Non-equilibrium dynamics and entropy production in the human brain

Christopher W. Lynn\textsuperscript{1,2}, Eli J. Cornblath\textsuperscript{2,3}, Lia Papadopoulos\textsuperscript{1}, Maxwell A. Bertolero\textsuperscript{3}, & Danielle S. Bassett\textsuperscript{1,2,4,5,6,7,*}

\textsuperscript{1}Department of Physics & Astronomy, College of Arts & Sciences, University of Pennsylvania, Philadelphia, PA 19104, USA
\textsuperscript{2}Department of Bioengineering, School of Engineering & Applied Science, University of Pennsylvania, Philadelphia, PA 19104, USA
\textsuperscript{3}Department of Neuroscience, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA
\textsuperscript{4}Department of Electrical & Systems Engineering, School of Engineering & Applied Science, University of Pennsylvania, Philadelphia, PA 19104, USA
\textsuperscript{5}Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA
\textsuperscript{6}Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA
\textsuperscript{7}Santa Fe Institute, Santa Fe, NM 87501, USA
It has long been suspected that living systems operate out of thermodynamic equilibrium\textsuperscript{1}, consuming energy and producing entropy in the environment in order to perform biological functions. Recent efforts at the microscale have established that non-equilibrium processes are critical for molecular and cellular operations\textsuperscript{2–8}. However, it remains unclear whether non-equilibrium dynamics manifest at macroscopic scales, and if so, how such dynamics support higher-order biological functions. Here we present a framework to probe for non-equilibrium processes by quantifying entropy production in macroscopic systems. We apply our method to the human brain, an organ whose immense metabolic consumption drives a diverse range of cognitive functions\textsuperscript{9,10}. Using whole-brain imaging data, we demonstrate that the brain fundamentally operates out of equilibrium at large scales. Moreover, we find that the brain produces more entropy – operating further from equilibrium – when performing physically and cognitively demanding tasks. To understand how the brain operates at different distances from equilibrium, we use simulations to show that the non-equilibrium nature of macroscopic systems is modulated by the strength of interactions at the microscale. Together, these results suggest that non-equilibrium dynamics are vital for cognition, and provide a general tool for quantifying the non-equilibrium nature of macroscopic systems.
The functions that support life – from processing information to generating forces and maintaining order – require organisms to operate far from thermodynamic equilibrium. For a system at equilibrium, the fluxes of transitions between different states vanish (Fig. 1a), a property known as detailed balance; the system ceases to produce entropy and its dynamics become reversible in time. By contrast, living systems exhibit net fluxes between states or configurations (Fig. 1b), thereby breaking detailed balance and establishing an arrow of time. Critically, such non-equilibrium dynamics lead to the production of entropy, a fact first recognized by Sadi Carnot in his pioneering studies of irreversible processes. At the molecular scale, enzymatic activity drives non-equilibrium processes that are crucial for intracellular transport, high-fidelity transcription, and biochemical patterning. At the level of cells and subcellular structures, non-equilibrium activity enables sensing, adaptation, force generation, and structural organization. However, despite the importance of non-equilibrium processes at small scales, there remain basic questions concerning how – and even whether – non-equilibrium dynamics unfold in macroscopic systems composed of many interacting components. Indeed, the amount of entropy produced by a system can only decrease with coarse-graining, leading to the possibility that complex living systems, despite operating far from equilibrium at small scales, may appear to regain equilibrium at large scales.

Perhaps the most promising system in which to probe for macroscopic non-equilibrium dynamics is the human brain, which consumes up to 20% of the body’s energy to perform an array of cognitive functions, from computations and attention to planning and motor execution. In fact, enzymatic and metabolic activity in the brain drives a number of non-equilibrium processes at the microscale, including molecular cycles, cellular housekeeping, and neuronal firing. But does the brain operate out of equilibrium at large scales? And if so, does the non-equilibrium nature of the brain vary with physical or cognitive demands? The answers to these questions will provide insights into the non-equilibrium basis of neural computations specifically, and will yield a framework for studying macroscopic non-equilibrium processes generally.
**Fig. 1** Macroscopic non-equilibrium dynamics in the brain.

**a-b.** A simple four-state system, with states represented as circles and transition rates as arrows. 
**a.** At equilibrium, there are no net fluxes of transitions between states—a condition known as detailed balance—and the system does not produce entropy. 
**b.** Systems that are out of equilibrium exhibit net fluxes of transitions between states, breaking detailed balance and producing entropy in the environment.

**c.** Brain states defined by the first two principal components of the neuroimaging time-series of regional activity, calculated for all time points and all subjects. Colors indicate the z-scored activation of different brain regions, ranging from high-amplitude activity (green) to low-amplitude activity (orange). Arrows represent possible fluxes between states.

