Control flow in active inference systems Part II: Tensor networks as general models of control flow

Chris Fields[a*], Filippo Fabrocini[b,c], Karl Friston[d,e], James F. Glazebrook[f,g], Hananel Hazan[a], Michael Levin[a,h] and Antonino Marcianò[i,j,k]

[a] *Allen Discovery Center at Tufts University, Medford, MA 02155 USA*

[b] *College of Design and Innovation, Tongji University, 281 Fuxin Rd, 200092 Shanghai, CHINA*

[c] *Institute for Computing Applications "Mario Picone", Italy National Research Council, Via dei Taurini, 19, 00185 Rome, ITALY*

[d] *Wellcome Centre for Human Neuroimaging, University College London, London, WC1N 3AR, UK*

[e] *VERSES Research Lab, Los Angeles, CA, 90016 USA*

[f] *Department of Mathematics and Computer Science, Eastern Illinois University, Charleston, IL 61920 USA*

[g] *Adjunct Faculty, Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA*

[h] *Wyss Institute for Biologically Inspired Engineering at Harvard University, Boston, MA 02115, USA*

[i] *Center for Field Theory and Particle Physics & Department of Physics Fudan University, Shanghai, CHINA*

[j] *Laboratori Nazionali di Frascati INFN, Frascati (Rome), ITALY*

[k] *INFN sezione Roma "Tor Vergata", I-00133 Rome, ITALY*

April 24, 2023

Abstract

Living systems face both environmental complexity and limited access to free-energy resources. Survival under these conditions requires a control system that can activate, or deploy, available perception and action resources in a context specific way. In the accompanying Part I, we introduced the free-energy principle (FEP) and the idea of active inference as Bayesian prediction-error minimization, and show how the control problem arises in active inference systems. We then reviewed classical and quantum formulations of the FEP, with the former being the classical limit of the latter. In this Part II, we show that when systems are described as executing active inference driven by the FEP, their control flow systems can always be represented as tensor networks (TNs). We show how TNs as control systems can be implemented within the general framework of quantum topological neural networks, and discuss the implications of these results for modeling biological systems at multiple scales.

Keywords

Bayesian mechanics; Dynamic attractor; Free-energy principle; Quantum reference frame; Scale-free model; Topological quantum field theory

1 Introduction

The framework of *active inference* provides a completely general, scale-free formal framework for describing interactions between physical systems in cognitive terms. In Part I of

*Corresponding author at: Allen Discovery Center at Tufts University, Medford, MA 02155 USA; *E-mail address*: fieldsres@gmail.com

this paper, we reviewed how active inference – a combination of learning with active exploration of the environment – emerges in systems compliant with the Free Energy Principle (FEP), a general least-action principle initially developed in neuroscience [1, 2, 3, 4, 5, 6, 7]. We then showed how the control flow problem arises in active inference systems, and reviewed classical and quantum formulations of the problem. Control flow can be represented as switching between classical dynamical attractors, between deployed quantum reference frames (QRFs) [8, 9], and between computational processes represented by TQFTs [10, 11]. Implementing control flow has a free-energy cost; hence any control-flow system must trade off its own processing costs against the expected benefits of switching between input/ouput modes. The time and memory dependence of control flow generically leads to context effects on both perception and action.

In this Part II, we develop a fully-general tensor representation of control flow in §2, and prove that this tensor can be factored into a TN if, and only if, the separability (or conditional statistical independence) conditions needed to identify distinct features of, or objects in, the environment are met. We show how TN architectures allows classification of control flows, and give two illustrative examples. We then discuss several established relationships between TNs and artificial neural network (ANN) architectures in §3, and show how these generalize to topological quantum neural networks [11, 12], of which standard deep-learning (DL) architectures are a classical limit [13]. Having developed these formal results, we turn to implications of these results for biology in §4, and discuss how TN architecture correlates with the observational capabilities of the system being modeled, particularly as regards abilities to detect spatial locality and mereology. We consider how to classify known control pathways in terms of TN architecture and how to employ the TN representation of control flow in experimental design. We conclude by looking forward to how these FEP-based tools can further integrate the physical and life sciences.

2 Tensor network representation of control flow

2.1 Tensor networks and holographic duality

Entanglement and quantum error correction, two concepts developed in quantum information theory, have been proved to have a fundamental role in unveiling quantum gravity [14]. At the origin of this consideration is the discovery by Bekenstein and Hawking [15, 16, 17, 18] that the second law of thermodynamics can be preserved in the gravitational field of a black hole, if this latter has an entropy proportional to the area of its horizon, by the inverse of the Newton gravitational constant G. This entropy is maximal, as implied by the second law itself, providing an upper bound for possible configurations of matter within a region of the same size [19, 20].

Nonetheless, the scaling of the local degrees of freedom counted by the entropy does not increase as the volume, hinging toward the formulation of the holographic conjecture [21], suggesting a division between the information that can only be retrieved on the boundary world, and a merely apparent bulk world. AdS/CFT realized the holographic conjecture, postulating a duality between gravity in asymptotically AdS space and quantum field theory on the spatial infinity of the AdS space [22]. Giving literal meaning to the duality, Ryu and Takayanagi (RT, [23]) proposed that entanglement of a boundary region fulfils the same law as for the black hole entropy, replacing the area of the black hole horizon with an extremal surface area that bounds the bulk region under scrutiny.

While on the boundaries the theory can be individuated by assigning a specific conformal field theory (CFT), in the bulk the geometry can be associated to specific entanglement structures of the quantum systems. This is, for instance, what happens to the ground states of a CFT associated to an AdS space: the RT surface area increases less fast than the volume of the boundary. When the boundary is at equilibrium, in a thermal state of finite temperature, the bulk geometry corresponds to that of a black hole, its horizon being parallel to the boundary and its size increasing with the temperature. The RT surface is then confined between the boundary and the back hole horizon, approaching the boundary at higher temperature and increasing its entropy. These considerations suggest the existence of a subtle link interconnecting the structure of spacetime and quantum entanglement, and hence that a theory of quantum gravity must be fundamentally holographic, where its states satisfy the RT formula for some bulk geometry.

The existence of an exact correspondence between bulk gravity and quantum theory at the boundary may hinge toward possible inconsistencies with locality. This has been discussed in the literature, in terms of local reconstruction theory [24, 25, 26]: variables in the bulk (e.g. bulk spins) can be controlled instantaneously from the boundary, but this requires simultaneous access to a large portion of the boundary: locality and an upper speed of light do not hold exactly in this theory. Nonetheless, local observers confined in small regions at the boundary still fulfill locality and the existence of an upper limit of the speed of information exchange, in a way that is reminiscent of quantum error correction codes (QECCs) in quantum information theory: information is stored redundantly, in such a way that when part of it is corrupted, a reconstruction of information is still possible. Locality in the bulk is therefore a QECC property of the encoding map that realizes the duality between bulk and boundary. On the other hand, these properties are strictly connected to RT, which provides the necessary resource of entanglement for QEEC to emerge [27].

The RT formula and QECC are properties fulfilled by different classes of models, among which TNs [28]. These have been first introduced in condensed matter physics as variational wave-functions of strongly correlated systems [29, 30]. TNs are many-body wavefunctions that can be derived by composing few-body quantum states, which are indeed tensors. A prototype TN is, e.g., a collection of Einstein-Podolsky-Rosen (EPR) entangled pairs of qubits: in a nonentangled basis, the measured qubits are in some entangled pure state, and can be composed with additional qubits to create states with increasing complexity. Indeed, complicated quantum entanglement can be derived by entangling only a few qubits [31].

Particularly relevant for its implications on the reconstruction of spacetime structure is the multi-scale entanglement renormalization ansatz (MERA) [32]. TNs can be naturally related to holography duality by considering that their entanglement entropy can be controlled by their graph geometry. Some versions of TNs that are characterized by RT entanglement entropy and QEEC have been constructed resorting to stabilizer codes [33, 34] and random tensors with large bond dimensions [35]. TNs with random tensors at each node can be regarded as random states restricted by the topology of the network. Exactly as random states are almost maximally entangled, random TNs show, through the RT formula, an almost maximal entanglement, providing a large family of states with interesting properties to explore holographic duality. Furthermore, for random TNs, the RT formula holds in generic spaces with not necessarily hyperbolic geometry, hinging toward an extension of holographic duality beyond AdS, to more general configurations in quantum gravity. Nonetheless, at least in three dimensions, random tensor networks have been related to the gravitational action, by means of the Regge calculus [36].

