

An energy-efficient process of non-deterministic computation drives the emergence of predictive models and exploratory behavior

Anonymous for peer review*
Anonymous for peer review
(Dated: February 14, 2023)

Cortical neural networks encode information about the environment, combining data across sensory modalities to form predictive models of the world, which in turn drive behavioral output. Cortical population coding is probabilistic, with synchronous firing across the neural network achieved in the context of noisy inputs. The system-wide computational process, which encodes the likely state of the local environment, is achieved at a cost of only 20 Watts, indicating a deep connection between neuronal information processing and energy-efficient computation. This report presents a new framework for modeling non-deterministic computation in cortical neural networks, in terms of thermodynamic laws. Initially, free energy is expended to produce von Neumann entropy, then predictive value is extracted from that thermodynamic quantity of information. The extraction of predictive value during a single computation yields a percept, or a predictive semantical statement about the local environment, and the integration of sequential neural network states yields a temporal sequence of percepts, or a predictive syntactical statement about the cause-effect relationship between perceived events. The amount of predictive value available for computation is limited by the total amount of energy entering the system, and will always be incomplete, due to thermodynamic constraints. This process of thermodynamic computation naturally produces a rival energetic cost function, which minimizes energy expenditure: the system can either explore its local environment to gain potential predictive value, or it can exploit previously-acquired predictive value by triggering a contextually-relevant and thermodynamically-favored sequence of neural network states. The system grows into a more ordered state over time, as it physically encodes the predictive value acquired by interacting with its environment.

I. INTRODUCTION

Animals collect data about the world through multiple sensory modalities, transducing physical events in the local environment into electrical signals, then passing these data to dedicated processing centers within the central nervous system [1, 2]. In the cerebral cortex, information is integrated across multiple sensory modalities and compared with contextually-relevant predictions or expectations [3, 4]. This computational process informs the selection of behavior [5, 6]. However, the mechanisms underlying information integration and predictive modeling at the systems level are currently not well understood.

Cortical neural networks encode both the state of the local environment and the divergence from expectations set by a predictive model [7, 8]. Critically, any information that is found to have useful predictive value is encoded into memory storage through synaptic remodeling, which favors those same patterns of neural activity to re-occur in a similar context [9, 10]. Thus, neurons coordinate memory storage and processing functions within a single computational unit. Yet this synaptic remodeling is primarily a feature of cortical neurons, not spinal neurons. Although some spike timing-dependent plasticity occurs upon repetitive paired stimulation [11-13] and after injury [14-16] in spinal circuits, cortical neurons engage in spontaneous synaptic remodeling under standard physiological conditions [17-19]. The emergence of

a more ordered system state in cortical neural networks during learning and development correlates with the construction of predictive cognitive models [20, 21].

Uncovering the natural processes by which cortical neural networks form predictive models over time may drive advances in both the cognitive sciences and the field of computational intelligence. As information is processed in cortical neural networks, predictive value is extracted and stored, and output behavior is selected in the context of both incoming sensory data and relevant past experience. This massively parallel computing process, which accesses predictive models stored in local memory, is highly energy-efficient. Yet it is not currently well understood how the mammalian cerebral cortex achieves this exascale computing power and a generalized problem-solving ability at an energetic cost of only 20 Watts [22, 23]. This energetic efficiency is particularly surprising, given that cortical neurons retain sensitivity to random electrical noise in gating signaling outcomes [24-27], a property that should be expected to reduce energy efficiency.

But importantly, prediction minimizes energetic expenditure. In cortical neural networks, prediction errors provide an energy efficient method of encoding how much a percept diverges from the contextual expectation [28, 29]. The ‘minimization of surprise’ was proposed by Karl Friston to be the guiding force driving the improvement of predictive models, with new information continually prompting the revision of erroneous priors [30, 31]. This view – embodied by ‘the free energy principle’ – asserts that biological systems gradually achieve a more ordered configuration over time by identifying a more compatible

* email; website

state with their environment, simply by reducing predictive errors [32]. It should be noted however that ‘free energy’ is a statistical quantity, not a thermodynamic quantity, in this particular neuroscientific context. Likewise, the concept of ‘free energy’ has been employed in the machine learning field to solve optimization problems through a process of gradient descent [33]. For the past forty years, researchers have striven to explain computation, in both biological and simulated neural networks, in terms of ‘selecting an optimal system state from a large probability distribution’ [34-36]. Yet a direct connection between noisy coding and energy efficiency has remained elusive. A deeper explanation of cortical computation, in terms of thermodynamics, is needed.

In this report, a new theoretical model is presented, tying together the concepts of computational free energy and thermodynamic free energy. Here, the macrostate of the cortical neural network is modeled as the mixed sum of all component microstates – the physical quantity of information, or von Neumann entropy, held by the system. The extraction of predictive value, or consistency, then compresses this thermodynamic quantity of information, as an optimal system state is selected from a large probability distribution. This computational process not only maximizes free energy availability; it also yields a multi-sensory percept, or a predictive semantical statement about the present state of the local environment, which is encoded by the cortical neural network state. These multi-sensory percepts are then integrated over time to form predictive syntactical statements about the relationship between perceived events, encoded by a sequence of cortical neural network states. This energy-efficient computational process naturally leads to the construction of predictive cognitive models.

Since energy is always conserved, the compression of information entropy must be paired with the release of thermal free energy, local to any reduction in uncertainty, in accordance with the Landauer principle [37-40]. This free energy can then be used to do work within the system, physically encoding the predictive value that was extracted from some thermodynamic quantity of information. This computational process allows a non-dissipative thermodynamic system to grow into a more ordered state over time, as it encodes predictive value. The amount of predictive value available during a computation is limited by the total amount of energy entering the system over that period of time, and will always be incomplete, due to thermodynamic constraints. The system will therefore respect the second law of thermodynamics and maintain some amount of uncertainty during predictive processing.

This new approach offers a theoretical framework for achieving energy-efficient non-deterministic computation in cortical neural networks, in terms of thermodynamic laws. It is also consistent with the extraordinary energy efficiency of the human brain. In addition to this broad explanatory power, this new approach offers testable hypotheses relating probabilistic computational processes to network-level dynamics and behavior.

