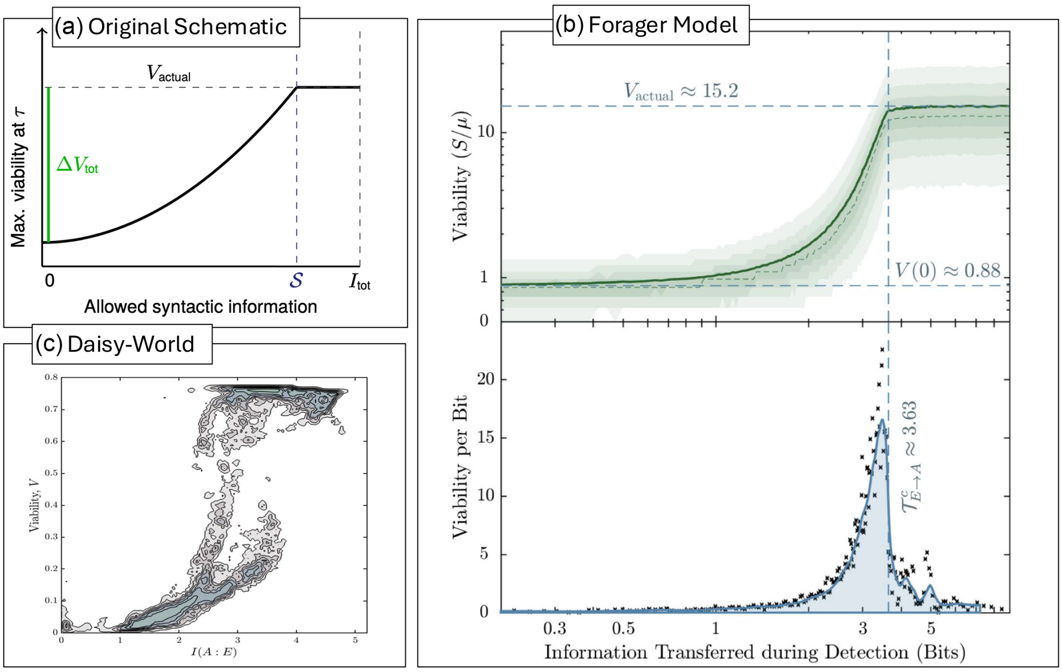

FIG. 1. Semantic thresholds. The SI identifies semantic thresholds where only specific mutual information between an agent and its environment impacts viability. (a) The foundational study [14] introduced this concept through counterfactual simulations assessing viability as information is scrambled. (b) Forager models [58] demonstrated such thresholds by adding noise to food detection sensors, revealing a peak in viability per bit of information at the threshold. (c) Similar thresholds appear in biosphere models like Daisy World, responding to stellar forcing [59].

viability through real-time interactions with their environments, thus encompassing artificial cells, neural systems, and novel biosignatures.

Finally, SI can be used in an agnostic fashion to study information stored not only in genetic sequences, but in a wide variety of substrates, both internal (e.g., epigenetic, allosteric, compositional, and neural information) and external (e.g., niche construction and collective constraints).

At a high level, SI may be considered as a kind of "information knockout" that can detect functional correlations in a wide variety of systems, without knowledge of the precise mechanism by which these correlations affect viability or which precise degrees of freedom encode these correlations. For deeper insights, mechanism identification and localizations of SI can be done by testing different types of perturbations and by integrating with knowledge of the underlying causal architecture.

Recently, this framework has been explored in several modeling studies (see Fig. 1). In a foraging model [58], viability was quantified as the expected lifetime of a forager. By introducing noise into the forager's sensory inputs, SI was quantified as the subset of environmental correlations critical for maintaining viability. One of the key insights from [58] is the discovery of a viability threshold, a plateau in the viability curve where certain correlations between an agent and its environment do not influence survival. Below this threshold, some of the information is purely syntactic, devoid of semantic value. Above the threshold, survival declines monotonically as noise increases, indicating that only a subset of environmental correlations holds semantic significance. The mechanistic cause of the viability threshold was shown to be the geometry of the forager; in particular, its finite size imposes a resolution limit on the useful information that can be extracted from sensory input.

Similar thresholds have also been identified in other contexts. In the Daisy World model [59], such thresholds emerged within the biosphere-planet feedback mechanisms that regulate planetary conditions, relating SI to ecological stability. In networks of coupled Kuramoto-style oscillators [60], the emergence of semantic thresholds was shown to depend on the underlying network topology, emphasizing how both structural and dynamical aspects shape SI.

Recent work has proposed extending these theoretical insights into experimental domains. Synthetic cells have been proposed as a novel platform to investigate SI [61]. Synthetic biology and molecular communication techniques enable the construction of programmable systems that encode specific chemical signals capable of inducing functional changes in a receiver system [32,62,63]; such systems provide a way to study SI (goal-oriented changes triggered by a message) by observing controlled self-organization and adaptive responses in synthetic cells. Active matter provides another promising platform for experimental investigation of SI theory, given recent work on minimal nonbiological particles that can acquire and process information [64–67].

B. Fitness value of information

As mentioned, organisms gather and process information to adapt their features and behaviors, enhancing viability and reproduction in dynamic conditions. One way in which this

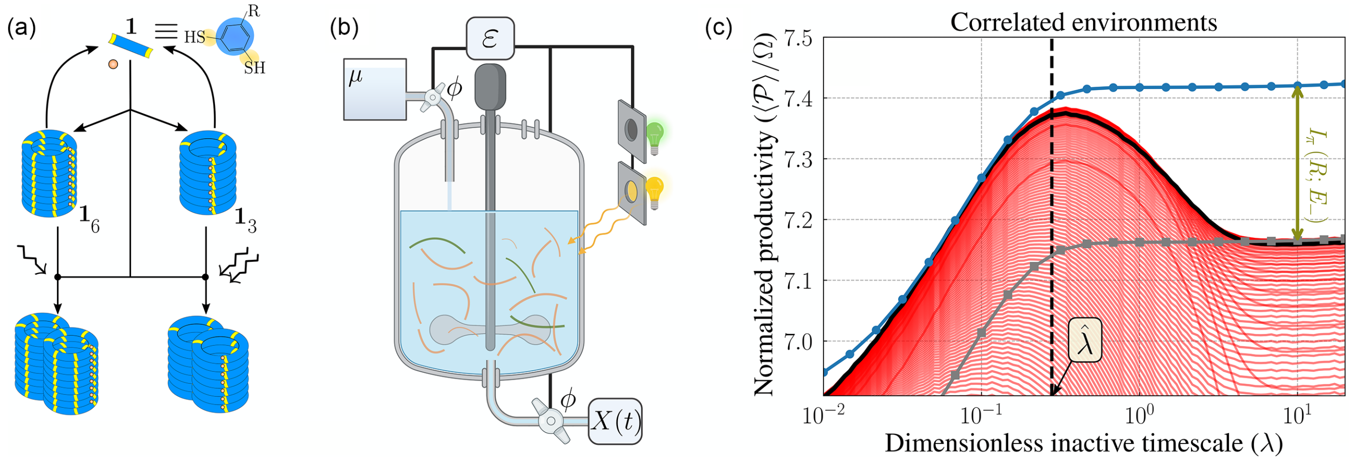

FIG. 2. SI in a flow reactor: a proposed experiment for studying SI inspired by the fitness value of information, from Ref. [72]. (a) The system consists of a population of simple replicators, such as the photocatalytic molecular replicators synthesized by Liu *et al.* [83]. (b) The replicators are placed in a flow reactor and subjected to fluctuating environments which favor different replicators, e.g., weak (green) light or strong (orange) light. The population is also allowed to reequilibrate during inactive periods (no growth). (c) Productivity (replicator production per time) depends on environmental statistics as well as internal parameters, such as exchange reactions and reequilibration timescale λ. At intermediate timescales, the system's memory may provide a source of side information, increasing productivity in proportion to the mutual information between successive environments [72].

can occur is when populations deploy different phenotypes in order to ensure survival in fluctuating environments [68].