On the other hand, since geometry emerges as a specification of the entanglement structure, one may consider that the Einstein equations should be connected as well to the dynamics of entanglement. For small perturbations around the ground state of a CFT on a boundary, linearized Einstein equations have been derived from the RT formula [37, 38]. Indeed, the conformal symmetry enables a relationship between the energy-momentum and the entanglement entropy, and consequently the area of the extremal surface can be connected to the energy-momentum distribution at the boundary – the result is equivalent to the linearized Einstein equations.

The dynamics on the boundary, on the other hand, shows a chaotic behaviour, with scrambling of the single-particle operators, which evolve into multi-particle operators [39]. Maximal chaotic behavior recovered in the growth of the commutator between ladder operators, as encoded in the out-of-time-ordered correlation (OTOC) functions, is characterized by exponential growth in time and temperature. A model endowed with this property is, e.g., the Sachdev-Ye-Kitaev model, developed to describe certain systems in condensed matter physics, such as Gapless spin-fluids [40, 41, 42]. On the other hand, operator scrambling is also related to QEEC: the chaotic dynamics at the boundary instantiates QECC preserving quantum information, which is efficiently hidden (and protected) behind the horizon. Nevertheless, this has led to many questions concerning the information behind the horizon being eventually accessible from the boundary though non-local measurements, the fate of the local degrees of freedom hitting the singularity, and the relation between the causal structure of the bulk and the smooth geometry across the horizon.

2.2 General results

We can now move to prove a general result:

Theorem 1. *A system A exhibits non-trivial control flow if, and only if, its control flow can be represented by a TN.*

and examine some of its corollaries. We begin by defining:

Definition 1. *Control flow is* trivial *if a system deploys only one QRF.*

As any collection of mutually-commuting QRFs can be represented as a single QRF [11, 79], any system that deploys only mutually-commuting QRFs exhibits trivial control flow.

Systems that deploy only a single QRF "do the same thing" regardless of context, and so do not qualify as "interesting" in the sense used here. As noted above, no finite physical system can measure the entire state of its boundary with a single QRF, so no such system can simultaneously measure and act on its entire context. Any system A that deploys multiple QRFs Q_i in sequence cannot, as noted in Part I, avoid contextuality due to unobservable effects, mediated by the action of H_B, of the action of Q_i on the state later measured by Q_j. Every action taken by an "interesting" system, in other words, at least transiently increases the VFE at its boundary.

Consider, then, a system A that deploys multiple, distinct QRFs $Q_1, Q_2, \ldots Q_n$ acting on its environment B, where $n \ll N = \dim(H_{AB})$. Classical control flow in A can then be represented by a matrix $\mathbf{CF} = [P_{ij}]$, where P_{ij} is the probability of the control transition $Q_i \rightarrow Q_j$. As noted in Part I, any such transition has an energetic cost, which must be paid with free energy sourced from the thermodynamic sector F of the A-B boundary $\mathscr{B}$.

The matrix $\mathbf{CF}$ is a 2-tensor. Theorem 1 states that this tensor can be decomposed into a TN. We prove it as follows:

Proof (Thm. 1). Suppose first that control flow in a system A can be represented by a TN. A TN is, by definition, a factorization of a tensor operator into a network of tensor operators. This network can be either hierarchical or flat; if it is hierarchical, each layer can be considered a flat TN. Hence no generality is lost in considering just the case of a flat TN, which is an operator contraction $T = \ldots T_{ij}T_{jk}T_{kl}\ldots$, where summation on shared indices is left implicit. In general, $T_{jk} \neq T_{jk}^T = T_{kj}$, hence these expressions do not commute. They therefore represent non-trivial control flow. Conversely, any non-trivial control flow can be written, at any fixed scale or level of abstraction, as a linear sequence of (in general probabilistic) operators. The fixed order of operators in the sequence can be encoded formally by adding "spatial" indices as needed to allow contraction over shared indices. Hence any non-trivial control flow at a fixed scale can be written as a flat TN. This construction can be repeated at each larger scale to produce a hierarchical TN over a collection of "lowest-scale" TNs. $\square$

We can now examine two corollaries of this result:

Corollary 1. *Decoherent reference sectors exist on a boundary $\mathscr{B}$ if and only if control flow can be implemented by a TN.*

Proof. Decoherence between sectors requires independently-deployable, non-commuting QRFs. This requires a control structure that factors, hence by Theorem 1, it requires a TN. Conversely, a TN factors the control structure, making QRFs independently deployable, which renders their sectors decoherent. □

Equivalently, the generative model (GM) implemented by a system [4] factors if, and only if, control flow can be implemented by a TN.

Corollary 2. *The TN of any system compliant with the FEP is a decomposition of the Identity.*

Proof. The FEP applies to systems with a NESS, and drives such systems to return to (the vicinity of) the NESS after any perturbation. Hence at a sufficiently large scale, the TN of any such system is a cycle, i.e., a decomposition of the Identity. □

Many standard TN models, e.g., MERAs, assume boundary conditions asymptotically far, in numbers of lowest-scale operators, from the region of the network that is of interest. Identifying such asymptotic boundary conditions yields a cyclic system.

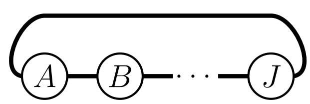
Theorem 1, together with its corollaries, provides a natural, formal means of classifying systems by their control architectures. At a high level, two characteristics distinguish systems with different architectures:

- Hierarchical depth, which indicates the number of "virtual machine" layers [43] the architecture supports. The interfaces between these layers implement coarse-graining, removing from the higher-level representation all dimensions, and hence all information, which is contracted out of the lower-level operators.

- Number and location of contractions that yield unitary operators, and hence build in entanglement between lower-level operators. The natural limit is a MERA, in which every pair of lower-level operators is entangled at every hierarchical level [44].

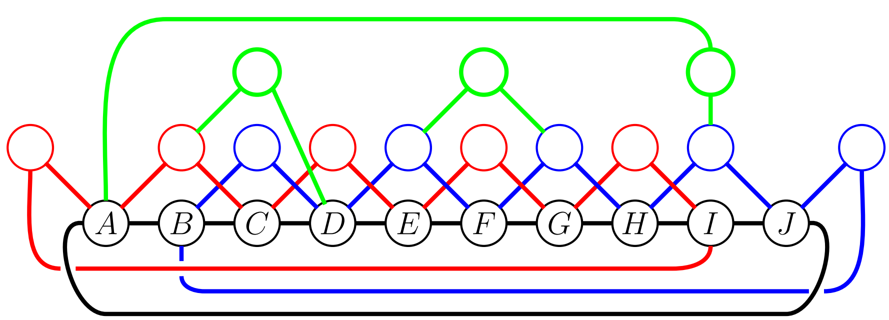
The control-flow architecture, in turn, specifies the structure of the "layout" of distinguishable sectors on $\mathscr{B}$ and hence of detectable features/objects in the environment. Locality on $\mathscr{B}$ requires a hierarchical TN; detectable entanglement requires a MERA-like TN. Locality is required for detectable features/objects to appear to have components with nested decompositions. Any QRF for geometric space, and hence for spacetime, must be hierarchical, and must be a MERA if entanglement in space is to be detected. A MERA is required, in particular, if the use of coherence between spatially-separated systems as a computational or communication resource is detectable.

To illustrate the classification of systems by hierarchical level, consider the ten-step cyclic TN shown in Diagram (1):



(1)

and its extension to a hierarchy as shown in Diagram (2):



(2)

where red, blue, and green colors indicate distinct hierarchical “layers” of tensor contractions. We have trained artificial neural networks (ANNs) to execute these TNs as the sequences of state transitions shown in Table 1. The first sequence (Dataset 1) is a ten-step cycle as shown Diagram (1); the second sequence (Dataset 2) layers the coarse-grained state transitions of Diagram (2) onto this ten-step cycle. In Dataset 2, a two-bit tag is used to differentiate the “low-level” from the coarse-grained “high-level” cycles. An example state state transition from a randomly-generated initial state is shown in Fig. 1; the red-on-green bit pattern effectively moves “up” one step on each state-transition cycle.

Dataset 1	Dataset 2			
A → B	00 A → B	01 A → C	10 B → D	11 A → D
B → C	00 B → C	01 C → E	10 D → F	11 D → H
C → D	00 C → D	01 E → G	10 F → H	11 H → A
D → E	00 D → E	01 G → I	10 H → J	
E → F	00 E → F	01 I → A	10 J → B	
F → G	00 F → G			
G → H	00 G → H			
H → I	00 H → I			
I → J	00 I → J			
J → A	00 J → A			

Table 1: Datasets used in ANN simulations. Dataset 1 specifies a ten-step cycle A $\rightarrow$ B $\rightarrow$ $\ldots$ $\rightarrow J \rightarrow A$. Dataset 2 specifies this same cycle, with three coarse-grained cycles layered on top. The tags (0,0), (0,1), (1,0), and (1,1) distinguish the data for the low- and high-level cycles.