II. METHODS

A. Entropy is a description of the system macrostate

For a classical thermodynamic system, the macrostate of the system is a distribution of microstates, given by the Gibbs entropy formula. Here, k_B is the Boltzmann constant, E_i is the energy of microstate i , p_i is the probability that microstate occurs, and the Gibbs entropy of the system is given by S :

$$S = -k_B \sum p_i \ln p_i. \quad (1)$$

Yet the behavior of cortical neural networks can be better described as a statistical ensemble of microstates [35, 36]. And so, in this model, the macrostate of the system is formally described as a statistical ensemble of all component pure microstates, given by the von Neumann entropy formula [41]:

$$S(\rho) = -\text{Tr} (\rho \ln \rho). \quad (2)$$

Here, entropy is a high-dimensional volume of possible system states, represented by the trace across a density matrix ρ . ρ is the sum of all mutually orthogonal pure states ρ_x , each occurring with some probability p_x . As the encoding Thermodynamic System ‘A’ is perturbed, by interacting with its surrounding environment, Thermodynamic System ‘B’, the density matrix undergoes a time evolution, from $\rho_{AB} \rightarrow \rho'_{AB}$. During this time evolution, consistencies can be identified between the two particle systems [42-45]. This non-deterministic process leads to the selection of a mutually-compatible system state. Here, entropy is additive in uncorrelated systems, but it is subtractive in correlated systems, as mutual redundancies in system states are recognized and reduced. This leads to the subadditivity rule:

$$\begin{aligned} S(\rho_{AB}) &= S(\rho_A) + S(\rho_B) \\ &= S(\rho'_A) + S(\rho'_B) \geq S(\rho'_{AB}). \end{aligned} \quad (3)$$

A cortical neural network is a far-from equilibrium thermodynamic system that actively acquires energy to accomplish work. That work involves *encoding the state of the surrounding environment into the physical state of the system itself*. This theoretical approach models the encoding process as a thermodynamic computation. The macrostate that results from this thermodynamic computation is not a deterministic outcome; rather, an optimal system state is selected in the present context from a large probability distribution. The most thermodynamically favored and ‘optimal’ system state is the one that is both most correlated with the surrounding environment and most compatible with existing anatomical and physiological constraints. The physical compression

of entropy, during a computation, causes the system to attain a more ordered state, as it becomes less random and more compatible with its surrounding environment.

B. Thermodynamic constraints on irreversible work

Thermodynamic System 'A' – a cortical neural network – identifies an optimal state in the context of its surrounding environment, Thermodynamic System 'B', by selecting the most consistent arrangement of microstates to encode the state of its surrounding environment. The entropy of the combined system is minimized over some time evolution, as consistencies are reduced. However, there are thermodynamic limits to this computational process. The Jarzynski equality, relating the quantity of free energy and work in non-equilibrium systems, holds [46, 47]. There simply cannot be more work dissipated toward entropy generation than the amount of energy expenditure. This relationship provides an equivalency between the average work dissipated to create entropy, an ensemble of possible measurements, given by \bar{W} , and the change in Helmholtz free energy, given by ΔG , after some time has passed, with $\beta = 1/k_B T$:

$$\exp(-\beta\bar{W}) = \exp(-\beta\Delta G). \quad (4)$$

Another strict thermodynamic equality, the Crooks fluctuation theorem, holds here as well - as in any path-dependent system state [48]. This rule relates the probability of any particular time-dependent trajectory ($a \rightarrow b$) to its time-reversal trajectory ($b \rightarrow a$). The Crooks fluctuation theorem requires all irreversible work ($\Delta W_{a \rightarrow b}$) to be greater than or equal to the difference in free energy between events and time-reversed events ($\Delta G = \Delta G_{a \rightarrow b} - \Delta G_{b \rightarrow a}$), and demonstrates the probability of forward events in a non-equilibrium system to be exponentially more likely than time-reversed events:

$$\frac{P_{a \rightarrow b}}{P_{b \rightarrow a}} = \exp \beta (\Delta W_{a \rightarrow b} - \Delta G). \quad (5)$$

Any physical structure engaging in physical work will dissipate some amount of free energy toward entropy. However, this energy is not irreversibly lost [49]. Indeed, the energy expended to create entropy cannot disappear, since energy can never be created nor destroyed. And since energy is always conserved, any compression of information entropy must be paired with the release of free energy. This conservation law, embodied by the Landauer principle, has been empirically validated in recent years [37-40]. Any compression of information entropy, or reduction of uncertainty, must be paired with the release of free energy. This free energy can then be used to do thermodynamic work within the system, by encoding the most likely state of the surrounding environment into the state of the neural network.

C. Thermodynamic foundations of predictive inference

Entropy is a natural by-product of thermodynamic systems: a measure of inefficiency in doing work. Entropy is also a description of the physical system state: a vector map of all possible states for all atoms in a thermodynamic system after some time evolution. These two definitions are tied together in a system whose 'work' is computing the most optimal system state to encode its environment. It is worth noting again that information compression is a known physical process that generates a discrete quantity of free energy [37-40]. This conversion of physical quantities is given by the Landauer limit, which provides the amount of energy released by the erasure of a single digital bit:

$$\langle W \rangle = k_B T (\log 2). \quad (6)$$

This concept can be extrapolated to computational systems with any number of possible microstates that trap *energy* to do *computational work*. The recently-derived Still dissipation theorem demonstrates that the quantity of non-predictive information in a system is proportional to the energy lost when an external driving signal changes by an incremental amount within a non-equilibrium system [50]. Here, the amount of information *without* predictive value is equal to the amount of work dissipated to entropy as the signal changes over a time course from $x_{t=0}$ (the immediately previous system state) to x_t (the present state):

$$\begin{aligned} \langle W_{diss} [x_{t=0} \rightarrow x_t] \rangle &= k_B T [I_{mem}(t) - I_{pred}(t)] \\ &= k_B T [I_{non-pred}(t)] \\ &= k_B T [\Delta I]. \end{aligned} \quad (7)$$

This equation stipulates that the total amount of information stored in memory, subtracted by the amount of information with predictive value, is the amount of non-predictive information remaining after predictive value has been extracted. This quantity is dissipated to entropy upon the completion of the system-wide computation, and is therefore unavailable as free energy to do work [51]. I_{mem} is equivalent to the total amount of information generated, or the amount of uncertainty gained by the system, over the course of a thermodynamic computation. I_{pred} is the quantity of predictive value extracted during the thermodynamic computation. The net quantity remaining, $I_{non-pred}$, is the amount of work that has been dissipated to entropy upon completion of a thermodynamic computing cycle. That net quantity of non-predictive information is equal to the quantity of uncorrelated states between Systems A and B, which will always be greater than or equal to zero:

$$\begin{aligned} \langle W_{diss} [x_{t=0} \rightarrow x_t] \rangle &= k_B T [\Delta I] \geq 0 \\ &= k_B T [S(\rho'_{AB})]. \end{aligned} \quad (8)$$