**d-e.** Probability distribution (color) and net fluxes between states (arrows) for neural dynamics at rest (**d**) and during a gambling task (**e**). In order to use the same axes in panels **d** and **e**, the dynamics are projected onto the first two principal components of the combined rest and gambling time-series data. The flux scale is indicated in the upper right, and the disks represent two-standard-deviation confidence intervals for fluxes estimated using trajectory bootstrapping\(^{15}\) (see Methods; Fig. S1).
Fluxes and broken detailed balance in the brain

Here we develop tools to probe for and quantify non-equilibrium dynamics in macroscopic living systems. We apply our methods to analyze whole-brain dynamics from 590 healthy adults both at rest and across a suite of seven cognitive tasks, recorded using functional magnetic resonance imaging (fMRI) as part of the Human Connectome Project\textsuperscript{19}. For each cognitive task (including rest), the time-series data consist of blood-oxygen-level-dependent (BOLD) fMRI signals from 100 cortical parcels\textsuperscript{20} (see Methods), which we concatenate across all subjects. To visualize the neural dynamics, we project the time series onto the first two principal components, which are calculated for all data points and all subjects (Fig. 1c). In fact, this projection defines a natural low-dimensional state space\textsuperscript{21}, capturing over 30\% of the variance in the neural activity (Fig. S2). One can then probe for non-equilibrium dynamics by calculating the net fluxes of transitions between different regions of state space\textsuperscript{22} (see Methods). Moreover, one can repeat this analysis for different tasks to investigate whether the brain’s non-equilibrium behavior depends on the cognitive function being performed.

To probe for non-equilibrium dynamics, we first consider the brain’s behavior during resting scans, wherein subjects are instructed to remain still without executing a specific task. At rest, we find that the brain exhibits net fluxes between states (Fig. 1d), thereby breaking detailed balance and fundamentally departing from equilibrium at large scales. Given the longstanding hypothesis that biological functions arise from non-equilibrium dynamics\textsuperscript{1}, one might expect the brain to operate even further from equilibrium when performing a specific cognitive task. To test this hypothesis, we study task scans, wherein subjects respond to stimuli and commands that require attention, computations, and physical and cognitive effort. For example, during a gambling task in which subjects play a card guessing game for monetary reward, the brain’s dynamics form a distinct loop of fluxes (Fig. 1e) that are nearly an order of magnitude stronger than those present during rest. Such closed loops of flux are a characteristic feature of non-equilibrium steady-state systems\textsuperscript{23}, and we verify that the brain operates in a stochastic steady state (Fig. S3). Furthermore, to confirm that non-equilibrium dynamics encode the arrow of time, we show that if the time series
are shuffled – thereby destroying the temporal order of the system – then the fluxes between states vanish and equilibrium is restored (Fig. S4). Together, these results demonstrate that the brain fundamentally operates out of equilibrium at large scales, and moreover, that the nature of this non-equilibrium behavior depends critically on the cognitive function being performed.

Simulating macroscopic non-equilibrium dynamics

To understand how non-equilibrium dynamics arise at large scales, it is helpful to consider a canonical model of stochastic dynamics in complex systems. In the Ising model, the interactions between spins are typically constrained to be symmetric, yielding simulated dynamics that obey detailed balance and converge to equilibrium. However, if we relax this constraint to allow for asymmetric interactions, then the system diverges from equilibrium, displaying closed loops of flux between spin states at small scales (Fig. 2a). But can these fine-scale violations of detailed balance combine to generate macroscopic non-equilibrium dynamics? To answer this question, we study a system of \( N = 100 \) spins (matching the 100 parcels in our neuroimaging data), with the interaction between each directed pair of spins drawn independently from a zero-mean Gaussian (Fig. 2b). This model is the asymmetric generalization of the Sherrington-Kirkpatrick (SK) model of a spin glass. After simulating the system at three different temperatures, we perform the same procedure that we applied to the neuroimaging data (Fig. 1): projecting the time-series onto the first two principal components of the combined data and calculating net fluxes in this low-dimensional state space. At high temperature, stochastic fluctuations dominate the system, and we only observe weak fluxes between states (Fig. 2c, right). In contrast, as the temperature decreases, the interactions between spins overcome the stochastic fluctuations, giving rise to clear loops of flux (Fig. 2c, middle and left). These loops of flux demonstrate that large-scale non-equilibrium dynamics can emerge from fine-scale asymmetries in the interactions between elements. Moreover, by tuning the strength of interactions, a single system can transition from near equilibrium to far from equilibrium, suggesting that the brain may operate at different “effective” temperatures when performing distinct cognitive functions (Fig. 1d,e).
Fig. 2 | Simulating complex non-equilibrium dynamics using an asymmetric Ising model.

a, Two-spin Ising model with asymmetric interactions (left), where the interaction $J_{\alpha\beta}$ represents the strength of the influence of spin $\beta$ on spin $\alpha$. Simulating the model with synchronous updates, the system exhibits a clear loop of flux between spin states (right). b, Asymmetric version of the Sherrington-Kirkpatrick (SK) model, wherein directed interactions are drawn independently from a zero-mean Gaussian with variance $1/N$, where $N$ is the size of the system. c, For an asymmetric SK model with $N = 100$ spins, we plot the probability distribution (color) and fluxes between states (arrows) for simulated time-series at temperatures $T = 0.1$ (left), $T = 1$ (middle), and $T = 10$ (right). In order to visualize the dynamics, the time series are projected onto the first two principal components of the combined data across all three temperatures. The scale is indicated in flux-per-time-step, and the disks represent two-standard-deviation confidence intervals estimated using trajectory bootstrapping (see Methods).
Quantifying entropy production in macroscopic systems

While fluxes in state space reveal non-equilibrium dynamics, quantifying this non-equilibrium behavior requires measuring the “distance” of a system from equilibrium. One such measure is the entropy production, a central concept in non-equilibrium statistical mechanics, which quantifies the amount of entropy that a system generates. Consider a system with joint transition probabilities \( P_{ij} = \text{Prob}[x_{t-1} = i, x_t = j] \), where \( x_t \) is the state of the system at time \( t \). If the dynamics are Markovian (as, for instance, is true for the Ising system), then the entropy production is given by

\[
S = \sum_{ij} P_{ij} \log \frac{P_{ij}}{P_{ji}},
\]

where the sum runs over all states \( i \) and \( j \). If the system obeys detailed balance (that is, if \( P_{ij} = P_{ji} \) for all pairs of states \( i \) and \( j \)), then the entropy production vanishes. Conversely, any violation of detailed balance leads to an increase in entropy production, thereby reflecting the non-equilibrium nature of the system.