INPUT (T)												OUTPUT (T+1)										
A	1	1	1	0	0	0	1	0	0	1	→	B	1	1	1	0	1	1	1	1	0	1
B	1	1	1	0	1	1	1	1	0	1	→	C	0	1	0	0	1	1	0	0	0	1
C	0	1	0	0	1	1	0	0	0	1	→	D	1	1	1	1	1	0	0	1	0	1
D	1	1	1	1	1	0	0	1	0	1	→	E	1	1	0	0	0	0	0	0	0	1
E	1	1	0	0	0	0	0	0	0	1	→	F	1	1	1	1	0	1	1	0	0	1
F	1	1	1	1	0	1	1	0	0	1	→	G	1	1	1	0	0	1	1	0	1	0
G	1	1	1	0	0	1	1	0	1	0	→	H	0	1	0	1	0	0	0	1	0	1
H	0	1	0	1	0	0	0	1	0	1	→	I	1	0	1	0	0	0	0	1	0	1
I	1	0	1	0	0	0	0	1	0	1	→	J	0	1	1	1	1	1	1	1	0	1
J	0	1	1	1	1	1	1	1	0	1	→	A	1	1	1	0	0	0	1	0	0	1

Figure 1: Example state transition from Dataset 1.

We trained two ANNs, one to execute each of the control cycles shown in Table 1. The networks are each composed of three layers, as illustrated in Fig. 2, with network sizes of [10, 50, 10] and [10, 200, 10], respectively, for the input, hidden, and output layers. The units in the hidden layer use the rectified linear unit (ReLU) nonlinear activation function and the neurons in the output layer use the hyperbolic tangent activation function. The network is connected in a feedforward way where a neuron in one layer connects to every neuron in the next layer. Since the ANN serves as a switch state controller, we use a training scheme, similar to one-class classification [45], where the training data are the only data that the network learns to produce. In so doing, the network learns to overfit the training data, and any input outside of the designated state-encoding is discarded. The network is, therefore, not expected to deviate from the learned pattern. The network learns both control regimes with 100% accuracy after training with 3,000 randomly-generated 10-bit inputs.

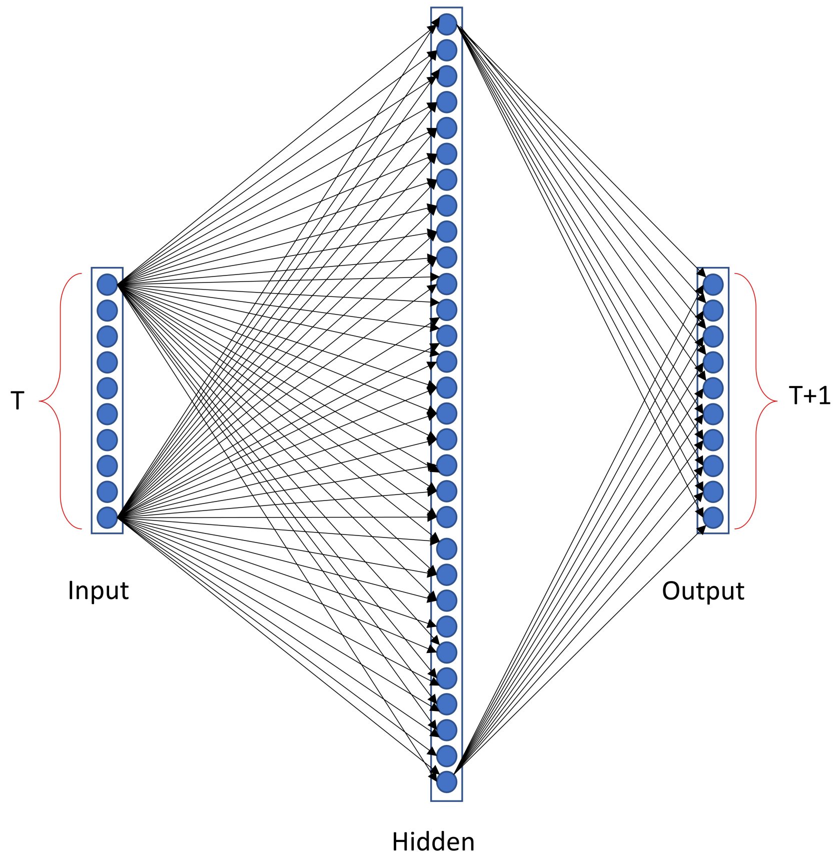


Figure 2: Feed-forward network archtecture used to learn the control cycles specified in Table 1. Each node is connected to every node of the next layer, as shown here for the first and last nodes only. The labels 'T' and 'T+1' indicate time steps in the executed control flow.

In the more realistic case of noisy input data, where binary states can be flipped, the Bidirectional Associative Memory (BAM), a minimal two-layer nonlinear feedback network [46], is a viable alternative to a shallow feed-forward ANN. The architecure is shown in Fig. 3. This BAM network learns to associate between the two initial and final states in Table 1, with similar performance to that of the feed-forward network.

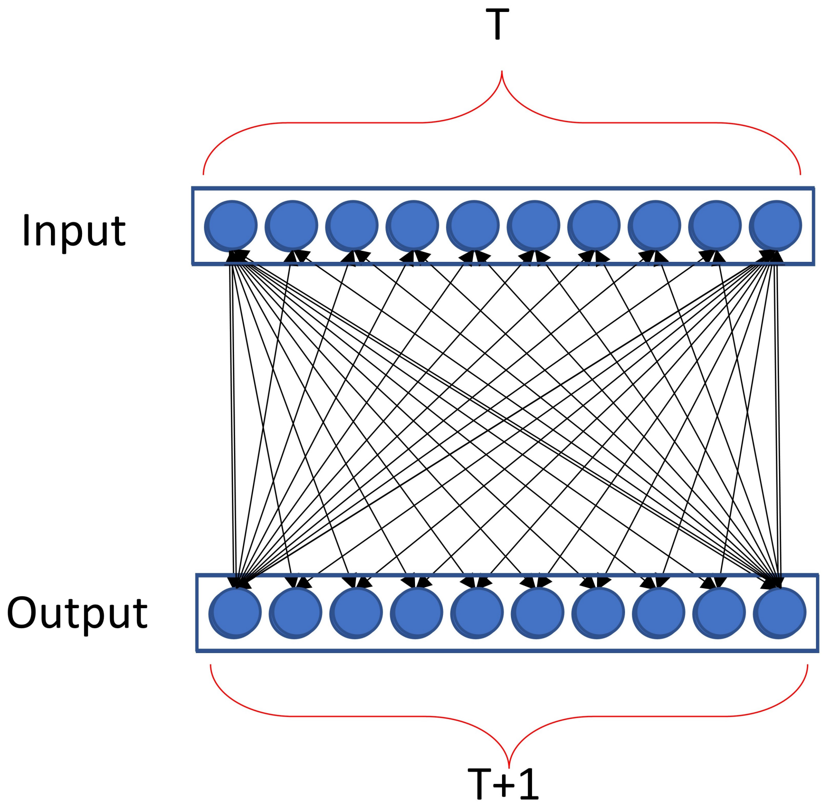


Figure 3: Architecture of the Bidirectional Associative Memory (BAM) network employed here. As in Fig. 2, only the connections of the first and last nodes are shown explicitly.

3 Implementing control flow with TQNNs

Tensor Networks can be naturally associated to the matrix elements of physical scalar products among topological quantum neural networks (TQNNs). Physical scalar products encode indeed the dynamics of TQFTs, since they fulfill their constraints of imposing flatness of the curvature and gauge invariance. Thus, the matrix elements associated to scalar products can be seen as evolution matrix elements for the spin-network states that span the Hilbert spaces of TQNNs.

3.1 Tensor networks as classifiers for TQNNs

A notable example is provided by BF theories [47], a class of TQFTs particularly well studied in the literature of mathematical physics that enables expressing effective theories of particle physics, gravity and condensed matter, and provides as well a general framework for implementations of models of quantum information and quantum computation, machine learning (ML), and neuroscience. These are defined on the principal bundle M of a connection A for some internal gauge group G, with algebra $\mathfrak{g}$, according to the action on a d-dimensional manifold $\mathcal{M}_d$:

$$\mathcal{S} = \int_{\mathcal{M}_d} \mathrm{Tr}[B \wedge F]\,, \tag{3}$$

where B is an ad($\mathfrak{g}$)-valued d-2-form, F denotes the field-strength of A, which is a 2-form, and the trace Tr is over the internal indices of $\mathfrak{g}$, ensuring gauge invariance of the density Lagrangian $\mathcal{L} = \mathrm{Tr}[B \wedge F]$ of the BF theory.