As uncertainty *increases*, entropy is produced, and less free energy is available within the system to do work. As uncertainty *decreases*, entropy is compressed, and more free energy is available within the system to complete tasks, such as remodeling synapses. A system that encodes its environment can therefore become more ordered by interacting with its local environment and identifying correlations. It is even *thermodynamically favored* for a system to take on a more ordered state over time, as it encodes predictive value gained by interacting with its surrounding environment [52-54]; indeed, the pruning of neural pathways during learning and development would not spontaneously occur if this were not the case. Here, the spontaneous emergence of a more ordered system state naturally emerges over time, from a process of thermodynamic computation – as some quantity of information entropy is generated and compressed, through the extraction of predictive value. However, the first and second laws of thermodynamics are always respected - energy is never created nor destroyed, and there is always some net dissipation of energy to entropy over time.

III. RESULTS

A. The first law limits the amount of predictive value in a quantity of information

In accordance with the second law of thermodynamics, entropy must always increase over time. However, far-from-equilibrium particle systems - which trap heat to do work - can either increase or decrease disorder locally. To prevent a violation of thermodynamic laws, free energy must be dissipated toward work whenever the system state becomes more ordered. Any reduction of uncertainty ($k_B T[\Delta I] = T\Delta S$) must be balanced by an increase in free energy (ΔG), to balance the total energetic account. Any inefficiency in operations increases entropy, while a more ordered system state increases the amount of energetic currency available to do work.

The lower the value of $T\Delta S$, the greater the amount of free energy available to the system. The quantity of free energy can be maximized in two ways: 1) if the system does not generate much uncertainty or entropy (with the value of $I_{mem}(t)$ approaching zero), or 2) if the entropy generated by the system over time t has a large amount of predictive value that can be extracted (with the value of $I_{mem}(t) - I_{pred}(t)$ approaching zero). The reduction of possible system macrostates, or the compression of information entropy, is equivalent to the extraction of predictive value. This computational process is thermodynamically favored, because it provides the system with more available free energy.

If the quantity of information contains internal consistencies, that quantity is naturally reduced or compressed. Any redundant or compressible entropy (e.g. possible system states that are identical to previous system states) are highly predictable and are easily accepted. Con-

versely, information that is unreliable or anomalous (e.g. neural signals that are more likely to be errors than accurate reporting of stimuli) reduce the amount of predictive value available to the system, providing a measure of inefficiency or wasted energy.

It is useful for a thermodynamic computing system to direct attention toward a novel stimulus, in case that event provides useful predictions about subsequent events. If predictive value is successfully extracted from the information generated, there is little net energetic cost to the initial energetic expenditure. The free energy that is released by reducing uncertainty and maximizing predictive value is recovered and made available to do work within the system. In this model, a cortical neural network continually acquires information about the local environment through the sensory apparatus, encoding physical events into the probabilistic macrostate of an ensemble of computational units. Extracting predictive value from this ‘information’ maximizes the free energy that is available to the system. This free energy can be used to implement structural change, leading to a more ordered system state that encodes the predictive value acquired during the computation.

Yet there is a limit to the amount of predictive value a system can have. Any system that gains a more ordered state has reduced the quantity of disorder or entropy; this may be true for a local, non-closed thermodynamic system, but it cannot be true for a closed system or the universe as a whole, because there must be some external energy source driving this computational work. The total information created by the system is limited by the amount of energy entering the system ΔE . Of the total information generated by the thermodynamic system, some amount will have predictive value, but the predictive value available will be limited by the total quantity of information held by the system. If the quantity of predictive value gained is greater than the amount of caloric energy expended to generate entropy, then the quantities in Eqs. 7 and 8 would be negative, and energy would be created by the thermodynamic system through the act of information processing alone. Since the creation of energy is explicitly forbidden by the first law of thermodynamics, the amount of predictive value is limited by the amount of energy available to the system. As such, any thermodynamic system gaining a more ordered state by extracting predictive value must generate an incomplete predictive model.

If a thermodynamic computing system were ever to reach perfect predictability for all computable statements, with $I_{mem}(t) = I_{pred}(t)$, all memory in the system would be erased, and there would be no information remaining to have predictive value. Therefore, no memory can continue to exist for any period of time with 100% predictive efficiency. If the value of $I_{mem}(t) - I_{pred}(t)$ were to reach zero, all system memory would be erased, and the system would spontaneously become disordered.

This simple model demonstrates that extracting predictive value is both a thermodynamic process and a com-

putational process. The first law of thermodynamics limits the amount of energy expended toward and recovered from entropy; this total quantity of energy turnover is constrained by the net quantity of energy entering the system. Since the physical process of compressing information is dependent on parsing predictive value, as provided in Still's equation, the quantity of disorder lost through computation must be equivalent to the amount of predictive value extracted by the system. Yet some disorder must always remain. If a system is perfectly ordered and perfectly predictive, it will cease to operate. Continued thermodynamic operation requires some uncertainty.

B. The predictive value in any set of computable statements will always be incomplete

In mathematical logic, any set of axioms which makes statements of truth is inherently limited and necessarily incomplete. It is useful to consider this rule on its own terms, before assessing whether it applies to physical systems performing logical operations. Gödel's first incompleteness theorem provides that there will always be statements within a set of axioms which are true but cannot be proven or disproven within the set of axioms itself. Gödel's second incompleteness theorem states that any set of axioms cannot contain any proof for the consistency of the set of axioms itself. As such, any logical system that evaluates both *the internal consistency or 'truth' of computable statements* and *the relationship of these statements to each other* is necessarily incomplete [55, 56].

These two incompleteness theorems assert that a set of axioms cannot be both completely true and completely known. Any number of statements can be held to be *semantically* true; the limit applies to full knowledge of their *syntactical* relationships. Eqs. 7 and 8 show that this law of mathematical logic must be true for any thermodynamic computing system. Over time, any number of semantical statements may be decided by a system. If the system does *not* hold all semantical statements about the universe, then there will also be syntactical relationships that are not known, and its knowledge is incomplete. If the system *does* hold all semantical statements about the universe, then knowing all syntactical relationships between these axiomatic statements would complete the predictive value of the system. Once all knowledge is complete, the thermodynamic system becomes completely ordered. With no uncertainty, there is no remaining entropy. The memory of the system would be completely erased, as all entropy is converted to free energy, and the system would become disordered once again. Continued thermodynamic operation requires uncertainty, and thus, incompleteness.