Calculating the entropy production requires estimating the transition probabilities \( P_{ij} \). However, for complex systems the number of states grows exponentially with the size of the system, making a direct estimate of the entropy production infeasible. To overcome this hurdle, we employ a hierarchical clustering algorithm that groups similar states in our observed data into a single cluster, yielding a reduced number of coarse-grained states (Fig. 6a; see Methods). Estimating the entropy production this way yields two desirable properties: First, because a system’s entropy production can only decrease with coarse-graining, in order to establish that a system is fundamentally out of equilibrium, one must simply demonstrate that the coarse-grained entropy production is significantly greater than zero. Second, by defining the clusters hierarchically, we prove that the estimated entropy production becomes more accurate (ignoring finite data effects) as the number of clusters increases (Fig. S5). Indeed, across all temperatures in the Ising system, the estimated entropy production increases with the number of clusters \( k \), thereby providing an improving lower bound on the true entropy production (Fig. 6b). Moreover, as the temperature decreases the entropy production grows dramatically, revealing the stark difference in the non-equilibrium nature of the system at high versus low temperature.
Fig. 6. Estimating entropy production using hierarchical clustering.

a. Schematic of clustering procedure where axes represent the activities of individual components (e.g., brain regions in the neuroimaging data or spins in the Ising model), points reflect individual states observed in the time-series, shaded regions define clusters (or coarse-grained states), and arrows illustrate possible fluxes between clusters. b. Entropy production in the asymmetric SK model as a function of the number of clusters $k$ for the same time-series studied in Fig. 2c, with error bars reflecting two standard deviations estimated using trajectory bootstrapping (see Methods).

Entropy production in the brain

We are now prepared to investigate whether the brain operates at different distances from equilibrium when performing distinct functions. We study seven tasks, each of which engages a specific cognitive process and associated anatomical system: emotional processing, working memory, social inference, language processing, relational matching, gambling, and motor execution. To estimate the entropy production, we cluster the neuroimaging data (combined across all subjects and task settings, including rest) into $k = 8$ coarse-grained states, the largest number for which all transitions were observed at least once in each task (Fig. S6). Across all tasks and rest, the brain produces a significant amount of entropy, confirming that the brain operates out of equilibrium (Fig. 4a). Specifically, for all task settings the entropy production is significantly greater than the noise floor that arises due to finite data (one-sided $t$-test with $p < 0.001$). Furthermore, the
**Fig. 4** Entropy production in the brain varies with physical and cognitive demands.

- **a.** Entropy production at rest and during seven cognitive tasks, estimated using hierarchical clustering with $k = 8$ clusters.
- **b.** Entropy production as a function of response rate (i.e., the frequency with which subjects are asked to physically respond) for the tasks listed in panel a. Each response induces an average $0.07 \pm 0.03$ bits of produced entropy (Pearson correlation $r = 0.774, p = 0.024$).
- **c.** Entropy production for low cognitive load and high cognitive load conditions in the working memory task, where low and high loads represent 0-back and 2-back conditions, respectively, in an n-back task. The brain produces significantly more entropy during high-load than low-load conditions (one-sided t-test, $p < 0.001, t > 10, df = 198$). Across all panels, raw entropy productions (Eq. (1)) are divided by the fMRI repetition time $\Delta t = 0.72$ s to compute an entropy production rate, and error bars reflect two standard deviations estimated using trajectory bootstrapping (see Methods).
brain produces more entropy during all of the cognitive tasks than at rest, with each task inducing a distinct pattern of fluxes between states (Fig. S7). In fact, the motor task (wherein subjects are prompted to perform specific physical movements) induces a 20-fold increase in entropy production over resting-state dynamics, thereby demonstrating that the brain is capable of operating at a wide range of distances from equilibrium.

To carry out the physical and cognitive functions required for each task – from focusing attention to performing computations and responding to stimuli – the brain consumes large amounts of energy. In living systems generally, such energy consumption is often critical for supporting non-equilibrium dynamics. Therefore, it is natural to hypothesize that increases in physical or cognitive effort drive the brain away from equilibrium. Indeed, across tasks, we find that entropy production increases with the frequency of physical responses (Fig. 4b), with each response producing an additional 0.07 ± 0.03 bits of entropy. Furthermore, within the working memory task (which controls for the frequency of physical responses), the brain produces more entropy during conditions that require greater cognitive effort (Fig. 4c). We verify that these findings do not depend on the Markov assumption in Eq. (1) (Fig. S8), are robust to reasonable variation in the number of clusters k (Fig. S9), and cannot be explained by head motion in the scanner (a common confound in fMRI studies) nor variance in the activity time-series (Fig. S10). Together, these results suggest that physical and cognitive demands, which are supported by the consumption of energy, lead the brain to produce more entropy, thereby driving neural dynamics away from equilibrium.