When the fit between phenotypes and environments affects the multiplicative growth rate, the optimal population strategy in uncertain environments is bet hedging, which was first derived information theoretically by Kelly [69]. Furthermore, when the phenotypic response can depend on an external signal or cue Z [70,71], the mutual information $I(Z;Y)$ between environment states Y and cues Z controls the increase of maximal growth rate. In such cases, which sometimes go under the name of "fitness value of information" in the literature [70,71], there is a direct quantitative relationship between information and functional outcomes (growth) [72–74]. In this sense, the fitness value of information can be seen as a special kind of SI.

This approach is further generalized by rate-distortion theory [75], which introduces a distortion function $d(x,z)$ to quantify the cost of mapping environment x to cue z. For example, $d(x,z)$ might represent the negative logarithmic growth rate when an organism observes cue z in environment x. While bet hedging directly maximizes the growth rate [70], rate-distortion theory provides a broader framework for analyzing diverse strategies. For instance, an organism requiring multiple essential but noninterchangeable nutrients assigns different values to sensing each nutrient based on its current state [76,77]. This framework has been considered in both information-theoretic and biophysical contexts [71,78,79] and it suggests that living systems optimize rate-distortion limits [27,80]. Organisms near these limits exhibit efficient trade-offs, using limited resources to harvest SI that most contributes to increasing growth [27].

Promising experimental tools include chemostats (continuous flow reactors), which measure growth rates in controlled environments [81,82]. Such systems could be adapted to study the fitness value of information in synthetic and biological systems. In this setting, a recent work [72] proposed a theoretical and experimental framework for studying the trade-offs between information and replicator production (see Fig. 2 for the experimental setup). The framework is applicable to modern microbial organisms as well as minimal molecular replicators.

III. IMPLICATIONS FOR ORIGIN-OF-LIFE RESEARCH

The informational perspective shifts the focus of origin-of-life research. Rather than emphasizing the emergence of specific molecules or structures, e.g., RNA, ribosomes, or metabolic pathways, it emphasizes the transition from information-neutral systems to systems with SI, that is, systems that maintain their viability by sensing and responding to their environments and acting as agents in service of their own self-maintenance [5,13]. Identifying the necessary and sufficient conditions for such transitions suggests new possibilities for understanding abiogenesis, not only on Earth [84] but also on Mars, subsurface ocean worlds, exoplanets, and even in artificial systems [42,85].

In fact, models of emergent learning and adaptive behavior in protocells or chemical systems illustrate how simple systems may leverage information for functional purposes. In other words, even simple systems can exhibit minimal SI, responding to environmental perturbations in ways that prolong their existence [13], thereby suggesting that SI may have emerged during the early stages of life, potentially facilitating other critical transitions in abiogenesis [42,86]. If extant life on Earth is one of many possible instantiations of living systems [84], then extraterrestrial life and even digital life could represent additional members of this category [42,85] and could evince similar features.

Recent proposals suggest that origins experiments could focus on detecting complexity and information processing rather than specific molecules or replication [87]. This could

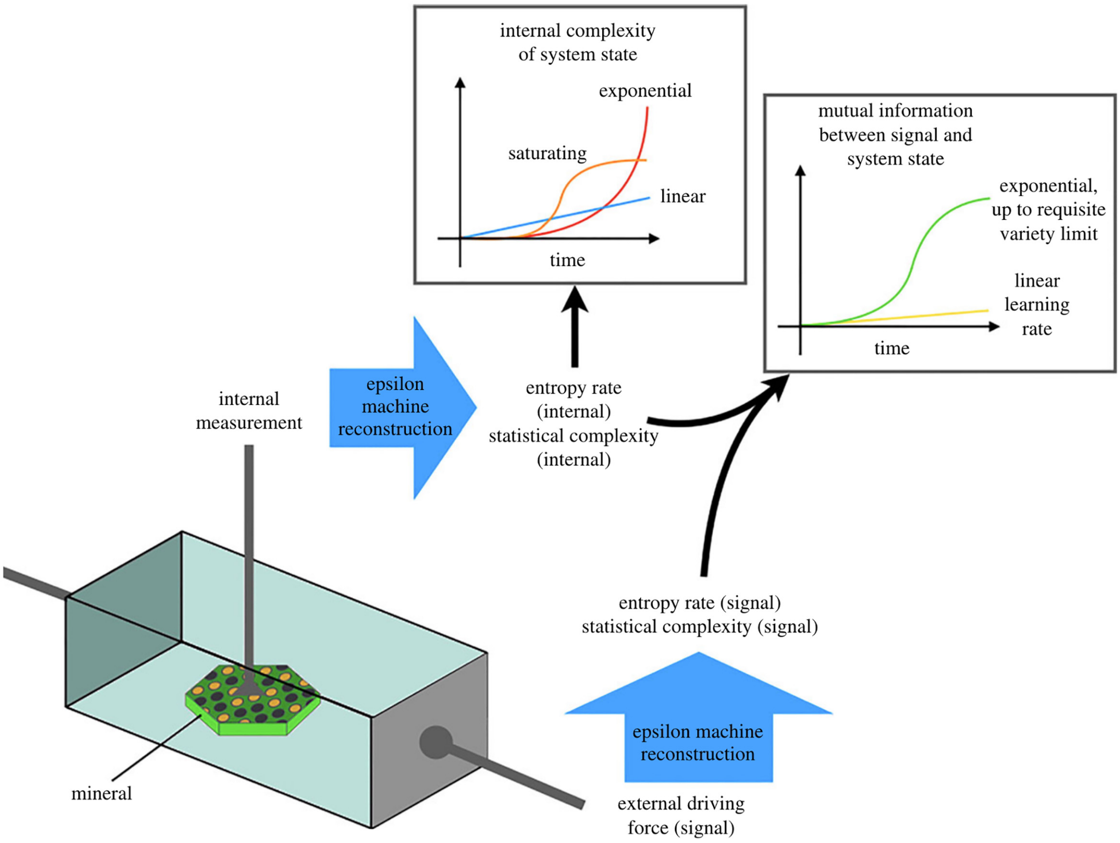

FIG. 3. Example schematic for an origin-of-life experiment focusing on information measures. Consult the text for full details. (Figure was reproduced with permission from [87].)

involve testing whether a system adapts to changing conditions by increasing its statistical complexity [88], indicating the ability to process and learn from information. Such learning has been observed in diverse systems [89,90], and this information-driven approach could guide experimental and computational investigations into origins [91,92].

For example, Fig. 3 presents a potential origin-of-life experiment, which focuses on information and complexity. In this setup, a growing chemical garden is driven away from equilibrium by electrodes, which produce a time-varying electrochemical gradient across the system; note that such gradients may have been crucial in facilitating the origin of life [93,94]. The variability of this electrochemical potential is controlled automatically according to an "epsilon machine". An epsilon machine is a hidden-variable dynamical model with associated measures of dynamical complexity, such as statistical complexity and entropy rate [88]. Measurements of internal state variables are taken, producing a time series. This time series also has an associated statistical complexity and entropy rate, derived through the algorithm of epsilon machine reconstruction.