This result is perhaps not surprising, as the relevance of incompleteness in relation to computing was formally addressed a century ago by the *Entscheidungsproblem*, or

decision-problem, proposed by Hilbert and Ackermann, which posed the question of whether an algorithm or computer could decide the universal validity of any mathematical statement [57]. Church and Turing quickly recognized there could be no computable function which determines whether two computational expressions are equivalent, thereby demonstrating the fundamental unsolvability of the *Entscheidungsproblem* and the practicality of Gödel's incompleteness theorems in terms of computable statements [58, 59]. Here, it is demonstrated that incompleteness is a fundamental limitation during the thermodynamic extraction of predictive value.

If no true-false decision can be made with regard to the relationship between computable statements, then the predictive value I_{pred} of the system is incomplete. As a result, some amount of work must be dissipated to ΔI , as unresolved information which simply cannot be reduced (yet) into a decisional outcome. Given the physical impossibility of computing the universal truth of all syntactical relationships between semantical statements, there will always be some value of $I_{non-pred}$ and therefore a net positive amount of free energy dedicated to ΔI during a thermodynamic computation, in accordance with the second law of thermodynamics. Yet it is possible to minimize this quantity of ΔI , as predictive value is extracted: If a decision can be made, uncertainty is reduced and the system state encoding that predictive statement is selected from a distribution of possible system states.

C. A single thermodynamic computation yields a predictive semantical statement

In this model, a neural network selects an optimal system state, by extracting predictive value to make consistent statements about its surrounding environment. By parsing for predictive value, the encoding system takes on a more compatible state with its surrounding environment during the process of thermodynamic computation.

Although entropy is generated by the system, predictive value can be extracted, reducing the distribution of possible system states. During each computation, noise is filtered out and anomalies are detected; the most appropriate system state is selected, in the context of the surrounding environment. This computation is based on both incoming sensory data and the biophysical properties of each encoding unit.

During the computation, incoming sensory data is noisily encoded into the neural network and integrated into a system macrostate. The system macrostate (a physical quantity of information) is defined as the mixed sum of all component microstates, which evolve over time t . The system state is then resolved as consistencies are reduced and predictive value is extracted. During this computational process, information is compressed, free energy is released, an optimal system state is selected from a large probability distribution, and all other possible system states are discarded.

This computation effectively encodes a semantical statement about the local environment, or a ‘percept’, such as a simultaneous flash of light and a loud noise. Because this semantical statement is essentially a predictive model of the state of the world at the present moment, constructed from multiple sensory inputs, a reasonable decision in the context of any remaining uncertainty is to orient attention toward any novel or anomalous stimuli, in order to gather more information which might hold additional predictive value. Critically, any information gathered that is not consistent with previously-gathered information or concurrent information is likely to be discarded. This efficiency may lead to perceptual errors, particularly in the presence of unexpected events.

Here, a *multi-sensory percept* is a ‘semantical statement about reality’, which corresponds to synchronous firing across the neural network, as the system performs *work* to encode that predictive statement. A series of percepts can then be integrated across the time domain. A sequence of semantical statements about the surrounding environment, corresponding to a sequence of system states, enables the construction of ‘a syntactical statement about reality’ or a *predictive cognitive model of the likely cause-effect relationship between perceived events*.

D. A sequence of thermodynamic computations yields a predictive syntactical statement

While entropy can decrease during the course of a computational cycle, through the extraction of predictive value, some amount of non-predictive value will always remain. In short, not all entropy can be compressed in a single computational cycle, because not all predictive value can be extracted. *The relationship of that percept to subsequent events* remains uncertain. This uncertainty can only be reduced by extracting additional predictive value. The system may integrate a temporal sequence of neural network states, encoding a temporal sequence of semantical statements, and extract a prediction regarding the likely cause-effect relationship between these perceived events, thereby creating a *predictive model of the syntactical relationships between semantical statements*.

During a single thermodynamic computation, the neural network encodes the most likely state of its surrounding environment. These computations can also be integrated over time, corresponding to a sequence of neural network states which encode a likely sequence of events occurring in the environment. Now, the amount of information held by the neural network is given by integrating the distribution of system macrostates over a longer timescale t , with this thermodynamic quantity equivalent to the net amount of free energy expended to produce entropy over multiple computational cycles. Any entropy that is not compressed or discarded during a single computational cycle remains available for further parsing.

Over the course of a single computational cycle, only semantical truth statements can be ascertained (e.g.

there is a bright flash and a loud noise, with the coincident stimuli both emerging from the same approximate location). In that instant, there is no predictive value to assign syntactical relationships between these stimuli and any subsequent outcomes (e.g. these stimuli immediately precede some aversive or rewarding outcome for the observer). If these stimuli have been previously observed, with no adverse outcome, the stimuli may be tolerated (e.g. this combination of percepts is likely to be part of the annual fireworks display). Alternatively, if the stimuli have been previously observed, with an outcome that potentially affects survival, a decision can be taken to reduce the expected risk (e.g. this combination of percepts is likely to indicate gunfire nearby). If the stimuli are novel, they may be attended carefully. In the absence of contextually-relevant knowledge, gained from prior experience, the system must integrate information over multiple computational cycles and extract predictive value. Any new information gained by the system about its environment during a single computation has uncertain predictive value; this information cannot be fully reduced and therefore must be stored in working memory. However, the immediately subsequent moment may yield more information, permitting the system to assign some predictive syntactical relationship between the previously-acquired information and the incoming information (e.g. the neural network can decide whether this perceived stimulus combination is likely to indicate an innocuous or life-threatening event, and direct movement according to this cognitive model). If a stimulus has high predictive value, gained from prior experience, it may be thermodynamically favored for some particular sequence of neural network states to ensue, triggering stereotyped behavior in that context.

Overlapping computational cycles may occur, with different timescales, as neural network states are temporally integrated. Short sequences of events may be fused together tightly, with the amount of predictive value determined in part by the salience and temporal contingency of sensory cues. Longer sequences of events may be tied together by common contextual cues or patterns of sensorimotor input-output behavior; these extended computations may include a multitude of shorter sequences. In each case, the extended computation includes semantical statements which are predicted to be syntactically related in some way. By integrating individual computations, each with some value of $I_{non-pred}$, into a sequence of computations, the system can extract additional predictive value.