Discussion

In this study, we describe a method for investing non-equilibrium dynamics by quantifying the amount of entropy that a system produces in its environment. While microscopic non-equilibrium processes are known to be vital for molecular and cellular operations, here we show that non-equilibrium dynamics also arise at large scales in complex living systems. Analyzing whole-brain imaging data, we find not only that the human brain functions out of equilibrium, but
that the brain’s entropy production (that is, its distance from equilibrium) increases with physical and cognitive exertion. Notably, the tools presented are non-invasive, applying to any system with time-series data, and can be used to study stochastic steady-state dynamics, rather than deterministic dynamics that trivially break detailed balance. Furthermore, the framework is not limited to the brain, but instead can be applied broadly to probe for non-equilibrium dynamics in complex systems, including collective behavior in human and animal populations\textsuperscript{31}, correlated patterns of neuronal firing\textsuperscript{32}, and aggregated activity in molecular and cellular networks\textsuperscript{33,34}. 


Methods

Calculating fluxes. Consider time-series data gathered in a time window $t_{\text{tot}}$, and let $n_{ij}$ denote the number of observed transitions from state $i$ to state $j$. The flux rate from state $i$ to state $j$ is given by $\omega_{ij} = (n_{ij} - n_{ji})/t_{\text{tot}}$. For the flux currents in Figs. 1d,e and 2c, the states of the system are points $(x, y)$ in two-dimensional space, and the state probabilities are estimated by $p(x, y) = t(x,y)/t_{\text{tot}}$, where $t(x,y)$ is the time spent in state $(x, y)$. The magnitude and direction of the flux through a given state $(x, y)$ is defined by the flux vector

$$u(x, y) = \frac{1}{2} \begin{pmatrix} \omega(x-1,y),(x,y) + \omega(x,y),(x+1,y) \\ \omega(x,y-1),(x,y) + \omega(x,y),(x,y+1) \end{pmatrix}. \tag{2}$$

In a small number of cases, two consecutive states in the observed time-series $x(t) = (x(t), y(t))$ and $x(t + 1) = (x(t + 1), y(t + 1))$ are not adjacent in state space. In these cases, we perform a linear interpolation between $x(t)$ and $x(t + 1)$ in order to calculate the fluxes between adjacent states.

Estimating errors using trajectory bootstrapping. The finite length of time-series data limits the accuracy with which quantities can be estimated. In order to calculate error bars on all estimated quantities, we apply trajectory bootstrapping\textsuperscript{15,22}. We first record the list of transitions

$$I = \begin{pmatrix} i_1 & i_2 \\ i_2 & i_3 \\ \vdots \\ i_{L-1} & i_L \end{pmatrix}, \tag{3}$$

where $i_\ell$ is the $\ell$th state in the time-series, and $L$ is the length of the time-series. From the transition list $I$, one can calculate all of the desired quantities; for instance, the fluxes are estimated by

$$\omega_{ij} = \frac{1}{t_{\text{tot}}} \sum_{\ell} \delta_{i,\ell} \delta_{j,\ell+1} - \delta_{j,\ell} \delta_{i,\ell+1}. \tag{4}$$

We remark that when analyzing the neural data, although we concatenate the time-series across subjects, we only include transitions in $I$ that occur within the same subject. That is, we do not include the transitions between adjacent subjects in the concatenated time-series.

To calculate errors, we construct bootstrap trajectories (of the same length $L$ as the original time-series) by sampling the rows in $I$ with replacement. For example, to compute errors for the flux vectors $u(x)$ in Figs. 1d,e and 2c, we first estimate the covariance matrix Cov$(u_1(x), u_2(x))$ by averaging over bootstrapped trajectories. Then, for each flux vector, we visualize its error by plotting an ellipse with axes aligned with the eigenvectors of the covariance matrix and radii equal to twice the square root of the corresponding eigenvalues (Fig. S1). All errors throughout the manuscript are calculated using 100 bootstrap trajectories.
The finite data length also induces a noise floor for each quantity, which is present even if the temporal order of the time-series is destroyed. To estimate the noise floor, we construct bootstrap trajectories by sampling individual data points from the time-series. We contrast these bootstrap trajectories with those used to estimate errors above, which preserve transitions by sampling the rows in $I$. The noise floor, which is calculated for each quantity by averaging over the bootstrap trajectories, is then compared with the estimated quantities. For example, rather than demonstrating that the average entropy productions in Fig. 4a are greater than zero, we establish that the distribution over entropy productions is significantly greater than the noise floor using a one-sided $t$-test with $p < 0.001$.

**Simulating the asymmetric Ising model.** The asymmetric Ising model is defined by a (possibly asymmetric) interaction matrix $J$, where $J_{\alpha\beta}$ represents the influence of spin $\beta$ on spin $\alpha$ (Fig. 2a), and a temperature $T \geq 0$ that tunes the strength of stochastic fluctuations. Here, we study a system with $N = 100$ spins, where each directed interaction $J_{\alpha\beta}$ is drawn independently from a zero-mean Gaussian with variance $1/N = 0.01$ (Fig. 2b). One can additionally include external fields $h_{\alpha}$, but for simplicity here we set them to zero. The state of the system is defined by a vector $x = (x_1, \ldots, x_N)$, where $x_\alpha = \pm 1$ is the state of spin $\alpha$. To generate time series, we employ Glauber dynamics with synchronous updates, a common Monte Carlo method for simulating Ising systems. Specifically, given the state of the system $x(t)$ at time $t$, the probability of spin $\alpha$ being “up” at time $t + 1$ (that is, the probability that $x_\alpha(t + 1) = 1$) is given by

$$\text{Prob}[x_\alpha(t + 1) = 1] = \frac{\exp \left( \frac{1}{T} \sum_{\beta} J_{\alpha\beta} x_\beta(t) \right)}{\exp \left( \frac{1}{T} \sum_{\beta} J_{\alpha\beta} x_\beta(t) \right) + \exp \left( - \frac{1}{T} \sum_{\beta} J_{\alpha\beta} x_\beta(t) \right)}.$$  \hspace{1cm} (5)