Note that epsilon machines are employed here as controlled and tunable generators of structured temporal correlations. This allows the modulation of the informational complexity of environmental inputs in a precise and quantifiable way command such that the complexity of the environment and the emergent information processing systems within the experimental system, if any, can be compared.

The premise of this experiment is to see whether the system can transition to being information driven, if we incrementally increase the complexity of the external driving force, and observe whether the internal complexity changes as well. If no such information transition happens, there will be little to no mutual information between the system's response and the driving force (the electrochemical gradient). On the other hand, if information, or perhaps SI, becomes a determining factor in this system, then the internal statistical complexity will track the external statistical complexity in some way (the system will learn) [90]. This learning may be approximately linear (internal statistical complexity is a linear function of time, if the external statistical complexity increases linearly with time) or it might exhibit more exotic dynamics, such as starting out slow and then exhibiting exponential behavior, before saturating into an overall sigmoid. This could be due to the system learning as much as is feasible, reaching the so-called limit of requisite variety [95].

In this experiment, information may be said to drive a system's dynamics once certain conditions are met: (a) The environment presents learnable features, (b) the system is capable of information processing (e.g., protocells with basic learning capabilities [92]), and (c) feedback mechanisms (selection pressures) connect information processing to increased viability. These mechanisms could enhance access to free energy [96] or other resources for maintaining a nonequilibrium state, such as recognizing temporal patterns in energy availability or avoiding adverse conditions.

The primary goal of this experimental setup is to explore whether physical or chemical systems can exhibit information-driven behavior in response to structured external stimuli. Real abiotic environments may indeed contain rich statistical patterns, such as day-night cycles, oscillating chemical gradients, correlations that constitute associative learning tasks, or redox fluctuations. The epsilon machine serves as a way to explore how increasing environmental complexity affects the internal dynamics and possible learning behaviors of a candidate prebiotic system. In this sense, the epsilon machine acts as an abstraction of environmental structure, enabling experimental tests of SI acquisition and response without requiring a fully specified planetary model.

The informational perspective could also bring new insights to classic origins experiments command such as the RNA world and ribozyme replicator systems. These experiments normally comprise strands of RNA that are capable of some function, such as the catalysis of their own replication (ligation of smaller RNA fragments to form a new generation of copies). Additional environmental challenges or selective forces can be introduced, such as the need to bind with certain proteins to form complexes (protoribosomes), or coevolution with primitive protocellular compartments, such as microdroplet coacervates [97,98]. The products of these experiments are typically assessed by comparison with known cellular functions, primarily the minimal conditions for Darwinian evolution (replication, heredity, selection, or other required functions) [16,99,100]. However, this bias carries a significant risk of missing key emergent entities or phenomena that could have been essential to life's origins, but bore little to no resemblance to life as we know it (or even a simplified form of life as we know it).

We can instead screen such primordial systems for SI. For example, we can take a system of RNA strands with a source of peptides, such as might be used to explore the emergence of the ribosome. Normally, one would seek structures similar to the ancient core of the known ribosome [101]. However, the experimental system may exhibit a range of self-organized structures, including many that are unfamiliar from the perspective of known biology. In this case, the informational perspective offers a new lens: The system could be analyzed through a range of information-scrambling interventions. The interventions would erase the information within different length segments of the RNA molecules (randomize their sequences of bases).

Additionally, one could scramble the sequences of amino acids in the encoded peptide molecules. The resulting intervened system could then be assessed to determine whether disrupting informational sequences at this stage of the system's protolife cycle is critical for maintaining the viability of its replicators or collectives of replicators. In the case of high SI content, such interventions would disrupt self-replication, information retention, or other viability-supporting behaviors, leading to a general breakdown of structure toward equilibrium. Interventions targeting different RNA subsequences or peptide segments would reveal which regions are most critical to the system's integrity and function.

In these experiments, SI would be estimated by systematically perturbing the system's structured components (e.g., RNA or peptide domains) and quantifying the resulting loss in viability, such as diminished self-replication or irreversible loss of cooperation between replicators. The more severe and specific the loss under particular scrambling operations, the more semantically meaningful those components are to the system's persistence.

Conversely, if the system exhibits low SI, such scrambling interventions would have little discernible effect on its structure or dynamics, suggesting that any observed functionality is driven by simpler, lower-level physical or chemical constraints. Large-scale experimental programs of this kind, deemphasizing predetermined biological targets and instead probing the informational relevance of system components, could uncover alternative pathways from molecules to life [102].

Given the huge combinatorial space of prebiotic chemistry and the deep challenges of taking a direct path from abiotic processes to a primitive cell, SI could be a powerful guiding beacon, highlighting unorthodox routes from nonlife to life as we know it and even to alternative forms that could have formed on Earth but did not or forms that could emerge on other worlds.

IV. IMPLICATIONS FOR ASTROBIOLOGY

Astrobiology is a rapidly emerging field dedicated to addressing the following profound question: Are we alone? [103–105]. Defined by NASA as the study of the "origin, evolution, distribution, and future of life in the Universe" [106], the field aims to uncover life's potential elsewhere. Two main areas of research include understanding planetary habitability, that is, the potential of different extraterrestrial environments to support life [42,103,107–109], and developing reliable methods for identifying biosignatures, the indicators of extant or extinct life [42,110–112].

In the astrobiological context, the informational perspective elucidates the constraints and possibilities for the existence (habitability) and detection (biosignatures) of life, thereby adumbrating a unifying framework that may apply both on Earth and in extraterrestrial settings.

A. Informational constraints on habitability

Two central facets of information processing involve sensing [113–115] and transmitting [63,116,117] information. Microorganisms, due to their simplicity and ubiquity, offer a key case study for understanding these processes. Mechanisms like chemotaxis, phototaxis, and thermotaxis allow microbes to detect gradients of chemicals, light, and temperature, respectively, offering critical insights into their environments [118–120]. These sensory capabilities enable behaviors such as nutrient acquisition, toxin avoidance, biofilm formation, and symbiosis [120–123]. In parallel, molecular communication through signaling molecules [32,62,124] fosters cooperation, cognition, and multicellularity [7,9,117,125–128].

Potential habitats for the origins, evolution, and sustenance of life span diverse physicochemical conditions, ranging from hydrothermal vents to soda lakes and pumice rafts [42,129–132] on Earth. Understanding how environmental parameters (e.g., temperature and viscosity) shape informational

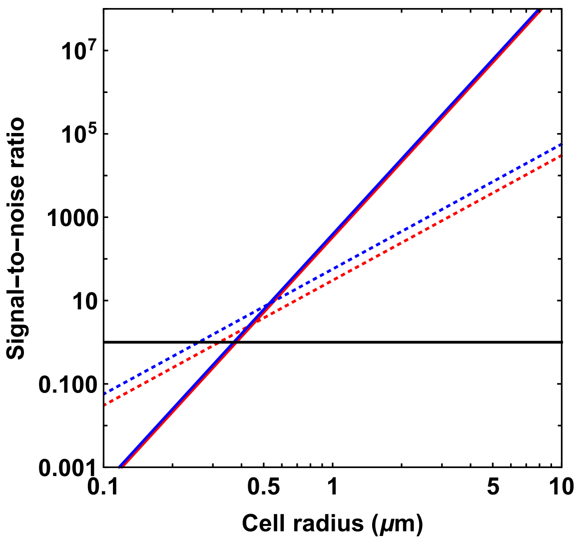

FIG. 4. Basic SNR achievable for chemosensing plotted as a function of cell radius for Earth's oceans (blue) and the hydrocarbon lakes of Saturn's moon Titan (red); the dotted and solid (bold) lines correspond to spatial and temporal modes of sensing, respectively [134]. A heuristic minimum cell radius may be inferred by determining when the SNR exceeds the value of unity (horizontal black line).

constraints is crucial for discerning which settings are conducive to the existence and sustenance of information-driven life.