A series of events with high predictive value will compress the quantity of entropy, thereby releasing free energy into the system to do work. Indeed, any time the system gains predictive value, it must balance the account by dissipating free energy toward work. Free energy is dissipated toward work, for example, as the system undertakes spontaneous structural remodeling to gain a more ordered system state, thereby encoding the predictive value acquired through lived experience. This

synaptic remodeling process thermodynamically favors the re-occurrence of sequential neural network states and stereotyped behaviors in predictable contexts. This emergence of a more ordered state allows the organism to more easily navigate familiar environments, minimizing inefficiency and maximizing available free energy.

E. Thermodynamic computation yields a trade-off between exploration and exploitation

In this model of thermodynamic information processing, a single computation yields a ‘semantical’ statement, while temporal sequences of neural network states encode ‘syntactical’ statements. Over a lifetime, the neural network steadily acquires a more ordered state but cannot achieve full predictive value, due to the constraints discussed in Sections IIIA and IIIB.

Any information generated will be parsed for predictive value – and the more predictive value extracted from the information, the more free energy is released to do work. As the system performs work to remodel itself into a more organized state, disorder is reduced, less entropy or uncertainty is produced, and more free energy is available to the system - as long as the system continues to navigate familiar contexts.

This computational process is highly energy-efficient, because consistency minimizes uncertainty and maximizes free energy availability. In this model, a thermodynamic computing system will both *expand* its quantity of information, by engaging in exploratory behavior to acquire predictive value, and *reduce* its quantity of information, by exploiting previously-acquired predictive value and favoring previous patterns of neural activity. These two processes perform in opposition, with any decisional outcome relying on the minimization of energy dissipation. The rival energetic cost function leads a neural network to explore its environment, gaining predictive value, and to engage in habitual patterns of behavior, exploiting that predictive value in familiar contexts.

Each computation cycle culminates in a decision, with any output behavior dependent upon this rival energetic cost function. If the system predicts its own knowledge to be *insufficient* to navigate the present situation, the system will further explore its environment to gain information. If predictive value can be extracted, the information is compressed and there is little net energy expenditure to do that work, although there is an up-front expenditure. The system then remodels itself to encode that newly-acquired predictive value. If the system predicts its own knowledge to be *sufficient* to navigate the situation, the system will exploit any relevant previously-gained predictive value to select a contextually-appropriate action. If patterns of neural firing and behavior are readily activated by a familiar stimulus, with minimal uncertainty during the task, little net energy expenditure is needed to compute the decision.

F. Predictions

This theoretical approach to modeling non-deterministic computation in cortical neural networks makes some specific predictions with regard to the energy-efficiency of the brain [60], the spontaneous release of thermal free energy upon information compression [61], and the contribution of these localized thermal fluctuations to cortical neuron signaling outcomes [62]. This approach also makes specific predictions about the expected effects of electromagnetic stimulation and pharmacological interventions on perceptual content [63]. Some further predictions of the theory, prompted by the present model, include:

1. Thermodynamic computation drives synchronous firing of cells at a range of frequencies

In this model, a thermodynamic computation yields a predictive semantical statement regarding the most likely state of the surrounding environment, in a manner limited by the range and sensitivity of the sensory apparatus. As information is compressed, a single system state is realized and all other possible system states are discarded. Thus, the realization of a multi-modal ‘percept’ corresponds to the synchronous firing of neurons across the network, at a periodic frequency tied to the length of the computation. Yet the syntactical relationship of that semantical statement to ensuing events remains uncertain. As a result, these individual computations, yielding ‘percepts’, may contribute to temporal sequences of neural network states, paired with information content, yielding a predictive cognitive model of causal relationships between perceived events. As predictive value is extracted, information is compressed over these longer timescales. Thus, these sequences of computations are predicted to result in synchronous firing at much slower frequencies. During wakeful awareness, both semantical and syntactical information is being parsed, so both fast and slow oscillations should be observed. This prediction of synchronous neural activity at nested frequencies is consistent with empirical observations in the cerebral cortex [64-66]. During sleep, incoming sensory data is not being evaluated and so only syntactical relationships between previously-stored information is being parsed; as a result, sleep should be characterized by synchronous activity only at slower frequencies. This prediction is also consistent with empirical observation [67, 68]. These synchronous firing events are predicted to be disrupted by the absorption of thermal free energy release at relevant frequencies, or the random introduction of free energy.

2. *Thermodynamic computation drives spontaneous remodeling of the encoding structure*

The quantity of predictive value extracted during a computation is equivalent to the amount of information entropy compressed during that computation; this quantity must be exactly balanced by the amount of free energy released into the system and the amount of work then dissipated toward creating a more ordered system state. This work is expected to involve locally optimizing the system to encode the acquired predictive value (e.g. by remodeling the neural network structure to thermodynamically favor a certain pattern of neural activity in a predictive context). In short, *the extraction of predictive value* should be paired with *the dissipation of work* toward creating a more ordered system state *which encodes that predictive value*. Of course, cortical neurons are already known to spontaneously remodel their synaptic connections to encode learned information [9]. This structural remodeling is prompted by participation in network-wide synchronous activity [19] and has been shown to affect each synapse in proportion to its initial contribution [10]. While the molecular mechanisms underlying synaptic plasticity are well-established, this theoretical model describes the relationship between free energy, entropy minimization, and spontaneous directed work, as uncertainty is resolved. Here, it is predicted that spontaneous self-remodeling is a natural output of a thermodynamic computation, in which predictive synaptical relationships are encoded in sequences of neural network states. Any unique sequence of neural activity is therefore expected to uniquely match representational information content. Alterations in information content should be paired with a conformational change in the encoding structure, to store the freshly-parsed information; conversely, any impairment to the encoding structure should be paired with a loss of stored information content. If instead the null hypothesis is true, and activity-dependent synaptic remodeling can be entirely explained by classical mechanisms, then this process should *not* be localized to synapses that have reduced uncertainty and should *not* be dependent on parsing predictive value to maximize free energy availability.