Variation with respect to the conjugated variables, the connection A and the B frame-field, closing a canonical symplectic structure, provide the equations of motion of the theory [47]:

$$F = 0\,, \qquad d_A B = 0\,, \tag{4}$$

which are, respectively, the curvature constraint, imposing the flatness of the connection, and the Gauß constraint, imposing invariance under gauge transformations, having denoted with d_A the covariant derivative with respect to the connection A.

At the quantum level, the states of the kinematical Hilbert space of the theory, fulfilling by construction the Gauß constraint, can be represented in terms of cylindrical functionals Cyl, supported on graphs Γ that are unions of segments γ_i, the end points of which meet in nodes n, and with holonomies – elements of the group G – $H_{\gamma_i}[A]$ of the connection A

assigned to γ_i and intertwiner operators – invariant tensor products of representations – v_n assigned to the nodes n.

For $G = \mathrm{SU}(2)$, spin-networks $|\Gamma, j_\gamma, \iota_n\rangle$, supported on Γ and labelled by the spin j_γ of the irreducible representations of the group elements assigned to γ and by the quantum intertwiner numbers ι_n associated to v_n, represent a basis of the kinematical Hilbert space of the theory. In terms of functionals of Cyl, one can provide the holonomy representation, which is related to the "spin and intertwiner" representation of $|\Gamma, j_\gamma, \iota_n\rangle$ by means of the Peter-Weyl transform. This allows us to decompose the spin-network cylindric functional as [48]:

$$\Psi_{j_{\gamma_{ij}}, \iota_{n_i}}(h_{\gamma_{ij}}) = \left(\bigotimes_n \iota_n \right) \cdot \left(\bigotimes_{\gamma_{ij}} D^{(j_{\gamma_{ij}})}(h_{\gamma_{ij}}) \right) , \tag{5}$$

with $D^{(j)}$ are Wigner matrices providing representation matrices of the SU(2) group elements.

The functorial evolution among spin-networks is ensured by the projector operator [11], which implements the curvature constraint in the physical scalar product among states, i.e.

$$\langle \mathrm{in} | P \, | \mathrm{out} \rangle \, , \qquad \text{with} \qquad P = \int \mathcal{D}[N] \exp(\imath \int \mathrm{Tr}[NF]) \, . \tag{6}$$

We may then regard $|\mathrm{in}\rangle$ as elements of the Hilbert space, and without loss of generality pick up those ones resulting from composing tensorially in Cyl k-representations of holonomies. We may further denote them as $|j_1 \dots j_k\rangle$, with some ordering prescription to associate the topological structure of Γ to the sequence of spin labels. Physically evolving states $P|\mathrm{in}\rangle$ are distinguished from the former ones by labelling them as $|\widetilde{j_1 \cdots j_k}\rangle$. Similarly, we introduce $|\mathrm{out}\rangle$ as the tensor product of (n-k)-representations of holonomies, and denote these states as $|i_1 \dots i_{n-k}\rangle$. Then the matrix elements of $\langle \mathrm{in} | P \, | \mathrm{out} \rangle$ naturally give rise [27] to an n-tensor, i.e.

$$\langle i_1 \dots i_{n-k} | \widetilde{j_1 \cdots j_k} \rangle = T_{i_1 \dots i_{n-k} j_1 \dots j_k} \,. \tag{7}$$

3.2 Geometric RG flow for TQNNs and TNs

The mathematical structures of TQNNs we summarized in Sec. 3.1 are picturing systems "at equilibrium", for which TQFTs characterize a topological stability that percolates into the related transition amplitudes. Nonetheless, it is worth considering as well how stochastic noise might interfere with the topological order ensured by TQFTs, and study the role of "out-of-equilibrium" physics in the analysis of the evolution of the systems under scrutiny.

Out-of-equilibrium dynamics is instantiated considering a heat-flow evolution of the fundamental fields of the theory, with respect to a thermal time τ. Typical Langevin equations, complemented with stochastic noise, provide through their convergence toward the equations of motion of the theory the relaxation toward equilibrium of the field configurations representing specific systems [49]. In general, given some fields ϕ_σ, with a classical equation of motion derived, according to the variational principle $\delta\mathcal{S}/\delta\phi_\sigma$, from an action $\mathcal{S}$ over a Euclidean manifold $\mathcal{M}$, the associated Langevin equations read:

$$\frac{\partial}{\partial\tau}\phi_\sigma = -\frac{\delta\mathcal{S}}{\delta\phi_\sigma} + \eta_\sigma \,, \tag{8}$$

with η_σ a stochastic noise term. The theory at equilibrium is characterized by the symmetries of the equations of motion $\delta\mathcal{S}/\delta\phi_\sigma = 0$ that are broken in the transient phase [50]; these symmetries are consistent with – and in the case of BF theories, actually generated by – the theories at equilibrium.

A prototype of geometric heat-flow was introduced by Hamilton, and then used by Perelman to prove the Poincaré conjecture, which goes under the name of Ricci flow. Here the gravitational field $g_{\mu\nu}$ is the basic configurational space field, while the drift terms are the

Einstein equations of motion in the vacuum, which indeed are expressed by requiring that the components of the Ricci tensor vanish, i.e. $R_{\mu\nu} = 0$. The Ricci flow then reads

$$\imath \frac{\partial}{\partial \tau} g_{\mu\nu} = -2 R_{\mu\nu} \,, \tag{9}$$

having considered now a Lorentzian manifold $\mathcal{M}$. The Ricci flow equations can be further complemented introducing the Ricci target $R^T_{\mu\nu} = \kappa^2 (T_{\mu\nu} - 1/2 g_{\mu\nu} T)$, expressed in terms of the Newton constant $G = \kappa^2/(8\pi)$ and the energy-momentum tensor of matter $T_{\mu\nu}$, so as to obtain at equilibrium the Einstein equations:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \kappa^2 T_{\mu\nu} \,, \qquad \text{or equivalently} \qquad R_{\mu\nu} = R^T_{\mu\nu} \,. \tag{10}$$

The stochastic version of the Ricci flow, with heat equation turning into a Langevin equation, has been introduced and deepened in [50] for a generic gravitational system in the presence of matter fields, describing an action $\mathcal{S}$ for gravity and matter. Moving then from:

$$\imath \frac{\partial}{\partial \tau} g_{\mu\nu} = -\frac{1}{\kappa^2} \frac{\delta \mathcal{S}}{\delta g^{\mu\nu}} + \eta \, g_{\mu\nu} \,, \tag{11}$$

in which a multiplicative noise $\eta_{\mu\nu} = \eta \, g_{\mu\nu}$ appears, the Hamiltonian analysis of the stochastic Ricci flow (SRF) in the Adomian decomposition method (ADM) variables has been derived [50].

An essential by-product of the discussion, from the Ricci flow perspective, is that the equilibration trajectories correspond to those of a renormalization group (RG) flow. The thermal time τ plays the role of scale parameter that individuates a dimension in the bulk, which is out-of-equilibrium. The boundaries are recovered asymptotically in τ, in the infrared regime, and are by definition at equilibrium and thus symmetric.

For a particular class of TQFTs, the BF theories we have introduced in Sec. 3.1 for implementing TQNNs and TNs, the geometric RG flow acquire a specific expression as the TQFT equivalent of the gravitational Ricci flow [51].

3.3 TNs as a generalization of the main model architectures in ML

The use of TNs is an emerging topic in the ML community. The integration between the two appears quite immediate. A TN structure can be viewed as a ML model in which the parameters are properly adjusted to learn the classification of a data set. Yet, as Ref. [52] mentions, machine learning can aid, in turn, in determining a factorization of a TN approximating a data set. Moreover, TNs are also used to compress the layers of ANN architectures, besides a variety of other uses. Tensor networks are becoming more and more popular to the extent that they are a powerful tool for representing and manipulating high-dimensional data, as in the case of image and video classification tasks in which the data are represented as a high-dimensional tensor. High efficiency, flexibility, and ease of use are making them a dominant choice for many AI applications. Furthermore, besides being used to represent data, TNs can be used to process data by exploiting a number of operators. This feature makes them an effective technique for processing data in ML applications.

As is well known, TNs are particularly well suited for representing quantum many-body states in which the dimension of the Hilbert space is exponentially large in the number of particles. The corresponding ML approach consists in:

- Lifting data to exponentially higher spaces;
- Applying any linear classifier $f(x) = W^{\star}\Phi(X)$ to a non-linear space;
- Compressing the weights by using TNs.