Stochastically updating each spin $\alpha$ according to Eq. (5), one arrives at the new state $x(t + 1)$. For each temperature in the Ising calculations in Figs. 2c and 6b, we generate a different time-series of length $L = 100,000$ with 10,000 trials of burn-in.

**Hierarchical clustering.** To estimate the entropy production of a system, one must first calculate the transition probabilities $P_{ij} = n_{ij}/(L - 1)$. For complex systems, the number of states $i$ (and therefore the number of transitions $i \rightarrow j$) grows exponentially with the size of the system $N$. For example, in the Ising model each spin $\alpha$ can take one of two values ($x_\alpha = \pm 1$), leading to $2^N$ possible states and $2^{2N}$ possible transitions. In order to estimate the transition probabilities $P_{ij}$, one must observe each transition $i \rightarrow j$ at least once, which requires significantly reducing the number of states in the system. Rather than defining coarse-grained states a priori, complex systems (and the brain in particular) often admit natural coarse-grained descriptions that are uncovered through dimensionality-reduction techniques.

Although one can use any coarse-graining technique to implement our framework and estimate entropy production, here we employ hierarchical $k$-means clustering for two reasons: (i) generally, $k$-means is perhaps the most common and simplest clustering algorithm, with demonstrated effectiveness fitting neural dynamics and (ii) specifically, by defining the clusters hierarchically we prove that the estimated entropy production becomes more
accurate as the number of clusters increases (ignoring finite data effects; Fig. S5).

In $k$-means clustering, one begins with a set of states (for example, those observed in our time-series) and a number of clusters $k$. Each observed state $x$ is randomly assigned to a cluster $i$, and one computes the centroid of each cluster. On the following iteration, each state is re-assigned to the cluster with the closest centroid (here we use cosine similarity to determine distance). This process is repeated until the cluster assignments no longer change. In a hierarchical implementation, one begins with two clusters; then one cluster is selected (typically the one with the largest spread in its constituent states) to be split into two new clusters, thereby defining a total of three clusters. This iterative splitting is continued until one reaches the desired number of clusters $k$. In Supplementary Sec. 6, we show that hierarchical clustering provides an increasing lower-bound on the entropy production; and in Supplementary Sec. 7, we demonstrate how to choose the number of clusters $k$.

**Neural data.** The whole-brain dynamics used in this study are measured and recorded using blood-oxygen-level-dependent (BOLD) functional magnetic resonance imaging (fMRI) collected from 590 healthy adults as part of the Human Connectome Project\[19\][29]. BOLD fMRI estimates neural activity by calculating contrasts in blood oxygen levels, without relying on invasive injections and radiation\[37\]. Specifically, blood oxygen levels (reflecting neural activity) are measured within three-dimensional non-overlapping voxels, spatially contiguous collections of which each represent a distinct brain region (or parcel). Here, we consider a parcellation that divides the cortex into 100 brain regions that are chosen to optimally capture the functional organization of the brain\[20\]. After processing the signal to correct for sources of systematic noise such as head motion (see Supplementary Sec. 12), the activity of each brain region is discretized in time, yielding a time-series of neural activity. For each subject, the shortest scan (corresponding to the emotional processing task) consists of 176 discrete measurements in time. In order to control for variability in data size across tasks, for each subject we only study the first 176 measurements in each task.

**Data Availability**

Data were provided by the Human Connectome Project (HCP); the Washington University, University of Minnesota, and Oxford University Consortium (Principal Investigators David Van Essen and Kamil Ugurbil; grant no. 1U54MH091657) funded by 16 NIH institutes and centers that support the NIH Blueprint for Neuroscience Research; and the McDonnell Center for Systems Neuroscience at Washington University.
Code Availability

The code that supports the findings of this study is available from the corresponding author upon reasonable request.

Supplementary Information. Supplementary text and figures accompany this paper.

Acknowledgements. The authors thank Erin Teich, Pragya Srivastava, Jason Kim, and Zhixin Lu for feedback on earlier versions of this manuscript. The authors acknowledge support from the John D. and Catherine T. MacArthur Foundation, the ISI Foundation, the Paul G. Allen Family Foundation, the Army Research Laboratory (W911NF-10-2-0022), the Army Research Office (Basset-W911NF-14-1-0679, Falk-W911NF-18-1-0244, Grafton-W911NF-16-1-0474, DCIST- W911NF-17-2-0181), the Office of Naval Research, the National Institute of Mental Health (2-R01-DC-009209-11, R01-MH112847, R01-MH107235, R21-M MH-106799), the National Institute of Child Health and Human Development (1R01HD08688-01), National Institute of Neurological Disorders and Stroke (R01 NS099348), and the National Science Foundation (NSF PHY-1554488, BCS-1631550, and NCS-FO-1926829).