3. *Thermodynamic computation yields an exploration/exploitation dynamic in behavior*

In this model, the system identifies a compatible state with its surrounding environment, by maximizing predictive value or consistency. To gain predictive value, it is necessary to collect data from the environment, parse that dataset for predictive value, compress information, and subsequently encode that predictive value into the encoding structure, so this encoding system state is favored to repeat in a similar context. It is useful to attend to any incoming data that does not match existing predictive models, because this discrepancy signals a valuable

opportunity to update the cognitive model with this new information. Therefore, a decision on how to act, given the incoming information, will rely on a rival energetic cost function: If knowledge is predicted to be *sufficient* to navigate the situation, the organism is unlikely to engage in exploration to gain additional predictive capacity. If knowledge is predicted to be *insufficient* to navigate the situation, the organism is likely to engage in exploration. Any change in the predicted risk of energy dissipation for exploration or non-exploration will shift the probability of exploration.

4. *Thermodynamic computation carries the risk of predictive error*

This model of thermodynamic computation is extraordinarily energy-efficient. However, two types of inefficiencies can occur. If existing knowledge is not exploited in familiar situations, time and energy may be wasted. If the consistency and predictive power of a new cognitive model is ignored in favor of holding onto a previous cognitive model, time and energy may also be wasted. While the former leads to unnecessary dissipation of energy toward exploration, the latter leads to potentially missing critical information. Some organisms may select the short-term savings gained by holding onto a previous cognitive model over the potential long-term energetic savings gained by deconstructing that cognitive model and constructing an improved model in its place.

To explore what might be true about the world, it is necessary to sacrifice time and energy resources for uncertain reward, with the prediction that this process will be worthwhile in the long term. In this theory, mental exploration takes thermodynamic work – and some individuals may opt not to engage in the process, with the prediction there will be no payoff in doing so.

Exploration exposes the organism to uncertainty, thereby generating information with potential predictive value, while exploitation of knowledge that was previously gained through exploration makes good use of that predictive value. Ignoring incoming information may reduce energetic costs in the short term, but this strategy can impair prospects for survival in the longer term, so it is useful to accurately assess when knowledge is insufficient. The organism must compute the likely energetic cost of exploration against the potential cost of making a predictive error. For this reason, it is predicted that an individual capability for parsing information – and discarding inaccurate cognitive models in favor of more accurate cognitive models – will provide a selective advantage, which should lead to measurable differences in survival outcomes when environmental conditions suddenly shift.

IV. DISCUSSION

The extraordinary energetic efficiency of the central nervous system has been noted, particularly amongst theorists who query whether this competence is intrinsically linked to the physical production of information entropy [69, 70] or the exascale computing capacity of the brain [32, 71]. In this model, energy is indeed being expended on entropy production – but rather than being irreversibly lost, this quantity is parsed during a periodic computational cycle of information generation and compression. Since unlikely or unfamiliar system states may hold predictive value, the quantity of entropy may be expanded to accommodate a new system state, even at an energetic cost, in order to reduce uncertainty in the longer term.

Interestingly, there are limits to the amount of predictive value in a quantity of information. These limits are given by the first law of thermodynamics. Indeed, any system that takes on a more ordered state has essentially reduced its entropy over the course of a computational cycle. But the amount of free energy that is dedicated toward ordering the system state cannot be greater than the amount of free energy that is gained through information compression. Some amount of work must also be dissipated to uncertainty, since having complete predictive value and continued operation is thermodynamically impossible. As such, there are intrinsic thermodynamic limits to the amount of work being done by the system and thermodynamic limits to the completeness of its predictive capacity.

A single computational cycle will both resolve the neural network state in the present moment and generate a semantic statement about the external environment, which is validated by orienting toward incoming sensory information. These semantical statements are held in working memory, and then integrated with subsequent neural network states to compute any syntactical relationships. Holding untrue semantic statements as true, and attempting to use these statements to form syntactical relationships with other semantical statements, is equivalent to energetic inefficiency. Yet information content which may have accurately reflected reality in some context could simply be inaccurate in a different context; the ability of the system to recognize the difference is largely determined by how effectively the neural network structure had been remodeled to encode that information. If the neural network remains adaptable, not completely ordered, more nuanced complexities can be noted and more predictive value can be gained, at some energetic expense.

This process of thermodynamic computation results in a rival energetic cost function, to maximize free energy. This rival energetic cost function drives a system to either *maximize* entropy production through exploration, to gain predictive value, or *minimize* entropy production through exploitation, to make effective use of previously-acquired predictive value. The system must first expend

energy to generate information, then recover energy by compressing information. That free energy is then used to remodel the encoding structure itself, to store that predictive value, thus thermodynamically favoring the recurrence of sequential system states in a relevant or familiar context. The maximization of free energy is readily accomplished through cycles of information generation and compression, or by relying on the optimized system state that was already achieved by that process. The timescales of computations may overlap, with individual computations that encode multi-sensory percepts being integrated into short sequences of neural network states which encode possible relationships between perceived events. The overall temperature and energy flux of the system remain constant, permitting the non-equilibrium system to act as a net heat sink. In this view, neural networks – acting in accordance with the laws of thermodynamics – are expected to physically process information, undergoing non-deterministic computation.

Physically compressing information, by extracting predictive value, yields a single neural network state, encoding the likely present state of the local environment. The incompleteness of information compression in any single computational cycle leaves some entropy remaining, *which can then be parsed for predictive value over longer timescales*. By integrating information over longer timescales, the neural network may integrate temporal sequences of *consistent* system states to build predicted cause-effect relationships between perceived events. These mental models – which make predictions about the general structure and operation of the physical world; the likely behavior of other people, animals, and things in the environment; and the most effective way of acting in response to certain stimuli – are expected to be formed by thermodynamic computations occurring across the lifetime.

There is a long-standing ontological issue regarding the relationship between information and entropy. Both concepts describe the sum of all possible system states, so resolving this issue lies at the heart of thermodynamic computing. Biological neural networks are thermodynamic systems that produce entropy, like any other particle system. Yet unlike other entropy-producing thermodynamic systems – such as the steam engine – a biological neural network computes information, reducing disorder or randomness into a more optimal system state, with the system growing increasingly ordered over time as it interacts with its environment. In this model, neural networks both *generate information* and *parse that information for predictive value*. It is useful here to consider ‘information’ as both the sum of all possible system states (its mathematical definition) and the *meaning* extracted from a messy dataset (its colloquial definition). In this theoretical model, these are two stages of the same computational process. As neural networks produce von Neumann entropy, they are physically producing disorder – a set of possible system states, or the mathematical definition of information. However that information is only

useful if predictive value is extracted, a distribution of system states is compressed into a single outcome, free energy is released back into the system, and that free energy is used to remodel the encoding structure to store that predictive value. In other words, any thermodynamic system can create a distribution of possible system states, thereby producing *entropy*. But only a thermodynamic system capable of compressing that probability distribution into a single actualized system state will then parse *consistency* or *predictive value* from that information, store that predictive value for future use, and recall those predictions in relevant contexts.