This intersection of biophysics and astrobiology, with explicit connections to planetary environments [105,133], remains distinctly underexplored [42]. For instance, Ref. [134] modeled heuristic lower bounds on cell sizes capable of gradient sensing and motility in various environments. By performing simple signal-to-noise ratio (SNR) calculations, which can be further extended to include memory effects [26], a rough cell-radius limit of $R_{\min} \sim 0.3$–0.4 μm for chemosensing in Earth's oceans was estimated, consistent with empirical bounds of approximately 0.4 μm [135]. Similar calculations for Titan's hydrocarbon lakes revealed comparable thresholds, despite the extreme conditions. Figure 4 depicts these size limits for Earth and Titan.

Another study [136] estimated data rates $\mathcal{I}$ for molecular communication via signaling molecules between microbes across habitats. The $\mathcal{I}$ and its dependence on intercellular distances d were demonstrated to be broadly consistent with certain laboratory experiments [137] and numerical simulations [138]. The findings of this work, spanning Earth's upper oceans ($\mathcal{I} \sim 4.7 \times 10^{-2}$ bits/s) to Titan's lakes ($\mathcal{I} \sim 2.6 \times 10^{-1}$ bits/s), illustrated how intracellular distances and environmental variables influence informational dynamics and were subsequently harnessed to explore ramifications for Earth's biosphere [139].

These sample studies highlight the putative major role of informational constraints in shaping habitability.

B. Information-centric biosignatures

Agnostic biosignatures, which are designed in principle to transcend Earth-centric assumptions, have emerged as a promising tool [105,140,141]. These approaches prioritize universal features of life over specific molecular markers, addressing the challenges of abiotic false positives [111,142,143]. For instance, abiotic processes can mimic cell-like morphologies [144,145], complicating microfossil identification. Agnostic biosignatures aim to avoid such pitfalls by focusing on life's core characteristics, making them particularly valuable for detecting non-Earth-like biomarkers [85,105].

Proposals for agnostic biosignatures encompass both *in situ* analyses and remote sensing. These include binding pattern analysis of nucleic acid molecules [146], abundance distribution of monomers (e.g., amino acids) [147], assembly theory for molecular construction steps [148], and machine learning for distinguishing abiotic and biological materials [149]. Information-centric approaches, such as epsilon machine reconstruction [150] and Jensen-Shannon divergence for spectral analysis [151], have likewise shown promise in identifying biosignatures through time series data and atmospheric spectra.

Moreover, *in situ* strategies, such as the "poke it and look for a response" method, could leverage dynamic responses to stimuli to differentiate living systems from abiotic materials. Possible measures in this context include information flow [152,153] and information efficiency [154,155], each of which can perhaps distinguish living organisms from dead ones or abiotic materials.

We emphasize that the biosignature approaches discussed here differ significantly in their experimental tractability and theoretical grounding. While some rely on general information-theoretic metrics (e.g., compressibility, redundancy), others—particularly machine learning classifiers, may depend on specific training datasets or biochemical priors. This heterogeneity reflects the current exploratory state of the field. Our aim in surveying these methods is not to endorse any particular one, but to highlight a growing interest in moving beyond narrow biochemical definitions of life toward more abstract, information-based criteria.

By integrating such approaches with informational constraints on habitability, astrobiologists can refine their methods to account for the unique challenges posed by extraterrestrial environments. For example, cell size thresholds derived from information sensing models can serve as heuristic filters in biosignature evaluation.

V. SUMMARY AND FUTURE DIRECTIONS

This paper explored the informational perspective, which posits that life's unique ability to acquire, process, and utilize information is a fundamental dimension that distinguishes it from non-life. This perspective complements existing approaches by emphasizing underlying principles rather than specific functions or mechanisms. We reviewed several emerging formalisms, such as SI and the fitness value of information, that provide quantitative method for studying the relationship between information, function, and self-maintenance. We also discussed implications for ongoing research in fields such as origin-of-life, astrobiology, and artificial life.

We acknowledge that different perspectives exist regarding the relationship between semantic information and more established information-theoretic constructs such as mutual information, as well as biologically grounded notions like fitness and genetic encoding. While our framework builds upon and intersects with these concepts, it is intended to extend them by tying information to function in a system-specific, coarse-grain-sensitive, and viability-dependent manner. Rather than offering a replacement, we aim to provide a complementary and operational tool kit that can clarify the role of information in contexts where traditional notions may prove insufficient, particularly in origins-of-life and agnostic biosignature settings.

Further progress may exploit recent advances in fields like active matter, synthetic biology, and molecular-communication systems. These developments offer powerful test beds for constructing experimental systems that acquire and use information for functional purposes, while exhibiting simple forms of adaptation, learning, and feedback control in complex environments. Such work may provide experimental measurements of functional information flow and reveal novel forms of information-driven nonequilibrium organization.

A promising direction for future work is to extend the information-theoretic frameworks discussed in this paper to evolving and open-ended systems, where the relevant state space may change dynamically due to the emergence (or destruction) of new agents, niches, and evolutionary transitions. Another theoretical avenue would explore deeper connections to information-theoretic fundamentals, including the formulation of SI and fitness value of information using algorithmic (Kolmogorov) information theory [156]. Algorithmic measures may allow for the data-driven quantification of functional information in individual trajectories, without requiring the definition statistical ensembles, and they may uncover connections to existing measures of complexity [157]. We note that, although most measures of algorithmic information are uncomputable, in practice, useful estimates can be calculated using real-world compression algorithms [158,159].

As with many information-theoretic measures, quantifying semantic information presents methodological challenges that are also opportunities for future work. For instance, calculating SI depends on how the system is coarse grained: Different choices in data collection, labeling, or state representation can yield substantially different results. In experimental or observational settings, this dependence on modeling decisions poses limitations, but also points to an avenue for progress. Given an accurate model, the effects of coarse graining can in principle be bounded or estimated, and this could be integrated into experimental design protocols.

Quantifying a system's viability may also depend on extended temporal dynamics. In such cases, the semantics of the information may themselves depend on the evolving internal state of the system, making the functional relevance of structure time dependent. This raises further questions about how best to define and measure viability in systems where environmental or internal fluctuations occur on long timescales [51].

Finally, calculating SI can become computationally expensive or even intractable when the state space or the set of perturbations is large. Addressing these challenges will require careful methodological choices, potentially involving heuristics, approximations, or surrogate metrics. Developing such tools is an important direction for advancing the practical utility and empirical grounding of semantic information.

In the context of astrobiology, an important open challenge is to adapt the frameworks of SI and fitness value of information to remote sensing, such as the detection of planetary biosignatures. For instance, while the original definition of SI is based on causal interventions, future work may develop observation measures that serve as useful proxies for detecting information-driven systems at the planetary scale. This could inform mission-design strategies and support the search for life in nonterrestrial and unconventional forms.