Citation diversity statement. Recent work in neuroscience and other fields has identified a bias in citation practices such that papers from women and other minorities are under-cited relative to the number of such papers in the field. Here we sought to proactively consider choosing references that reflect the diversity of the field in thought, form of contribution, gender, race, geographic location, and other factors. Excluding self-citations to the authors of this paper and single-author citations, the first and last authors of references are 57% male/male, 23% female/male, 13% male/female, and 7% female/female. We look forward to future work that could help us better understand how to support equitable practices in science.

Author Contributions. C.W.L. and D.S.B. conceived the project. C.W.L. designed the framework and performed the analysis with input from E.J.C., L.P., and D.S.B. M.A.B. processed the neural data. C.W.L. wrote the manuscript and Supplementary Information; and E.J.C., L.P., M.A.B., and D.S.B. edited the manuscript and Supplementary Information.

Competing Interests. The authors declare no competing financial interests.
**Corresponding Author.** Correspondence and requests for materials should be addressed to D.S.B. (dsb@seas.upenn.edu).
Supplementary Information

Non-equilibrium dynamics and entropy production in the human brain

1 Introduction

In this Supplementary Information, we provide extended analysis and discussion to support the results presented in the main text. In Sec. 2, we describe how the flux vectors (in Figs. 1d,e and 2c of the main text) are calculated and illustrated. In Sec. 3, we show that principal component analysis (PCA) provides a natural low-dimensional embedding of neural dynamics that we can use to visualize fluxes between brain states. In Sec. 4, we show that, although the brain functions out of equilibrium, it does operate at a steady state. Demonstrating that the brain operates at a non-equilibrium steady-state opens the door for future investigations using tools and intuitions that have recently been generalized from equilibrium statistical mechanics\textsuperscript{40,42}. In Sec. 5, we show that if one shuffles the order of neural time-series data (thereby destroying the arrow of time), then the fluxes between states vanish and the system returns to equilibrium. In Sec. 6, we establish that estimating entropy production using hierarchical clustering yields two desirable properties. First, because a system’s entropy production can only decrease with coarse-graining\textsuperscript{13}, in order to establish that a system is fundamentally out of equilibrium, one must simply demonstrate that the coarse-grained entropy production is significantly greater than zero. Second, by defining the clusters hierarchically\textsuperscript{43}, we prove that the estimated entropy production becomes more accurate as the number of clusters increases. In Sec. 7, we demonstrate how to choose the number of clusters (or coarse-grained states) $k$ when estimating the entropy production. In Sec. 8, we present the flux between coarse-grained states in the neural dynamics as directed networks, which we refer to as flux networks. We demonstrate that these flux networks vary in structure across different cognitive tasks. In Secs. 9-11, we show that the entropy production results in the main text (Fig. 4) do not depend on the assumption that the neural dynamics are Markovian (Sec. 9), are robust to reasonable variation in the number of coarse-grained states $k$ (Sec. 10), and cannot be explained by head movement within the scanner nor variance in the neural time-series (Sec. 11). Finally, in
Sec. 12, we detail how the neural data was processed.

2 Visualizing flux currents

In order to visualize net fluxes in neural dynamics, we project the dynamics onto the first two principal components and employ a technique known as probability flux analysis\textsuperscript{22}. The net flux of transitions from a given state \((x, y)\) to its neighboring states can be visualized using the flux vector

\[
\mathbf{u}(x, y) = \frac{1}{2} \left( \begin{array}{c} 
\omega_{(x-1,y),(x,y)} + \omega_{(x,y),(x+1,y)} \\
\omega_{(x,y-1),(x,y)} + \omega_{(x,y),(x,y+1)} 
\end{array} \right),
\]

To compute the errors for a given flux vector \(\mathbf{u}(x)\), we calculate the covariance matrix \(\text{Cov}(u_1(x), u_2(x))\) by averaging over 100 bootstrapped trajectories. Then, we illustrate the errors by plotting an ellipse whose axes are aligned with the eigenvectors of the covariance matrix and whose radii are equal to twice the square root of the corresponding eigenvalues (Fig. S1).

![Fig. S1](visualizing_flux_vectors.png)

Schematic demonstrating how we illustrate the flux of transitions through a state (vector) and the errors in estimating the flux (ellipse).

3 Low-dimensional embedding using PCA

In order to visualize net fluxes between states in a complex system, we must project the dynamics onto two dimensions. While any pair of dimensions can be used to probe for broken detailed
balance, a natural choice is the first two principal components of the time-series data. Indeed, principal component analysis has been widely used to uncover low-dimensional embeddings of large-scale neural dynamics. Combining the time-series data from the rest and gambling task scans (that is, the data studied in Fig. 1 of the main text), we find that the first two principal components capture over 30% of the total variance in the observed recordings (Fig. S2a), thereby comprising a natural choice for two-dimensional projections. Moreover, we confirm that the projected dynamics capture approximately the same amount of variance in both the rest and gambling tasks, confirming that PCA is not overfitting the neural dynamics in one task or another (Fig. S2b).

**Fig. S2** PCA reveals low-dimensional embedding of neural dynamics.

- **a**, Cumulative fraction of variance explained by first ten principal components (line) and explained variance for each individual principal component (bars) in the combined rest and gambling data.
- **b**, For the same principal components (calculated for the combined rest and gambling data), we plot the cumulative fraction of variance explained (lines) and individual explained variance (bars) for the rest (red) and gambling (blue) data.