Here, the outcomes of cortical neuron computations are expected to be non-deterministic, but *not* statistically random; they will naturally encode the ‘best match’ between the system and its local environment. This approach fits well with other recent models emphasizing the

role of contextual cues [72] and cost calculations [73] in decision-making. Usefully, this new model introduces a thermodynamic basis for predictive processing, with the extraction of predictive value physically minimizing energy dissipation. As a result, any far-from-equilibrium thermodynamic system that traps heat to perform computational work will inevitably cycle between expanding the distribution of system states and pruning toward a more optimized system state. This computational process naturally yields an exploration/exploitation dynamic in behavior and the gradual emergence of a more ordered system state over time.

ACKNOWLEDGMENTS

The author declares no interests. The author will report funding sources after anonymous peer review.

-
- [1] Hubel, D.H. and Wiesel, T.N. (1959) Receptive fields of single neurones in the cat’s striate cortex. *J Physiol* 148: 574–91.
- [2] Fettiplace, R. (2017) Hair cell transduction, tuning and synaptic transmission in the mammalian cochlea. *Compr Physiol* 7(4): 1197–227.
- [3] Schultz, W. and Dickinson, A. (2000) Neuronal Coding of Prediction Errors. *Ann Rev Neurosci* 23: 473–500.
- [4] Lochmann, T. and Deneve, S. (2011) Neural processing as causal inference. *Current Opinion in Neurobiology* 21(5): 774–81.
- [5] Arnal, L.H., Wyart, V., and Giraud, A.L. (2011) Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nat Neurosci* 14: 797–801.
- [6] Wimmer, K., et al. (2015) Sensory integration dynamics in a hierarchical network explains choice probabilities in cortical area MT. *Nat Commun* 6:177
- [7] Melloni, L., et al., (2011) Expectations change the signatures and timing of electrophysiological correlates of perceptual awareness. *J Neurosci* 31(4): 1386–96.
- [8] Kok, P., Jehee, J.F. and de Lange, F.P. (2012) Less is more: expectation sharpens representations in the primary visual cortex. *Neuron* 75(2): 265–70.
- [9] Hebb, D.O. (1949) The Organization of Behavior. *New York: Wiley and Sons* Chapter 6.
- [10] Turrigiano, G.G., et al. (1998) Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature* 391(6670): 892–6.
- [11] Nishimura, Y., et al. (2013) Spike-timing-dependent plasticity in primate corticospinal connections induced during free behavior. *Neuron* 80(5): 1301–9.
- [12] Urbin, M.A., et al. (2017) Spike-timing-dependent plasticity in lower-limb motoneurons after human spinal cord injury. *J Neurophysiol* 118(4): 2171–80.
- [13] Ting, W.K., Huot-Lavoie, M., and Ethier, C. (2020) Paired Associative Stimulation Fails to Induce Plasticity in Freely Behaving Intact Rats. *eNeuro* 7(2): 0396-19-2020.
- [14] Simonetti, M. and Kuner, R. (2020) Spinal Wnt5a Plays a Key Role in Spinal Dendritic Spine Remodeling in Neuropathic and Inflammatory Pain Models and in the Proalgesic Effects of Peripheral Wnt3a. *J Neurosci* 40(35): 6664–77.
- [15] Tan, A.M., et al. (2008) Neuropathic pain memory is maintained by Rac1-regulated dendritic spine remodeling after spinal cord injury. *J Neurosci* 28(49): 13173–83.
- [16] Bradley, P.M., et al. (2019) Corticospinal circuit remodeling after central nervous system injury is dependent on neuronal activity. *J Exp Med* 216(11): 2503–14.
- [17] Huang, C.S., et al. (2005) Common molecular pathways mediate long-term potentiation of synaptic excitation and slow synaptic inhibition. *Cell* 123(1): 105–18.
- [18] Kapur, A., et al. (1997) GABAA-mediated IPSCs in piriform cortex have fast and slow components with different properties and locations on pyramidal cells. *J Neurophysiol* 78(5): 2531–45.
- [19] Zarnadze, S., et al. (2016) Cell-specific synaptic plasticity induced by network oscillations. *Elife* 5:e14912
- [20] Gotts, S.J. (2016) Incremental learning of perceptual and conceptual representations and the puzzle of neural repetition suppression. *Psychon Bull Rev* 23(4): 1055–71.
- [21] Kitano, K. and Fukai, T. (2004) Temporal characteristics of the predictive synchronous firing modeled by spike-timing-dependent plasticity. *Learn Mem* 11(3): 267–76.
- [22] Magistretti, P. and Pellerin, L. (1999) Cellular mechanisms of brain energy metabolism and their relevance to functional brain imaging. *Philos Trans Roy Soc Lond B* 354: 1155–63.
- [23] Sokoloff, L. (1981) Localization of functional activity in the central nervous system by measurement of glucose utilization with radioactive deoxyglucose. *J Cereb Blood Flow Metab* 1: 7–36.
- [24] Dorval, A.D. and White, J.A. (2005) Channel noise is essential for perithreshold oscillations in entorhinal stellate neurons. *J Neurosci* 25(43): 10025–8.
- [25] Stern, E.A., Kincaid, A.E., and Wilson, C.J. (1997) Spontaneous subthreshold membrane potential fluctuations and action potential variability of rat corticostriatal and striatal neurons in vivo. *J Neurophysiol* 77(4): 1697–715.