The output of the model is a separation of classes that would not be linearly separable in a linear space. In particular, the decision function is the overlap of the weight tensor W with the feature map tensor Φ as in Fig. 4. The weight tensor W can be approximated by the decomposition in Fig. 5.

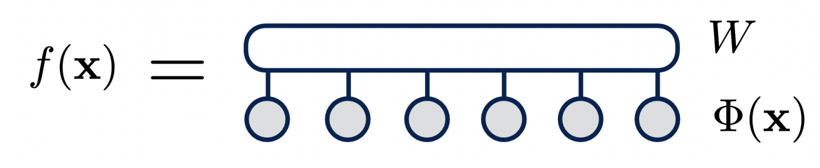


Figure 4: Representation of the decision function (see [53]).

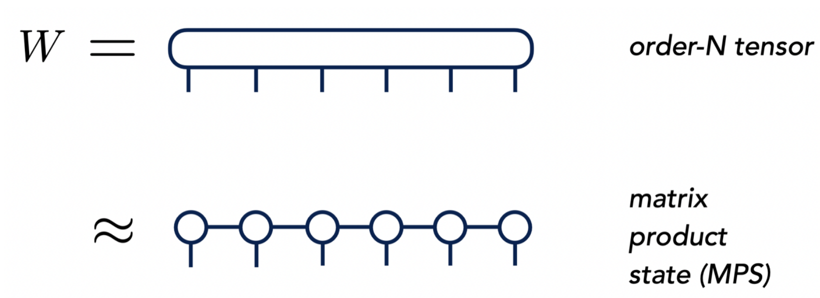


Figure 5: Matrix product decomposition (again see [53]).

Regularization and optimization are built as a constructive product of low-order tensors while weight compression is performed by using the Matrix Product States (MPS) decomposition. If we look at Deep Neural Networks as a piecewise composition of linear discriminators (logistic regression functions), then the TN framework appears as a generalization of the main model architectures found in the ML literature, e.g. Support Vector Machines, Kernel models, and Deep Neural Networks.

The literature concerning the use of tensor theory in traditional ML is becoming large. A short review starts with a seminal paper by Stoudenmire and Schwab [54], which demonstrated how algorithms for optimizing TNs can be adapted to supervised learning tasks by using MPS (tensor trains) to parametrize non-linear kernel learning models. Novikov, Trofimov, and Oseledets [55] have shown how an exponentially large tensor of parameters can be represented in a factorized format called Tensor Train (TT), with the consequence of obtaining a regularization of the model. van Glasser, Pancotti, and Cirac [56] explored the connection between TNs and probabilistic graphical models by introducing the concept of a "generalized tensor network architecture" for ML. Ref. [57] then designed a generative model, i.e. a traditional machine learning model that learns joint probability distributions from data and generates samples according to it, by using MPS. Ref. [58] made use of autoregressive MPSs for building an unsupervised learning model that goes beyond proof-of-concept by showing performance comparable to standard traditional models. Finally, Ref. [59] analyzes the contribution of polynomials of different degrees to the supervised learning performance of different architectures.

4 Implications for biological control systems

Scale-free biology requires a smooth transition from quantum-like to classical-like behavior. Typical representations of metabolic, signal-transduction, and gene-regulatory pathways are entirely classical, even though many of their steps involve electron-transfer or other mechanisms that are acknowledged to require a quantum-theoretic description [60, 61]. As noted earlier, free-energy budget considerations suggest that both prokaryotic and eukaryotic cells employ quantum coherence as a computational resource [62]. Emerging empirical evidence for longer-range entanglement in mammalian brains suggests that large-scale networks may also be using quantum coherence as a resource [63]. Control flow models must, therefore, support the possibility of quantum computation in biological systems. Hierarchical TNs that include unitary components, e.g., MERA-type models, provide this capability.

In prokaryotes, the primary tasks of control flow are adapting metabolism to available resources via metabolite-driven gene regulation [64] and initiating DNA replication and cell division when conditions are favorable. We can, therefore, expect shallow hierarchies of effectively classical control transitions in these organisms. Eukaryotes, however, are characterized by both intracellular compartmentalization and morphological degrees of freedom at the whole-cell scale. We have shown previously that the FEP will induce "neuromorphic" morphologies – i.e. morphologies that segregate inputs from outputs and enable a fan-in/fan-out computational architecture – in any systems with morphological degrees of freedom [65]. Such systems can be expected to have deep control hierarchies at the cellular level, with hierarchical structure correlating with morphological structure in morphologically-complex cells such as neurons [66], and in multicellular assemblages at all scales. These distinctions correlate with the orders-of-magnitude increase in classical computational power (estimated from total metabolic energy budget) as a function of cell-surface area in eukaryotes as compared to prokaryotes [62], as illustrated in Fig. 6.

As well as managing metabolism and replication, most eukaryotes implement active exploration of the environment, communication with other systems, and – crucially for cognition – the writing and reading of stigmergic memories. Thus we can expect such systems to implement QRFs for spacetime and for specific kinds of objects, e.g., conspecifics and suitable substrates for recording stigmergic memories. Such QRFs rely on symmetries, and hence on redundancy of encoded (or encodable) information; they depend, in other words, on the availability of error-correcting codes [67, 68]. The implementation of spacetime as a QECC by TNs has been extensively studied by physicists as noted above; see [69] for review and [27] for a detailed analysis using the present formalism. The use of spacetime as an error-correcting code by organisms – e.g., the implementation of translational and rotational invariance of objects by dorsal visual processing in mammals [70, 71] – is well-understood phenomenologically, but the details of neural implementation remain to be elucidated.

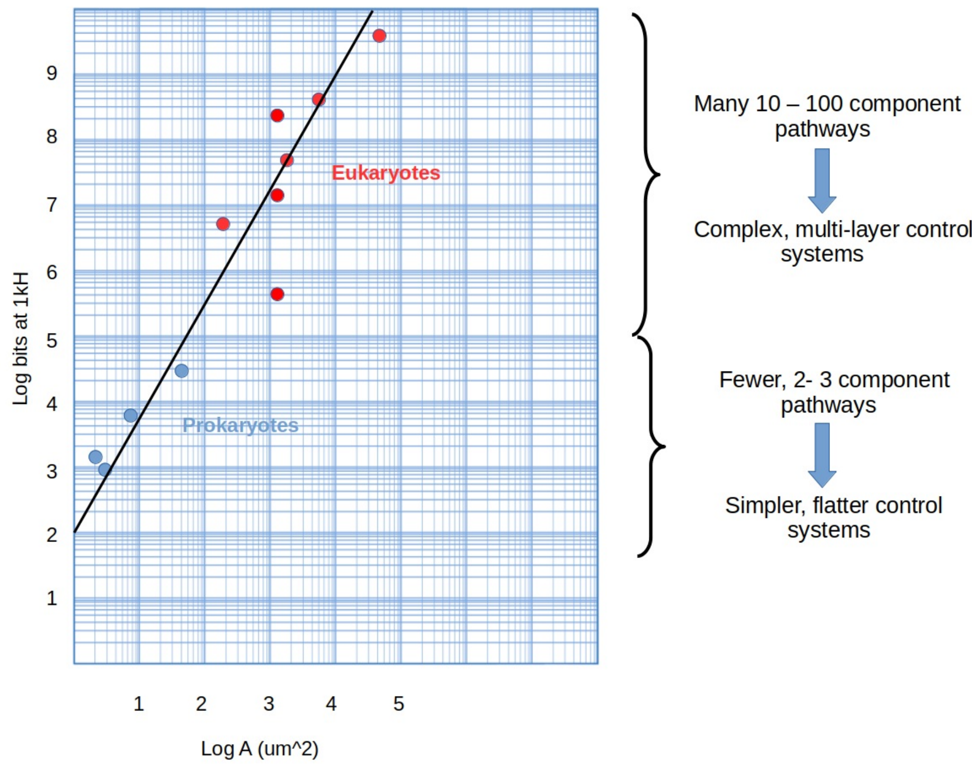


Figure 6: Power-law relation between maximum classical computation rate (vertical axis) and cell-surface area (horizontal axis) derived in [62]. Information processing in eukaryotes is implemented by complex , overlapping signaling pathways that require hierarchical control, which information processing in prokaryotes is implemented by comparatively simple, two or three component pathways that require only shallow control systems. Adapted with permission from [62] Fig. 3.