ACKNOWLEDGMENTS

G.G., A.F., and A.K. acknowledge support through Grant No. 62417 from the John Templeton Foundation. A.F. and M.L. acknowledge support from the NASA Exobiology program under Grant No. 80NSSC20K0622. The opinions expressed in this publication are those of the author(s) and do not necessarily reflect the views of the Foundation. A.K. was partly supported by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 101068029. M.E. acknowledges and appreciates the support of the University of Auckland's Research and Study Leave Programme. A.W.E. acknowledges support from Discovery Grant No. RGPIN-2016-05288 from the Natural Sciences and Engineering Research Council.

DATA AVAILABILITY

No data were created or analyzed in this study.

---

[1] H. R. Maturana and F. J. Varela, *Autopoiesis and Cognition: The Realization of the Living* (Springer, Dordrecht, 1980).

[2] F. J. Varela, in *Principles of Biological Autonomy*, 2nd ed., edited by E. A. Di Paolo and E. Thompson (MIT Press, Cambridge, 2025).

[3] A. Frank, M. Gleiser, and E. Thompson, *The Blind Spot: Why Science Cannot Ignore Human Experience* (MIT Press, Cambridge, 2024).

[4] R. D. Beer, The cognitive domain of a glider in the game of life, Artif. Life **20**, 183 (2014).

[5] X. E. Barandiaran and M. D. Egbert, Norm-establishing and norm-following in autonomous agency, Artif. Life **20**, 5 (2014).

[6] R. D. Beer and E. A. Di Paolo, The theoretical foundations of enaction: Precariousness, Biosystems **223**, 104823 (2023).

[7] P. Lyon, The cognitive cell: Bacterial behavior reconsidered, Front. Microbiol. **6**, 264 (2015).

[8] F. Baluška and M. Levin, On having no head: Cognition throughout biological systems, Front. Psychol. **7**, 902 (2016).

[9] J. A. Shapiro, All living cells are cognitive, Biochem. Biophys. Res. Commun. **564**, 134 (2021).

[10] X. E. Barandiaran, E. Di Paolo, and M. Rohde, Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action, Adapt. Behav. **17**, 367 (2009).

[11] S. A. Kauffman, *A World Beyond Physics: The Emergence and Evolution of Life* (Oxford University Press, Oxford, 2019).

[12] P. Ball, *How Life Works: A User's Guide to the New Biology* (University of Chicago Press, Chicago, 2023).

[13] M. Egbert, M. M. Hanczyc, I. Harvey, N. Virgo, E. C. Parke, T. Froese, H. Sayama, A. S. Penn, and S. Bartlett, Behaviour and the origin of organisms, Orig. Life Evol. Biosph. **53**, 87 (2023).

[14] A. Kolchinsky and D. H. Wolpert, Semantic information, autonomous agency and non-equilibrium statistical physics, Interface Focus **8**, 20180041 (2018).

[15] J. Collier, Information in biological systems, Philos. Inf. **8**, 763 (2008).

[16] R. M. Hazen, P. L. Griffin, J. M. Carothers, and J. W. Szostak, Functional information and the emergence of biocomplexity, Proc. Natl. Acad. Sci. USA **104**, 8574 (2007).

[17] C. L. Nehaniv, D. Polani, K. Dautenhahn, R. te Boekhorst, and L. Canamero, in *Artificial Life VIII: Proceedings of the Eighth International Conference on Artificial Life*, edited by R. Standish, M. A. Bedau, and H. A. Abbass (MIT Press, Cambridge, 2002), pp. 345–349.

[18] G. Schlosser, Self-re-production and functionality, Synthese **116**, 303 (1998).

[19] J. M. Smith and E. Szathmary, *The Origins of Life: From the Birth of Life to the Origin of Language* (Oxford University Press, Oxford, 2000).

[20] M. Mossio, C. Saborido, and A. Moreno, An organizational account of biological functions, Brit. J. Philos. Sci. **60**, 813 (2009).

[21] P. Godfrey-Smith and K. Sterelny, *Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta and U. Nodelman (Metaphysics Research Lab., Stanford University, 2007).

[22] R. A. Gatenby and B. R. Frieden, Information theory in living systems, methods, applications, and challenges, Bull. Math. Biol. **69**, 635 (2007).

[23] S. I. Walker and P. C. W. Davies, The algorithmic origins of life, J. R. Soc. Interface **10**, 20120869 (2013).

[24] G. Tkačik and W. Bialek, Information processing in living systems, Annu. Rev. Condens. Matter Phys. **7**, 89 (2016).

[25] R. E. Ulanowicz, Information theory in ecology, Comput. Chem. **25**, 393 (2001).

[26] C. Adami, *The Evolution of Biological Information: How Evolution Creates Complexity, from Viruses to Brains* (Princeton University Press, Princeton, 2024).

[27] W. Bialek, *Biophysics: Searching for Principles* (Princeton University Press, Princeton, 2012).

[28] S. A. Frank, Natural selection. V. How to read the fundamental equations of evolutionary change in terms of information theory, J. Evol. Biol. **25**, 2377 (2012).

[29] T. Tlusty, A colorful origin for the genetic code: Information theory, statistical mechanics and the emergence of molecular codes, Phys. Life Rev. **7**, 362 (2010).

[30] G. Tkacik, C. G. Callan, and W. Bialek, Information flow and optimization in transcriptional regulation, Proc. Natl. Acad. Sci. USA **105**, 12265 (2008).

[31] G. Tkačik, C. G. Callan, Jr., and W. Bialek, Information capacity of genetic regulatory elements, Phys. Rev. E **78**, 011910 (2008).

[32] A. Rhee, R. Cheong, and A. Levchenko, The application of information theory to biochemical signaling systems, Phys. Biol. **9**, 045011 (2012).

[33] T. Gregor, D. W. Tank, E. F. Wieschaus, and W. Bialek, Probing the limits to positional information, Cell **130**, 153 (2007).

[34] H. H. Mattingly, K. Kamino, B. B. Machta, and T. Emonet, *Escherichia coli* chemotaxis is information limited, Nat. Phys. **17**, 1426 (2021).

[35] K. D. Farnsworth, J. Nelson, and C. Gershenson, Living is information processing: From molecules to global systems, Acta Biotheor. **61**, 203 (2013).

[36] J. M. Smith, The concept of information in biology, Philos. Sci. **67**, 177 (2000).

[37] E. Mayr, *What Evolution Is* (Basic Books, New York, 2001).

[38] J. P. Schrum, T. F. Zhu, and J. W. Szostak, The origins of cellular life, Cold Spring Harbor Perspect. Biol. **2**, a002212 (2010).

[39] J. M. Smith and E. Szathmary, *The Major Transitions in Evolution* (Oxford University Press, Oxford, 1997).

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

[41] H. Morowitz and E. Smith, Energy flow and the organization of life, Complexity **13**, 51 (2007).

[42] M. Lingam and A. Loeb, *Life in the Cosmos: From Biosignatures to Technosignatures* (Harvard University Press, Cambridge, 2021).

[43] J. M. Smith and E. Szathmary, *The Origins of Life: From the Birth of Life to the Origin of Language* (Oxford University Press, Oxford, 2000).

[44] R. Solé, C. P. Kempes, B. Corominas-Murtra, M. De Domenico, A. Kolchinsky, M. Lachmann, E. Libby, S. Saavedra, E. Smith, and D. Wolpert, Fundamental constraints to the logic of living systems, Interface Focus **14**, 20240010 (2024).

[45] A. Celani and M. Vergassola, Bacterial strategies for chemotaxis response, Proc. Natl. Acad. Sci. USA **107**, 1391 (2010).