4 The brain operates at a stochastic steady state

Some of the tools and intuitions developed in traditional statistical mechanics to study equilibrium systems have recently been generalized to systems that exist at non-equilibrium steady states. For example, Evans et al. generalized the second law of thermodynamics to non-equilibrium steady-state systems by discovering the (steady state) fluctuation theorem. More recently, Dieterich et
Fig. S3 | Small changes in state probabilities imply steady-state dynamics.

Change in state probabilities \( \dot{p}_i \), normalized by the standard deviation \( \sigma_{\dot{p}_i} \), plotted as a function of the first two principal components at rest (a) and during the gambling task (b).

al. showed that, by mapping their dynamics to an equilibrium system at an effective temperature, some non-equilibrium steady-state systems are governed by a generalization of the fluctuation-dissipation theorem\(^{42}\). Thus, it is both interesting and practical to investigate whether the brain operates at a non-equilibrium steady state.

We establish in the main text that the brain operates out of equilibrium. To determine if the brain functions at a steady state, we must examine whether its state probabilities are stationary in time; that is, letting \( p_i \) denote the probability of state \( i \), we must determine whether \( \dot{p}_i = dp_i/dt = 0 \) for all states \( i \). The change in the probability of a state is equal to the net rate at which transitions flow into versus out of a state. For the two-dimensional dynamics studied in Fig. 1 in the main text, this relation takes the form

\[
\frac{dp_{(x,y)}}{dt} = \omega_{(x-1,y),(x,y)} - \omega_{(x,y),(x+1,y)} + \omega_{(x,y-1),(x,y)} - \omega_{(x,y),(x,y+1)},
\]

(7)

where \( \omega_{ij} = (n_{ij} - n_{ji})/t_{tot} \) is the flux rate from state \( i \) to state \( j \), \( n_{ij} \) is the number of observed transitions \( i \rightarrow j \), and \( t_{tot} \) is the temporal duration of the time-series\(^{21}\).

Here, we calculate the changes in state probabilities for both the rest and gambling scans.
Across all states in both task conditions, we find that these changes are indistinguishable from zero when compared to statistical noise (Fig. S3). Specifically, the changes in state probabilities are much less than twice their standard deviations, indicating that they cannot be significantly distinguished from zero with a $p$-value less than 0.05. Combined with the results from the main text, the stationarity of the neural state probabilities demonstrates that the brain operates at a non-equilibrium steady-state.

5 Shuffling time-series restores equilibrium

In the main text, we demonstrate that the brain operates out of equilibrium by exhibiting net fluxes between states (Fig. 1d,e in the main text). These fluxes break detailed balance and establish an arrow of time. Here we demonstrate that if the arrow of time is destroyed by shuffling the order of the neural time-series, then the fluxes vanish and equilibrium is restored. Specifically, for both the rest and gambling task scans, we generate 100 surrogate time-series with the order of the data randomly shuffled. Averaging across these shuffled time-series, we find that the fluxes between states are vanishingly small compared to statistical noise (Fig. S4), thus illustrating that the system has returned to equilibrium. We remark that other common surrogate data techniques, such as the random phases and amplitude adjusted Fourier transform surrogates, are not applicable here because they preserve the temporal structure of the time-series data.

6 Bounding entropy production using hierarchical clustering

Complex systems are often high-dimensional, with the number of possible states or configurations growing exponentially with the size of the system. In order to estimate the entropy production of a complex system, we must reduce the number of states through the use of coarse-graining, or dimensionality reduction, techniques. Interestingly, the entropy production admits a number of strong properties under coarse-graining. Of particular interest is the fact that the entropy production can only decrease under coarse-graining. Specifically, given two descriptions of a system, a “microscopic” description with states $\{i\}$ and a “macroscopic” description with states $\{i'\}$, we say that the second description is a coarse-graining of the first if there exists a surjective
Shuffled data do not exhibit net fluxes between brain states.

Probability distribution (color) and nearly imperceivable fluxes between states (arrows) for neural dynamics, which are shuffled and projected onto the first two principal components, both at rest (a) and during a gambling task (b). The flux scale is indicated in the upper right, and the disks represent two-standard-deviation confidence intervals for fluxes estimated using trajectory bootstrapping (see Methods in the main text).

map from the microstates \{i\} to the macrosctates \{i'\} (that is, if each microstate \(i\) is mapped to a unique macrostate \(i'\); Fig. S5a). Given such a coarse-graining, Esposito showed\(^{13}\) that the entropy production of the macroscopic description \(S'\) can be no larger than that of the microscopic description \(S\); in other words, the coarse-grained entropy production provides a lower bound for the original value, such that \(S' \leq S\).

The monotonic decrease of the entropy production under coarse-graining implies two desirable mathematical results. First, if one finds that any coarse-grained description of a system is out of equilibrium (that is, if the coarse-grained entropy production is significantly greater than zero), then one has immediately established that the full microscopic system is out of equilibrium (since the “true” microscopic entropy production is at least as large as the coarse-grained value). We use this fact in the main text to show – only by studying coarse-grained dynamics – that the brain fundamentally operates far from equilibrium.
a, Coarse-graining is defined by a surjective map from a set of microstates \{i\} to a set of macrostates \{i'\}. Under coarse-graining the entropy production can only decrease or remain the same. b, In hierarchical clustering, states are iteratively combined to form new coarse-grained states (or clusters). Each iteration defines a coarse-graining from \(k\) states to \(k - 1\) states, thereby forming a hierarchy of lower bounds on the entropy production.