Both the context-sensitivity of, and the occurence of context effects due to non-commutativity of QRFs in, control networks can be expected to increase with their complexity and hierarchical depth. "Bowtie" networks with high fan-in/fan-out to/from multi-use proteins or second messengers such as Ca^{2+} are increasingly recognized as ubiquitous in higher eukaryotic cells [72]. Such networks have the general form of the CCCD depicted in Part I, Diagram 3. Frequently, such networks evolve via compression of information (e.g. toward share second messengers, as in $[Ca^{2+}]$-based interactions [73, 74]) as an efficiency-increasing mechanism. Bowties introduce semantic ambiguities that must be resolved by context. Each incoming signal has its own governing semantics, but the relevant context can depend on boundary conditions which can be exceedingly difficult (if not impossible) to predetermine (see e.g., [75, 76] for general discussions of the history and semantic depth of this problem). As pointed out in [77], a context change $x \mapsto y$ is semantically problematic if for a fixed set $\{o_i\}$ of observations, the conditional probability distributions $\mathrm{P(o_i|x)}$ and $\mathrm{P(o_i|y)}$ are well defined, but the joint distribution $\mathrm{P(o_i|x \vee y)}$ is not [78]. This occurs whenever the QRFs for x and y do not commute [79, Th 7.1]. As suggested by Part I, Diagram 3, this context-switching problem affects deep learning using VAEs [80]; see e.g., the application to antimicrobial peptides in [81]. In general, the structure of Part I, Diagram 3 can serve as a convenient benchmark for distinguishing signal transduction networks that incorporate co-deployable versus non-co-deployable QRFs [79].

"Quantum" context effects due to non-commutativity have, interestingly, been reported even at the scale of human language use. The "Snow Queen" experiment [82] challenged subjects with distinct, mutually-inconsistent meanings of terms such as 'kind', 'evil', or 'beautiful' in different contexts, and detected statistically-significant context effects using the CbD formalism [83, 84]. Such effects cannot be explained by linguistic ambiguity, misreading, etc. Such language-driven contextuality is taken up in the setting of psycholinguistics and distributional semantics in [85], which combines CbD and the sheaf theoretic [86, 87] methods to systematically study semantic ambiguity as creating meaning/sense discrepancies in statements like "It was about time", "She had time on her hands to win the heat", "West led with a queen", etc.

While the notion of "languages" has thus far been applied to cells, tissues, and even non-vertebrates in a mostly metaphorical way, we can speculate that linguistic approaches to understanding the interplay between context dependence and semantic ambiguity may be useful to biology in general. Immune cells (e.g., T cells) are, for example, "programmed" or "trained" by their progenitor cells to respond to local cellular signals and ambient conditions in particular ways. Unexpected context changes may induce dysfunctional (at the organism scale) responses, including chronic disorders [88]; these can be considered consequenes of discrepancies between the "actual" semantics of incoming signals and the semantics expected by the immune system's "language." This suggestion of possible "linguistic" contextuality seems in consonance with the hypothesis of [89] that the immune system is a cognitive (living) system implementing its exclusive system of language-grammar, which may be prone to analogous disorders of communication as those discussed in [85]. Similar context effects have been observed in microbiological systems [90]; here discrepancies in experimentally derived classical probabilities arising from lactose-glucose interference signaling in *E. Coli* can only be explained in terms of non-classical probabilities. We note that the expression 'quantum-like' [91] is often used for such effects; however, their formal structure is exactly that given by quantum theory.

We expect that further research into quantum biology will unfold significant perspectives on human/mammalian physiology and cognitive capabilities along the lines suggested in the present article. For example, allostatic maintenance, as briefly alluded to in Part I, can be seen as a process regulating a body's physiological conditions relative to costs and benefits while dynamically allocating resources for the purpose of overall adaptability of an organism within its internal environment. Implementing the allostatic and anticipatory mechanisms are the visceromotor cortical regions generating autonomic, hormonal, and immunological predictions leading to interoceptive inference [92, 93, 94, 96, 97, 98, 102, 99]. This process of inference in humans and mammals putatively utilizes predictive coding for the processes of homeostasis-allostasis through a hierarchy of cellular to organ-level systems, in turn connecting interoception to the processes of extercoception and proprioception [92, 99, 100, 101, 102]. The basic principles follow from how allostasis provides protection against potential surprise by utilizing a framework somewhat beyond the error signaling necessary for homeostatic maintenance (it is essentially through minimizing the free energy of internal state trajectories towards combatting surprise, as discussed in Part I). The net effect of the process is consonant with the Good Regulator theorem of [95], showing how regulation of a given system requires an internal model of that system. A further perspective is to emphasize the predictive nature of an integrated, complex, allostatic-interoceptive cortical system capable of supporting a spectrum of psychological phenomena including memory and emotions [99] (cf. [102]). Accordingly, cognitive conditions such as depression and autism have been described as abnormalities of allostatic-interoceptive inference, so impairing predictive coding mechanisms due to aberrant assimilation and mistuning of prediction errors (putatively a connectivity issue), conceivably leading to a root cause of many known cognitive conditions [92, 100, 102].

We anticipate that this fully general, context sensitive model of control flow will be important for understanding morphogenesis, which is not simply a feed-forward emergent system, but rather a highly context-sensitive error-minimizing process [103]. Specifically, the collective intelligence of cells during embryonic development, organ regeneration, and metamorphosis can create and repair specific complex structures despite a wide range of perturbations [104]. Changes in the genome, the number of cells, or the starting configuration can often be overcome: bisected embryos result in normal twins, amputated salamander limbs re-grow back to normal, and planarian fragments result in perfect little worms [105]. The competency of cellular collectives to reach the correct target morphology despite even drastic interventions requires an understanding of how they navigate, via context-sensitive control flow, problem spaces including anatomical morphospace [106], physiological, and transcriptional spaces [68, 107]. Understanding the navigation policies used by unconventional collective intelligences can help not only understand creative problem-solving on rapid timescales (such as the ability to regulate genes to accommodate an entirely novel stressor [108] without evolutionary adaptation), but may also have implications for predicting and managing the goals and behavioral repertoires of synthetic beings [109].

5 Conclusion

We have shown here how the problem of defining control flow arises in active inference systems, and provided three formal representations of the problem. We have proved that control flow in such systems can always be represented by a tensor network, provided illustrative examples, and shown how the general formalism of topological quantum neural networks can be used to implement a general model of control flow. These results provide a general formalism with which to characterize context dependence in active inference systems at any scale, from that of macromolecular pathways to that of multi-organism communities. They suggest that the concept of communication by language is not just metaphorical when applied to biological systems in general, but rather an appropriate and productive description of interactional dynamics.

We view these results as a further step toward fully integrating the formal models, concepts, and languages of physics, biology, and cognitive science. This integration is not reductive. It rather allows us to classify systems using natural measures of organizational and computational complexity, and to understand how interactions between simpler systems can implement the more complex behavior of the larger systems that they compose.

Acknowledgements

K.F. is supported by funding for the Wellcome Centre for Human Neuroimaging (Ref: 205103/Z/16/Z), a Canada-UK Artificial Intelligence Initiative (Ref: ES/T01279X/1) and the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement No. 945539 (Human Brain Project SGA3). M.L. gratefully acknowledges funding from the Guy Foundation and the John Templeton Foundation, Grant 62230. A.M. wishes to acknowledge support by the Shanghai Municipality, through the grant No. KBH1512299, by Fudan University, through the grant No. JJH1512105, the Natural Science Foundation of China, through the grant No. 11875113, and by the Department of Physics at Fudan University, through the grant No. IDH1512092/001.

Conflict of interest

The authors declare no competing, financial, or commercial interests in this research.

References

[1] Friston, K. J. 2010 The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience 11*, 127–138.

[2] Friston, K. J. 2013 Life as we know it. *Journal of The Royal Society Interface 10*, 20130475.

[3] Friston KJ, FitzGerald T, Rigoli F, Schwartenbeck P, Pezzulo G. Active inference: a process theory. *Neural Comput* 2017;29:1–49.

[4] Friston, K. J. 2019 A free energy principle for a particular physics. Preprint arxiv:1906.10184 [q-bio.NC]. https://arxiv.org/abs/1906.10184

[5] Ramstead MJ, Sakthivadivel DAR, Heins C, Koudahl M, Millidge B, Da Costa L, Klein B, Friston KJ 2022 On Bayesian mechanics: A physics of and by beliefs. *Interface Focus* 13, 2022.0029.

[6] Fields C, Friston K, Glazebrook JF, Levin M 2022 A free energy principle for generic quantum systems. *Prog. Biophys. Mol. Biol.* 173, 36–59.

[7] Friston, K., Da Costa, L., Sakthivadivel, D. A. R., Heins, C., Pavliotis, G. A., Ramstead, M., Parr, T. 2022 Path integrals, particular kinds, and strange things. Preprint arxiv:2210.12761.