[46] V. H. Sridhar, C. Torney, and I. D. Couzin, The geometry of decision-making in individuals and collectives, Proc. Natl. Acad. Sci. USA **118**, e2102157118 (2021).

[47] K. Adamala and J. W. Szostak, Nonenzymatic template-directed RNA synthesis inside model protocells, Science **342**, 1098 (2013).

[48] H. Fellermann, S. Rasmussen, H.-J. Ziock, and R. V. Solé, Life cycle of a minimal protocell—a dissipative particle dynamics study, Artif. Life **13**, 319 (2007).

[49] Proceedings of the Summer Workshop on Information-Driven States of Matter, Rochester, 2024 (unpublished).

[50] C. Rovelli, in *Wandering Towards a Goal: How Can Mindless Mathematical Laws Give Rise to Aims and Intentions?*, edited by A. Aguirre, B. Foster, and Z. Merali, The Frontiers Collection (Springer, Cham, 2018), pp. 17–27.

[51] M. D. Egbert and J. Pérez-Mercader, Methods for measuring viability and evaluating viability indicators, Artif. Life **24**, 106 (2018).

[52] S. Panzeri, C. D. Harvey, E. Piasini, P. E. Latham, and T. Fellin, Cracking the neural code for sensory perception by combining statistics, intervention, and behavior, Neuron **93**, 491 (2017).

[53] W. Bialek, I. Nemenman, and N. Tishby, Predictability, complexity and learning, Neural Comput. **13**, 2409 (2001).

[54] P. W. Tang, P. S. Chua, S. K. Chong, M. S. Mohamad, Y. W. Choon, S. Deris, S. Omatu, J. M. Corchado, W. H. Chan, and R. A. Rahim, A review of gene knockout strategies for microbial cells, Recent Pat. Biotechnol. **9**, 176 (2016).

[55] B. Hall, A. Limaye, and A. B. Kulkarni, Overview: Generation of gene knockout mice, Curr. Protoc. Cell Biol. **44**, 19 (2009).

[56] T. D. Schneider, Evolution of biological information, Nucl. Acids Res. **28**, 2794 (2000).

[57] C. Adami, Information theory in molecular biology, Phys. Life Rev. **1**, 3 (2004).

[58] D. R. Sowinski, J. Carroll-Nellenback, R. N. Markwick, J. Piñero, M. Gleiser, A. Kolchinsky, G. Ghoshal, and A. Frank, Semantic information in a model of resource gathering agents, PRX Life **1**, 023003 (2023).

[59] D. R. Sowinski, G. Ghoshal, and A. Frank, Exo-Daisy World: Revisiting Gaia theory through an informational architecture perspective, Planet. Sci. J. **6**, 176 (2025).

[60] D. R. Sowinski, A. Frank, and G. Ghoshal, Information-theoretic description of a feedback-control Kuramoto model, Phys. Rev. Res. **6**, 043188 (2024).

[61] M. Magarini and P. Stano, Synthetic cells engaged in molecular communication: An opportunity for modelling Shannon- and semantic-information in the chemical domain, Front. Commun. Netw. **2**, 724597 (2021).

[62] T. Nakano, A. W. Eckford, and T. Haraguchi, *Molecular Communication*, 2nd ed. (Cambridge University Press, Cambridge, 2024).

[63] N. Farsad, H. B. Yilmaz, A. Eckford, C.-B. Chae, and W. Guo, A comprehensive survey of recent advancements in molecular communication, IEEE Commun. Surv. Tutor. **18**, 1887 (2016).

[64] Y. Hong, N. M. K. Blackman, N. D. Kopp, A. Sen, and D. Velegol, Chemotaxis of nonbiological colloidal rods, Phys. Rev. Lett. **99**, 178103 (2007).

[65] B. Liebchen and H. Löwen, Synthetic chemotaxis and collective behavior in active matter, Acc. Chem. Res. **51**, 2982 (2018).

[66] T. Bäuerle, A. Fischer, T. Speck, and C. Bechinger, Self-organization of active particles by quorum sensing rules, Nat. Commun. **9**, 3232 (2018).

[67] A. Ziepke, I. Maryshev, I. S. Aranson, and E. Frey, Multi-scale organization in communicating active matter, Nat. Commun. **13**, 6727 (2022).

[68] J. Yoshimura and V. A. A. Jansen, Evolution and population dynamics in stochastic environments, Res. Popul. Ecol. **38**, 165 (1996).

[69] J. L. Kelly, Jr., A new interpretation of information rate, Bell Syst. Tech. J. **35**, 917 (1956).

[70] M. C. Donaldson-Matasci, C. T. Bergstrom, and M. Lachmann, The fitness value of information, Oikos **119**, 219 (2010).

[71] O. Rivoire and S. Leibler, The value of information for populations in varying environments, J. Stat. Phys. **142**, 1124 (2011).

[72] J. Piñero, D. R. Sowinski, G. Ghoshal, A. Frank, and A. Kolchinsky, Information bounds production in replicator systems, arXiv:2501.00396.

[73] A. S. Moffett, N. Wallbridge, C. Plummer, and A. W. Eckford, Fitness value of information with delayed phenotype switching: Optimal performance with imperfect sensing, Phys. Rev. E **102**, 052403 (2020).

[74] S. E. Marzen and J. P. Crutchfield, Optimized bacteria are environmental prediction engines, Phys. Rev. E **98**, 012408 (2018).

[75] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression* (Prentice-Hall, Englewood Cliffs, 1971).

[76] T. Barker, P. J. Thomas, A. S. Moffett, A. W. Eckford, and M. Pierobon, *Proceedings of the 11th ACM International Conference on Nanoscale Computing and Communication* (ACM Press, New York, 2024), pp. 54–59.

[77] T. S. Barker, M. Pierobon, and P. J. Thomas, Subjective information and survival in a simulated biological system, Entropy **24**, 639 (2022).

[78] T. Berger, Living information theory: The 2002 Shannon lecture, IEEE Inf. Theory Soc. Newslett. **53**, 6 (2002).

[79] A. S. Moffett and A. W. Eckford, Minimal informational requirements for fitness, Phys. Rev. E **105**, 014403 (2022).

[80] A. S. Moffett and A. W. Eckford, Kelly bets and single-letter codes: Optimal information processing in natural systems, IEEE Trans. Mol. Biol. Multi-Scale Commun. 1 (2025).

[81] P. A. Hoskisson and G. Hobbs, Continuous culture–making a comeback? Microbiology **151**, 3153 (2005).

[82] N. R. Wright, N. P. Rønnest, and N. Sonnenschein, Single-cell technologies to understand the mechanisms of cellular adaptation in chemostats, Front. Bioeng. Biotechnol. **8**, 579841 (2020).

[83] K. Liu, A. Blokhuis, C. van Ewijk, A. Kiani, J. Wu, W. H. Roos, and S. Otto, Light-driven eco-evolutionary dynamics in a synthetic replicator system, Nat. Chem. **16**, 79 (2024).

[84] M. Lingam, R. Nichols, and A. Balbi, A Bayesian analysis of the probability of the origin of life per site conducive to abiogenesis, Astrobiology **24**, 813 (2024).

[85] S. Bartlett and M. L. Wong, Defining lyfe in the universe: From three privileged functions to four pillars, Life **10**, 42 (2020).

[86] C. Jeancolas, C. Malaterre, and P. Nghe, Thresholds in origin of life scenarios, iScience **23**, 101756 (2020).