- [26] Lee, D., Port, N.L., Kruse, W. and Georgopoulos, A.P. (1998) Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex. *J Neurosci* 18(3): 1161–70.
- [27] Averbeck, B.B., Latham, P.E. and Pouget, A. (2006) Neural correlations, population coding and computation. *Nat Rev Neurosci* 7: 358–66.
- [28] Adams, R.A., Shipp, S. and Friston, K.J. (2013) Predictions not commands: active inference in the motor system. *Brain Struct Funct* 218(3): 611–43.
- [29] Friston, K. and Kiebel, S. (2009) Predictive coding under the free-energy principle. *Philos Trans R Soc Lond B* 364(1521): p. 1211–21.
- [30] Feldman, H. and Friston, K. (2010) Attention, Uncertainty, and Free-Energy. *Front Hum Neurosci* 4: 215.
- [31] Friston, K., J. Kilner, and Harrison, L. (2006) A free energy principle for the brain. *J Physiol* 100(3): 70–87.
- [32] Bellec, G., et al., (2020) A solution to the learning dilemma for recurrent networks of spiking neurons. *Nat Commun* 11(1): 3625.
- [33] Scellier, B. and Bengio, Y. (2017) Equilibrium propagation: Bridging the gap between energy-based models and backpropagation. *Front Comput Neurosci* 11: 24.
- [34] Hopfield, J.J. (1987) Learning algorithms and probability distributions in feed-forward and feed-back networks. *Proc Natl Acad Sci U S A* 84(23): 8429–33.
- [35] Beck, J.M., et al. (2008) Probabilistic population codes for Bayesian decision making. *Neuron* 60(6): 1142–52.
- [36] Maoz, O., et al. (2020) Learning probabilistic neural representations with randomly connected circuits. *Proc Natl Acad Sci U S A* 117(40): 25066–73.
- [37] Landauer, R. (1961) Irreversibility and heat generation in the computing process. *IBM J Res Devel* 5: 183–91.
- [38] Berut, A., et al. (2012) Experimental verification of Landauer’s principle linking information and thermodynamics. *Nature* 483(7388): 187–9.
- [39] Jun, Y., Gavrilov, M., and Beckhoefer, J. (2014) High-precision test of Landauer’s principle in a feedback trap. *Phys Rev Lett* 113: 190601.
- [40] Yan, L.L., et al. (2018) Single-atom demonstration of the quantum Landauer principle. *Phys Rev Lett* 120: 210601.
- [41] Bengtsson, I. and Życzkowski, K. (2007) Geometry of Quantum States, First Edition. *Cambridge: Cambridge University Press* Chapter 12.
- [42] Bennett, C.H. and Shor, P.W. (1998) Quantum Information Theory. *IEEE TRANSACTIONS ON INFORMATION THEORY* 44(6): 2724–42.
- [43] Jozsa, R., et al. (1998) Universal quantum information compression. *Phys Rev Lett* 81: 1714.
- [44] Schumacher, B. (1995) Quantum coding. *Phys Rev A* 51: 2738.
- [45] Sciarra, S., Franco, R.L., and Compagno, G. (2017) Universality of Schmidt decomposition and particle identity. *Sci Rep* 7: 44675.
- [46] Jarzynski, C. (1997) Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. *Phys Rev E* 56: 5018.
- [47] Jarzynski, C. (1997) Nonequilibrium equality for free energy differences. *Phys Rev Lett* 78: 2690.
- [48] Crooks, G.E. (1998) Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems. *J Stat Phys* 90: 1481.
- [49] Callen, H.B. and Welton, T.A. (1951) Irreversibility and Generalized Noise. *Phys Rev* 83: 34.
- [50] Still, S., et al. (2012) Thermodynamics of prediction. *Phys Rev Lett* 109(12): 120604.
- [51] Sonntag, R.E., Borgnakke C., and Van Wylen, G.J (2003) Fundamentals of Thermodynamics, Sixth Edition. *New York: Wiley and Sons* Chapter 13.
- [52] Martyushev, L.M., Nazarova, A.S., and Seleznev, V.D. (2007) On the problem of minimum entropy production in the nonequilibrium stationary state. *Journal of Physics A* 40(3): 371–80.
- [53] Hillert, M. and Agren, J. (2006) Extremum principles for irreversible processes. *Acta Materialia* 54(8): 2063–6.
- [54] Glandsdorff, P. and Prigogine, I. (1971) Thermodynamic theory of structure, stability, and fluctuations. *London: Wiley-Interscience* Chapter 8.
- [55] Gödel, K. (1931) Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I, Translated into English in 2000 by M. Hirzel. *Monatshefte für Mathematik Physik* 38: 173–98.
- [56] Raatikainen, P. (2015) Gödel’s Incompleteness Theorems. *Stanford Encyclopedia of Philosophy*
- [57] Hilbert, D. and Ackermann, W. (1928) Grundzüge der theoretischen Logik, Translated into English in 1999 by L.M. Hammond, G.G. Leckie, and F. Steinhardt. *Berlin: Springer-Verlag*
- [58] Church, A. (1936) A Note on the Entscheidungsproblem. *J Symb Logic* 1(1): 40–1.
- [59] Turing, A.M. (1936) On Computable Numbers, with an Application to the Entscheidungsproblem. *Proc Lond Math Soc* 42(1): 230–65.
- [60] . Under Review .
- [61] . Under Review .
- [62] . Under Review .
- [63] . Under Review .
- [64] Buzsaki, G. and Draguhn, A. (2004) Neuronal oscillations in cortical networks. *Science* 304(5679): 1926–9.
- [65] Engel, A.K. and Singer, W. (2001) Temporal binding and the neural correlates of sensory awareness. *Trends Cogn Sci* 5(1): 16–25.
- [66] Harris, A.Z. and Gordon, J.A. (2015) Long-range neural synchrony in behavior. *Ann Rev Neurosci* 38: 171–94.
- [67] Mitra, A., Snyder, A.Z., Tagliazucchi, E., Laufs, H. and Raichle, M.E. (2015) Propagated infra-slow intrinsic brain activity reorganizes across wake and slow wave sleep. *ELife* 4:10781.
- [68] Steriade, M., Timofeev, I. and Grenier, F. (2001) Natural waking and sleep states: a view from inside neocortical neurons. *J Neurophysiol* 85(5): 1969–85.
- [69] Collell, G. and Fauquet, J. (2015) Brain activity and cognition: a connection from thermodynamics and information theory. *Front Psychol* 6: 818.
- [70] Street, S. (2016) Neurobiology as Information Physics. *Front Syst Neurosci* 10: 90.
- [71] Keyes, D.E. Ltaief, H., and Turkiyyah, G. (2020) Hierarchical algorithms on hierarchical architectures. *Philos Trans A Math Phys Eng Sci* 378(2166): 20190055.
- [72] Bruza, P.D. et al. (2023) Contextuality and context-sensitivity in probabilistic models of cognition. *Cogn Psychol* 140: 101529.
- [73] Jara-Ettinger, J., Schulz, L.E. and Tenenbaum, J.B. (2020) The Naïve Utility Calculus as a unified, quantitative framework for action understanding. *Cogn Psychol* 123: 101334.