[8] Aharonov Y, Kaufherr T. Quantum frames of reference. *Phys Rev D* 1984;30:368–385.

[9] Bartlett SD, Rudolph T, Spekkens RW. 2007 Reference frames, super-selection rules, and quantum information. *Rev Mod Phys* 79, 555–609.

[10] Atiyah, M. 1988 Topological quantum field theory. *Pub. Math. IHÈS* 68, 175–186.

[11] Fields, C.; Glazebrook, J. F.; Marcianò, A. (2022) Sequential measurements, topological quantum field theories, and topological quantum neural networks. *Fortschr. Phys.* 2022, 2200104.

[12] Marcianò, A.; Chen, D.; Fabrocini, F.; Fields, C.; Greco, E.; Gresnigt, N.; Jinklub, K.; Lulli, M., Terzidis, K.; Zappala, E. 2022 Quantum neural networks and topological quantum field theories. *Neural Networks* 153, 164–178.

[13] Marcianò, A., Chen, D., Fabrocini, F., Fields, C., Lulli, M., Zappala, E. 2022 Deep neural networks as the semi-classical Limit of topological quantum neural networks: The problem of generalisation. Preprint arXiv:2210.13741.

[14] Qi, X.-L. Does gravity come from quantum information? *Nature Phys.* 14, 984-987 (2018), https://doi.org/10.1038/s41567-018-0297-3.

[15] Bekenstein, J. D. Black holes and the second law. *Lett. Nuovo Cim.* 4, 737-740 (1972).

[16] Bekenstein, J. D. Black holes and entropy. *Phys. Rev. D* 7, 2333-2346 (1973).

[17] Hawking, S. W. Gravitational radiation from colliding black holes. *Phys. Rev. Lett.* 26, 1344-1346 (1971).

[18] Hawking, S. W. Black hole explosions? *Nature* 248, 30-31 (1974).

[19] Susskind, L. The world as a hologram. *J. Math. Phys.* 36, 6377-6396 (1995).

[20] Bousso, R. The holographic principle. *Rev. Mod. Phys.* 74, 825-874 (2002).

[21] 't Hooft, G. Dimensional reduction in quantum gravity. Preprint at https://arxiv.org/abs/gr-qc/9310026 (1993).

[22] Maldacena, J. The large-N limit of superconformal field theories and supergravity. *Int. J. Theor. Phys.* 38, 1113 (1999).

[23] Ryu, S. and Takayanagi, T. Holographic derivation of entanglement entropy from AdS/CFT, *Phys. Rev. Lett.* 96, 181602 (2006).

[24] Almheiri, A., Dong, X. and Harlow, D. Bulk locality and quantum error correction in AdS/CFT. *J. High Energy Phys.* 2015, 163 (2015).

[25] Hubeny, V. E. and Rangamani, M. Causal holographic information. *J. High Energy Phys.* 2012, 114 (2012).

[26] Headrick, M., Hubeny, V. E., Lawrence, A. and Rangamani, M. Causality and holographic entanglement entropy. *J. High Energy Phys.* 2014, 162 (2014).

[27] Fields, C.; Glazebrook, J. F.; Marcianò, A. 2023 Communication protocols and quantum error-correcting codes from the perspective of topological quantum field theory. Preprint arxiv:2303.16461 [hep-th].

[28] Swingle, B. Entanglement renormalization and holography. *Phys. Rev. D* 86, 065007 (2012).

[29] White, S. R. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.* 69, 2863-2866 (1992).

[30] Verstraete, F., Murg, V. and Cirac, J. I. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Adv. Phys.* 57, 143-224 (2008).

[31] DiVincenzo, D. P. et al. in *Quantum Computing and Quantum Communications* (ed. Williams, C. P.) 247-257 (Springer, Berlin, 1999).

[32] Vidal, G. Class of quantum many-body states that can be efficiently simulated. *Phys. Rev. Lett.* 101, 110501 (2008).

[33] Pastawski, F., Yoshida, B., Harlow, D. and Preskill, J. Holographic quantum error-correcting codes: toy models for the bulk/boundary correspondence. *J. High Energy Phys.* 2015, 149 (2015).

[34] Yang, Z., Hayden, P. and Qi, X.-L. Bidirectional holographic codes and sub-AdS locality. *J. High Energy Phys.* 2016, 175 (2016).

[35] Hayden, P. et al. Holographic duality from random tensor networks. *J. High Energy Phys.* 2016, 9 (2016).

[36] Han, M. and Huang, S. Discrete gravity on random tensor network and holographic Rényi entropy. *J. High Energy Phys.* 2017, 148 (2017).

[37] Lashkari, N., McDermott, M. B. and Van Raamsdonk, M. Gravitational dynamics from entanglement "thermodynamics". *J. High Energy Phys.* 2014, 195 (2014).

[38] Swingle, B. and Van Raamsdonk, M. Universality of gravity from entanglement. Preprint at https://arxiv.org/abs/1405.2933 (2014).

[39] Shenker, S. H. and Stanford, D. Black holes and the butterfly effect. *J. High Energy Phys.* 2014, 67 (2014).

[40] Sachdev, S. and Ye, J. Gapless spin-fluid ground state in a random quantum Heisenberg magnet. *Phys. Rev. Lett.* 70, 3339-3342 (1993).

[41] Kitaev, A. A simple model of quantum holography. KITP http://online.kitp.ucsb.edu/online/entangled15/kitaev/; http://online.kitp.ucsb.edu/online/ entangled15/kitaev2/ (2015).

[42] Maldacena, J. and Stanford, D. Remarks on the Sachdev-Ye-Kitaev model. *Phys. Rev. D* 94, 106002 (2016).

[43] Smith, J. E., Nair, R. 2005 The architecture of virtual machines. *IEEE Computer* 38(5), 32–38.

[44] Orús, R. 2019 Tensor networks for complex quantum systems. *Nat. Rev. Phys.* 1, 538–550.

[45] Manevitz, L. M., Yousef, M. 2002 One-class SVMs for document classification. *J. Mach. Learn. Res.* 2, 139–154.

[46] Kosko, B. 1988 Bidirectional associative memories. *IEEE Trans. Syst. Man Cybern.* 18, 49–60.

[47] J. Baez, Four-dimensional BF theory as a topological quantum field theory *Lett. Math. Phys.* 1996, 38, 129.

[48] Rovelli, C. *Quantum Gravity*, Cambridge University Press, Cam- bridge, UK 2004.

[49] Parisi, G. and Wu, Y.-S. Perturbation theory without gauge fixing, *Scientia Sinica* 24, 483 (1981); doi: 10.1360/ya1981-24-4-483, http://engine.scichina.com/doi/10.1360/ya1981-24-4-483.

[50] Lulli, M., Marcianò, A. and Shan, X. Stochastic Quantization of General Relativity à la Ricci-Flow, [arXiv:2112.01490 [gr-qc]].

[51] Marcianò, A. *in preparation*.

[52] Sengupta, R. Adhikary, S. Oseledets, I. and Biamonte, J. Tensor Networks in Machine Learning, https://arxiv.org/pdf/2207.02851.pdf, 2022.

[53] E. Miles Stoudenmire, Talk at the International Center for Theoretical Physics Trieste, August 2018.

[54] Stoudenmire, E.M. Schwab, D.J. Supervised Learning with Quantum-Inspired Tensor Networks, https://arxiv.org/abs/1605.05775, 2016.

[55] Novikof, A. Trofimov, M. and Oseledets, I. Exponential Machines, https://arxiv.org/abs/1605.03795, 2017.

[56] Glasser, I. Pancotti, N. and Cirac, J.I. From probabilistic graphical models to generalized tensor networks for supervised learning, https://arxiv.org/abs/1806.05964, 2018.

[57] Han, Z.-H. Wang, J. Fan, H. Wang, L. and Zhang, P. Unsupervised Generative Modeling Using Matrix Product States, *Physical Review X*, 8, 2018.

[58] Liu, J. Li, S.-J. Zhang, J. and Zhang, P. Tensor networks for unsupervised machine learning, https://arxiv.org/abs/2106.12974, 2021.

[59] Convy, I. and Whaley, K.B. Interaction Decompositions for Tensor Network Regression, https://arxiv.org/abs/2208.06029, 2022.

[60] Zweir MC, Chong LT. 2010 Reaching biological timescales with all-atom molecular dynamics simulations. *Curr. Opin. Pharmacol.* 10: 745–752.

[61] Groenhof, G. 2013 Introduction to QM/MM simulations. *Methods Mol. Biol.* 924, 43–66.

[62] Fields, C.; Levin, M. 2021 Metabolic limits on classical information processing by biological cells. *BioSystems 209*, 104513.