[87] S. J. Bartlett and P. Beckett, Probing complexity: Thermodynamics and computational mechanics approaches to origins studies, Interface Focus **9**, 20190058 (2019).

[88] J. P. Crutchfield, Between order and chaos, Nat. Phys. **8**, 17 (2012).

[89] C. L. Buckley, T. Lewens, M. Levin, B. Millidge, A. Tschantz, and R. A. Watson, Natural induction: Spontaneous adaptive organisation without natural selection, Entropy **26**, 765 (2024).

[90] W. Zhong, J. M. Gold, S. Marzen, J. L. England, and N. Yunger Halpern, Machine learning outperforms thermodynamics in measuring how well a many-body system learns a drive, Sci. Rep. **11**, 9333 (2021).

[91] M. Dueñas-Díez and J. Pérez-Mercader, How chemistry computes: Language recognition by non-biochemical chemical automata. From finite automata to turing machines, iScience **19**, 514 (2019).

[92] S. Bartlett and D. Louapre, Provenance of life: Chemical autonomous agents surviving through associative learning, Phys. Rev. E **106**, 034401 (2022).

[93] E. Branscomb, T. Biancalani, N. Goldenfeld, and M. Russell, Escapement mechanisms and the conversion of disequilibria; The engines of creation, Phys. Rep. **677**, 1 (2017).

[94] W. Nitschke, B. Schoepp-Cothenet, S. Duval, K. Zuchan, O. Farr, F. Baymann, F. Panico, A. Minguzzi, E. Branscomb, and M. J. Russell, Aqueous electrochemistry: The toolbox for life's emergence from redox disequilibria, Electrochem. Sci. Adv. **3**, e2100192 (2023).

[95] W. R. Ashby, Requisite variety and its implications for the control of complex systems, Cybernetica **1**, 83 (1958).

[96] J. Piñero, R. Solé, and A. Kolchinsky, Optimization of nonequilibrium free energy harvesting illustrated on bacteriorhodopsin, Phys. Rev. Res. **6**, 013275 (2024).

[97] B. Drobot, J. M. Iglesias-Artola, K. Le Vay, V. Mayr, M. Kar, M. Kreysing, H. Mutschler, and T. Y. D. Tang, Compartmentalised RNA catalysis in membrane-free coacervate protocells, Nat. Commun. **9**, 3643 (2018).

[98] T. Z. Jia, C. Hentrich, and J. W. Szostak, Rapid RNA exchange in aqueous two-phase system and coacervate droplets, Origins Life Evol. B. **44**, 1 (2014).

[99] J. M. Carothers, S. C. Oestreich, J. H. Davis, and J. W. Szostak, Informational complexity and functional activity of RNA structures, J. Am. Chem. Soc. **126**, 5130 (2004).

[100] J. W. Szostak, Functional information: Molecular messages, Nature (London) **423**, 689 (2003).

[101] A. S. Petrov, B. Gulen, A. M. Norris, N. A. Kovacs, C. R. Bernier, K. A. Lanier, G. E. Fox, S. C. Harvey, R. M. Wartell, N. V. Hud, and L. D. Williams, History of the ribosome and the origin of translation, Proc. Natl. Acad. Sci. USA **112**, 15396 (2015).

[102] R. Krishnamurthy, Life's biological chemistry: A destiny or destination starting from prebiotic chemistry? Chem. A Eur. J. **24**, 16708 (2018).

[103] D. Schulze-Makuch and L. N. Irwin, *Life in the Universe: Expectations and Constraints*, 3rd ed. (Springer, Cham, 2018).

[104] C. S. Cockell, *Astrobiology: Understanding Life in the Universe*, 2nd ed. (Wiley, Hoboken, 2020).

[105] M. Lingam and A. Balbi, *From Stars to Life: A Quantitative Approach to Astrobiology* (Cambridge University Press, Cambridge, 2024).

[106] https://astrobiology.nasa.gov/research/astrobiology-at-nasa/exobiology/.

[107] H. Lammer, J. H. Bredehöft, A. Coustenis, M. L. Khodachenko, L. Kaltenegger, O. Grasset, D. Prieur, F. Raulin, P. Ehrenfreund, M. Yamauchi, J. E. Wahlund, J. M. Grießmeier, G. Stangl, C. S. Cockell, Y. N. Kulikov, J. L. Grenfell, and H. Rauer, What makes a planet habitable? Astron. Astrophys. Rev. **17**, 181 (2009).

[108] C. S. Cockell, T. Bush, C. Bryce, S. Direito, M. Fox-Powell, J. P. Harrison, H. Lammer, H. Landenmark, J. Martin-Torres, N. Nicholson, L. Noack, J. O'Malley-James, S. J. Payler, A. Rushby, T. Samuels, P. Schwendner, J. Wadsworth, and M. P. Zorzano, Habitability: A review, Astrobiology **16**, 89 (2016).

[109] S. R. Kane, *Planetary Habitability* (IOP, Bristol, 2021).

[110] M. Neveu, L. E. Hays, M. A. Voytek, M. H. New, and M. D. Schulte, The ladder of life detection, Astrobiology **18**, 1375 (2018).

[111] *Biosignatures for Astrobiology*, edited by B. Cavalazzi and F. Westall, Advances in Astrobiology and Biogeophysics (Springer, Cham, 2019).

[112] E. W. Schwieterman and M. Leung, An overview of exoplanet biosignatures, Rev. Mineral. Geochem. **90**, 465 (2024).

[113] G. Alexandre, S. Greer-Phillips, and I. B. Zhulin, Ecological role of energy taxis in microorganisms, FEMS Microbiol. Rev. **28**, 113 (2004).

[114] B. Hu and Y. Tu, Behaviors and strategies of bacterial navigation in chemical and nonchemical gradients, PLoS Comput. Biol. **10**, e1003672 (2014).

[115] S. Bi and V. Sourjik, Stimulus sensing and signal processing in bacterial chemotaxis, Curr. Opin. Microbiol. **45**, 22 (2018).

[116] R. C. Raymond, Communications, entropy, and life, Am. Sci. **38**, 273 (1950).

[117] W. R. Loewenstein, *The Touchstone of Life: Molecular Information, Cell Communication, and the Foundations of Life* (Oxford University Press, Oxford, 1999).

[118] J. P. Armitage, Bacterial tactic responses, Adv. Microb. Physiol. **41**, 229 (1999).

[119] G. H. Wadhams and J. P. Armitage, Making sense of it all: Bacterial chemotaxis, Nat. Rev. Mol. Cell Biol. **5**, 1024 (2004).

[120] J. Wong-Ng, A. Celani, and M. Vergassola, Exploring the function of bacterial chemotaxis, Curr. Opin. Microbiol. **45**, 16 (2018).

[121] S. L. Porter, G. H. Wadhams, and J. P. Armitage, Signal processing in complex chemotaxis pathways, Nat. Rev. Microbiol. **9**, 153 (2011).

[122] D. L. Kirchman, *Processes in Microbial Ecology*, 2nd ed. (Oxford University Press, Oxford, 2018).

[123] J.-B. Raina, V. Fernandez, B. Lambert, R. Stocker, and J. R. Seymour, The role of microbial motility and chemotaxis in symbiosis, Nat. Rev. Microbiol. **17**, 284 (2019).

[124] R. A. Dixon and C. J. Lamb, Molecular communication in interactions between plants and microbial pathogens, Annu. Rev. Plant. Physiol. Plant. Mol. Biol. **41**, 339 (1990).