[63] Kerskens, C. M., Pérez, D. L. 2022 Experimental indications of non-classical brain functions. *J. Phys. Commun.* 6, 105001.

[64] Ledezma-Tejeida, D.; Schastnaya, E.; Sauer, U. 2021 Metabolism as a signal generator in bacteria. *Curr. Opin. Syst. Biol.* 28, 100404.

[65] Fields, C.; Friston, K.; Glazebrook, J. F.; Levin, M.; Marcianò, A. 2022 The free energy principle induces neuromorphic development. *Neuromorph. Comp. Engin.* 2, 042002.

[66] Fields, C.; Glazebrook, J. F.; Levin, M. 2022 Neurons as hierarchies of quantum reference frames. *Biosystems* 219, 104714.

[67] Fields, C.; Levin, M. 2018 Multiscale memory and bioelectric error correction in the cytoplasm-cytoskeleton-membrane system. *WIRES Syst. Biol. Med.* 10, e1410.

[68] Fields, C., Levin, M. 2022 Competency in navigating arbitrary spaces as an invariant for analyzing cognition in diverse embodiments. *Entropy* 24, 819.

[69] Bain, J. 2020 Spacetime as a quantum error-correcting code? *Stud. Hist. Phil. Mod. Phys.* 71, 26–36.

[70] Flombaum, J. I.; Scholl, B. J.; Santos, L. R. 2008 Spatiotemporal priority as a fundamental principle of object persistence. In: *The Origins of Object Knowledge*, eds B. Hood and L. Santos (Oxford: Oxford University Press), 135–164.

[71] Fields, C. 2011 Trajectory recognition as the basis for object individuation: A functional model of object file instantiation and object-token encoding. *Front. Psychol.* 2, 49.

[72] Niss, K.; Gomez-Casado, C.; Hjaltelin, J. X.; Joeris, T.; Agace, W. W.; Belling, K. G.; Brunak, S. 2020 Complete topological mapping of a cellular protein interactome reveals bow-tie motifs as ubiquitous connectors of protein complexes. *Cell Rep.* 31, 107763.

[73] Carafoli, E. and Krebs, J. Why calcium? How clacium became the best communicator. *J. Biol Chem* 40 2016, 20849–20857.

[74] Polouliakh, N., Nock, R., Nielsen, F. and Kitano, H. G-protein coupled receptor siganling architecture of mammalian immune cells. *PLoS ONE* 4(1) 2009, e4189.

[75] Friedlander, T., Mayo, A. E., Tlusty, T. and Alon. U. Evolution of bow-tie architectures in biology. *PLOS Computational Biology* 11(3) 2015, e1004055

[76] Boniolo. G., D'Agostino, M., PIazza, M. and Pulcini, G. Molecular biology meets logic: Context-senstivity in focus. *Foundations of Science* 2021 in press, https://doi.org/10.1007/s10699-021-09789-y

[77] Fields, C.; Glazebrook, J. F.; Levin, M. 2021 Minimal physicalism as a scale-free substrate for cognition and consciousness. *Neurosci. Cons. 7(2)*, niab013.

[78] Kochen, S., Specker, E. P. 1967 The problem of hidden variables in quantum mechanics. *J. Math. Mech.* 17, 59–87.

[79] Fields, C.; Glazebrook, J. F. 2022 Information flow in context-dependent hierarchical Bayesian inference. *J. Expt. Theor. Artif. intell. 34*, 111–142.

[80] Kingma, D. P and Welling, M. An Introduction to variational autoencoders. *Foundations and Trends in Machine Learning* 12(4) (2019), 307–392.

[81] Dean, S. N. and Walper, S. A. Variational autoenecoder for generation of antimicrobial peptides. *ACS Omega* 5 2020, 20746–20754.

[82] Cervantes, V. H., Dzhafarov, E. N. (2018). Snow Queen is evil and beautiful: Experimental evidence for probabilistic contextuality in human choices. *Decision* 5(3), 193–204.

[83] Dzhafarov, E. N.; Kujala, J. V. (2017a). Contextuality-by-Default 2.0: Systems with binary random variables. In: J. A. Barros, B. Coecke and E. Pothos (eds.) *Lecture Notes in Computer Science* 10106, Springer, Berlin, 16–32.

[84] Dzharfarov, E. N. & Kon, M. 2018 On universality of classical probability with contextually labeled random varaibles. *J. Math. Psychol.* 85, 17–24.

[85] Wang, D., Sadrzadeh, Abramsky, S. and Cervantes, V. H. On the quantum-like contextulaity of ambiguous phrases. *Proceedings of the 2021 Workshop on Semantic Spaces at the Intersection of NLP, Physics and Cognitive Science*, pp. 42–52. Association for Computational Linguistics 2021.

[86] Abramsky, S., Brandenburger, A. 2011 The sheaf-theoretic structure of non-locality and contextuality. *New J. Phys.* 13, 113036.

[87] Abramsky, S., Barbosa, R. S., Mansfield, S. 2017 Contextual fraction as a measure of contextuality. *Phys. Rev. Lett.* 119, 050504.

[88] Editorial Focus. 2019 A matter of context. *Nature Immunol.* 20, 769.

[89] Atlan, H. and Cohen, I. R. 1998 Immune information, self-organiztion and meaning. *Int. Immunology* 10, 711–717.

[90] Basieva I, Khrennikov A, Ohya M, Yamato O. Quantum-like interference effect in gene expression: glucose-lactose destructive interference. *Syst Synth Biol* 2011;5:59–68.

[91] Khrennikov, A. Quantum-like modeling of cognition. *Front Phys* 3 2015:77.

[92] Barrett, L. F., and Simmons, W. K. (2015) Interoceptive predictions in the brain. *Nat. Rev. Neurosci.* **16**(7): 419–429.

[93] Barrett, L. F., Quigley, K. S. and Hamilton, P. An active inference theory of allostasis and interoception in depression. *Phil. Trans. R. Soc. B* 371: 20160011.

[94] Barrett, L. F. The theory of constructed emotion: an active inference account of interoception and categorization. *Soc. Cogn. Affect. Neurosci.* 12 2017, 1–23.

[95] Conant RC, Ashby WR. Every good regulator of a system must be a model of that system. *Int. J. Syst. Sci.* 1970;1(2):89–97.

[96] Corcoran, A. W., Pezzulo, G., Hohwy, J. 2020 From allostatic agents to counterfactual cognisers: Active inference, biological regulation, and the origins of cognition. *Biol. Philos.* 35(3), 32.

[97] Hohwy, J. *The Predictive Mind*. Oxford University Press, Oxford, UK, 2013

[98] Hohwy, J. 2016 The self-evidencing brain. *Noûs* 50(2), 259–285.

[99] Kleckner, I. R. et al. Evidence for a large-scale brain system supporting allostasis and interoception in humans. *Nature Human Behaviour* 11 2017, Article 0069, 14 pages.

[100] Seth, A. K., Suzuki, K, and Critchley, H. D. An interoceptive predictive coding model of conscious presence. *Frontiers in Psychology* 2 Article 395, 16 pages.

[101] Seth, A. K. (2013) Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Science* **17** (11): 565–573.

[102] Seth, A. K. and Friston, K. J. Active interoceptive inference and the emotional brain.*Phil. Trans. R. Soc. B* 371, 20160007.

[103] Levin, M. 2022 Technological approach to mind everywhere: An experimentally-grounded framework for understanding diverse bodies and minds. *Front. Syst. Neurosci.* 16, 768201.

[104] Pezzulo, G., Levin, M. 2016 Top-down models in biology: explanation and control of complex living systems above the molecular level. *J. R. Soc. Interface* 13(124), 20160555.

[105] Birnbaum, K. D., Sánchez Alvarado, A. 2008 Slicing across kingdoms: Regeneration in plants and animals. *Cell* 132(4), 697–710.

[106] Levin, M. 2022 Collective intelligence of morphogenesis as a teleonomic process. Preprint PsyArXiv hqc9b.

[107] Biswas, S., Clawson, W., Levin, M. 2023 Learning in transcriptional network models: Computational discovery of pathway-level memory and effective interventions. *int. J. Molec. Sci.* 24(1), 285.

[108] Emmons-Bell, M., Durant, F., Tung, A. et al. 2019 Regenerative adaptation to electrochemical perturbation in planaria: A molecular analysis of physiological plasticity. *iScience* 22, 147–165.

[109] Clawson, W., Levin, M. 2022 Endless forms most beautiful 2.0: Teleonomy and the bioengineering of chimaeric and synthetic organisms. *Biol. J. Linnean Soc.* 2022, blac073.