[125] B. L. Bassler, Small talk: Cell-to-cell communication in bacteria, Cell **109**, 421 (2002).

[126] C. M. Waters and B. L. Bassler, Quorum sensing: Cell-to-cell communication in bacteria, Annu. Rev. Cell Dev. Biol. **21**, 319 (2005).

[127] E. Ben-Jacob, Learning from bacteria about natural information processing, Ann. N.Y. Acad. Sci. **1178**, 78 (2009).

[128] G. Witzany, What is life? Front. Astron. Space Sci. **7**, 7 (2020).

[129] E. E. Stüeken, R. E. Anderson, J. S. Bowman, W. J. Brazelton, J. Colangelo-Lillis, A. D. Goldman, S. M. Som, and J. A. Baross, Did life originate from a global chemical reactor? Geobiology **11**, 101 (2013).

[130] D. D. Sasselov, J. P. Grotzinger, and J. D. Sutherland, The origin of life as a planetary phenomenon, Sci. Adv. **6**, eaax3419 (2020).

[131] F. Westall, A. Brack, A. G. Fairén, and M. D. Schulte, Setting the geological scene for the origin of life and continuing open questions about its emergence, Front. Astron. Space Sci. **9**, 1095701 (2023).

[132] L. E. Rodriguez, T. Altair, N. Y. Hermis, T. Z. Jia, T. P. Roche, L. H. Steller, and J. M. Weber, A geological and chemical context for the origins of life on early earth, Astrobiology **24**, S76 (2024).

[133] C. Scharf and O. Witkowski, Rebuilding the habitable zone from the bottom up with computational zones, Astrobiology **24**, 613 (2024).
[134] M. Lingam, Theoretical constraints imposed by gradient detection and dispersal on microbial size in astrobiological environments, Astrobiology **21**, 813 (2021).
[135] E. A. Martens, N. Wadhwa, N. S. Jacobsen, C. Lindemann, K. H. Andersen, and A. Visser, Size structures sensory hierarchy in ocean life, Proc. R. Soc. B **282**, 20151346 (2015).
[136] M. Lingam, Information transmission via molecular communication in astrobiological environments, Astrobiology **24**, 84 (2024).
[137] L. Grebenstein, J. Kirchner, R. S. Peixoto, W. Zimmermann, F. Irnstorfer, W. Wicke, A. Ahmadzadeh, V. Jamali, G. Fischer, R. Weigel, A. Burkovski, and R. Schober, Biological optical-to-chemical signal conversion interface: A small-scale modulator for molecular communications, IEEE Trans. Nanobiosci. **18**, 31 (2018).
[138] S. Sarkar, M. Z. Ali, and S. Choubey, Efficacy of information transmission in cellular communication, Phys. Rev. Res. **5**, 013092 (2023).
[139] M. Lingam, A. Frank, and A. Balbi, Planetary scale information transmission in the biosphere and technosphere: Limits and evolution, Life **13**, 1850 (2023).
[140] P. G. Conrad and K. H. Nealson, A non-earthcentric approach to life detection, Astrobiology **1**, 15 (2001).
[141] M. A. Chan, N. W. Hinman, S. L. Potter-McIntyre, K. E. Schubert, R. J. Gillams, S. M. Awramik, P. J. Boston, D. M. Bower, D. J. Des Marais, J. D. Farmer *et al.*, Deciphering biosignatures in planetary contexts, Astrobiology **19**, 1075 (2019).
[142] L. M. Barge, L. E. Rodriguez, J. M. Weber, and B. P. Theiling, Determining the "biosignature threshold" for life detection on biotic, abiotic, or prebiotic worlds, Astrobiology **22**, 481 (2022).
[143] C. Malaterre, I. L. ten Kate, M. Baqué, V. Debaille, J. L. Grenfell, E. J. Javaux, N. Khawaja, F. Klenner, Y. J. Lara, S. McMahon, K. Moore, L. Noack, C. H. L. Patty, and F. Postberg, Is there such a thing as a biosignature? Astrobiology **23**, 1213 (2023).
[144] J. M. García Ruiz, A. Carnerup, A. G. Christy, N. J. Welham, and S. T. Hyde, Morphology: An ambiguous indicator of biogenicity, Astrobiology **2**, 353 (2002).
[145] S. McMahon and J. Cosmidis, False biosignatures on Mars: Anticipating ambiguity, J. Geol. Soc. **179**, jgs2021-050 (2022).
[146] S. S. Johnson, E. V. Anslyn, H. V. Graham, P. R. Mahaffy, and A. D. Ellington, Fingerprinting non-terran biosignatures, Astrobiology **18**, 915 (2018).
[147] E. D. Dorn, K. H. Nealson, and C. Adami, Monomer abundance distribution patterns as a universal biosignature: Examples from terrestrial and digital life, J. Mol. Evol. **72**, 283 (2011).
[148] S. M. Marshall, C. Mathis, E. Carrick, G. Keenan, G. J. T. Cooper, H. Graham, M. Craven, P. S. Gromski, D. G. Moore, S. I. Walker, and L. Cronin, Identifying molecules as biosignatures with assembly theory and mass spectrometry, Nat. Commun. **12**, 3033 (2021).
[149] H. J. Cleaves, G. Hystad, A. Prabhu, M. L. Wong, G. D. Cody, S. Economon, and R. M. Hazen, A robust, agnostic molecular biosignature based on machine learning, Proc. Natl. Acad. Sci. USA **120**, e2307149120 (2023).
[150] S. Bartlett, J. Li, L. Gu, L. Sinapayen, S. Fan, V. Natraj, J. H. Jiang, D. Crisp, and Y. L. Yung, Assessing planetary complexity and potential agnostic biosignatures using epsilon machines, Nat. Astron. **6**, 387 (2022).
[151] S. Vannah, M. Gleiser, and L. Kaltenegger, An information theory approach to identifying signs of life on transiting planets, Mon. Not. R. Astron. Soc. Lett. **528**, L4 (2023).
[152] T. Schreiber, Measuring information transfer, Phys. Rev. Lett. **85**, 461 (2000).
[153] K. Hlaváčková-Schindler, M. Paluš, M. Vejmelka, and J. Bhattacharya, Causality detection based on information-theoretic approaches in time series analysis, Phys. Rep. **441**, 1 (2007).
[154] A. E. Allahverdyan, D. Janzing, and G. Mahler, Thermodynamic efficiency of information and heat flow, J. Stat. Mech. (2009) P09011.
[155] A. C. Barato, D. Hartich, and U. Seifert, Efficiency of cellular information processing, New J. Phys. **16**, 103024 (2014).
[156] M. Li and P. Vitányi, *An Introduction to Kolmogorov Complexity and its Applications*, 3rd ed., Texts in Computer Science (Springer, New York, 2008).
[157] P. M. Vitányi, Meaningful information, IEEE Trans. Inf. Theory **52**, 4617 (2006).
[158] S. Ro, B. Guo, A. Shih, T. V. Phan, R. H. Austin, D. Levine, P. M. Chaikin, and S. Martiniani, Model-free measurement of local entropy production and extractable work in active matter, Phys. Rev. Lett. **129**, 220601 (2022).
[159] I. G. Johnston, K. Dingle, S. F. Greenbury, C. Q. Camargo, J. P. K. Doye, S. E. Ahnert, and A. A. Louis, Symmetry and simplicity spontaneously emerge from the algorithmic nature of evolution, Proc. Natl. Acad. Sci. USA **119**, e2113883119 (2022).