Second, here we show that hierarchical clustering provides a hierarchy of lower bounds on the true entropy production. In hierarchical clustering, each cluster (or coarse-grained state) at one level of description (with \(k\) clusters) maps to a unique cluster at the level below (with \(k - 1\) clusters; Fig. S5b). This process can either be carried out by starting with a large number of clusters and then iteratively picking pairs of clusters to combine (known as agglomerative clustering), or by starting with a small number of clusters and then iteratively picking one cluster to split into two (known as divisive clustering, which we employ in our analysis\(^{[4]}\)). In both cases, the mapping from \(k\) clusters to \(k - 1\) clusters is surjective, thereby defining a coarse-graining of the system. Thus, letting \(S^{(k)}\) denote the entropy production estimated with \(k\) clusters, hierarchical clustering defines a hierarchy of lower bounds on the true entropy production \(S\):

\[
0 = S^{(1)} \leq S^{(2)} \leq S^{(3)} \leq \ldots \leq S. \tag{8}
\]

This hierarchy, in turn, demonstrates that the estimated entropy production \(S^{(k)}\) becomes more accurate with increasing \(k\).
We remark that the discussion above neglects finite data effects. We recall that estimating the entropy production requires first estimating the transition probabilities $P_{ij}$ from state $i$ to state $j$. This means that for $k$ clusters, one must estimate $k^2$ different probabilities. Thus, while increasing $k$ improves the accuracy of the estimated entropy production in theory, in practice increasing $k$ eventually leads to sampling issues that decrease the accuracy of the estimate. Given these competing influences, when analyzing real data the goal should be to choose $k$ such that it is as large as possible while still providing accurate estimates of the transition probabilities. We discuss how to choose $k$ in a reasonable manner in the following section.

7 Choosing the number of coarse-grained states

As discussed above, when calculating the entropy production, we wish to choose a number of coarse-grained states $k$ that is as large as possible while still arriving at an accurate estimate of the transition probabilities. One simple condition for estimating each transition probability $P_{ij}$ is that we observe the transition $i \to j$ at least once in the time-series. For all of the different tasks, Fig. S6a shows the fraction of the $k^2$ state transitions that are left unobserved after coarse-graining with $k$ clusters. We find that $k = 8$ is the largest number of clusters for which the fraction of unobserved transitions equals zero (within statistical errors) for all tasks; that is, the largest number of clusters for which all state transitions across all tasks were observed at least once. This is the primary reason why we used $k = 8$ coarse-grained states to analyze the brain’s entropy production in the main text (Fig. 4).

Interestingly, we find that $k = 8$ coarse-grained states is a good choice for two additional reasons. The first comes from studying the amount of variance explained by $k$ clusters (Fig. S6b). We find that the increase in explained variance from $k - 1$ to $k$ clusters is roughly constant for $k = 3$ and 4, then $k = 5$ to 8, and then $k = 9$ to 16. This pattern means that $k = 4$, 8, and 16 are natural choices for the number of coarse-grained states, since any further increase (say from $k = 8$ to 9) will yield a smaller improvement in explained variance. Similarly, the second reason for choosing $k = 8$ comes from studying the average distance between states within a cluster,
Choosing the number of coarse-grained states $k$.

**a.** Fraction of the $k^2$ state transitions that remain unobserved after hierarchical clustering with $k$ clusters for the different tasks. Error bars represent two standard deviations over 100 bootstrap trajectories for each task. **b.** Percent variance explained (top) and the increase in explained variance from $k - 1$ to $k$ clusters (bottom) as functions of $k$. **c.** Dispersion, or the average distance between data points within a cluster (top), and the decrease in dispersion from $k - 1$ to $k$ clusters (bottom) as functions of $k$.

which is known as the dispersion (Fig. S6c). Intuitively, a coarse-grained description with low dispersion provides a good fit of the observed data. Similar to the explained variance, we find that the decrease in dispersion from $k - 1$ to $k$ clusters is nearly constant for $k = 3$ to 4, then $k = 5$ to 8, and then $k = 9$ to 16, once again suggesting that $k = 4$, 8, and 16 are natural choices for the number of clusters. Together, these results demonstrate that the coarse-grained description with $k = 8$ states provides a good fit to the neural time-series data while still allowing for an accurate estimate of the entropy production in each task.

8 Flux networks: Visualizing flux between coarse-grained states

In the main text (Fig. 4), we demonstrated that the brain has the capacity to operate at a wide range of distances from equilibrium. We did so by estimating the amount of entropy the brain produces during different cognitive tasks. In addition to investigating the entropy production, one can also examine the specific neural processes underlying the brain’s non-equilibrium behavior, which are
encoded in the fluxes between coarse-grained states.

We find that each of the $k = 8$ states corresponds to high-amplitude activity in one or two cognitive systems \(^{29}\) (Fig. S7a). For each task, we can visualize the pattern of fluxes as a network, with nodes representing the coarse-grained states and directed edges reflecting net fluxes between states (Fig. S7b-i). These flux networks illustrate, for example, that the brain nearly obeys detailed balance during rest (Fig. S7b). Interestingly, in the emotion, working memory, social, relational, and gambling tasks (Fig. S7c-e,g,h) – all of which involve visual stimuli – the strongest fluxes connect visual (VIS) states. By contrast, these fluxes are weak in the language task (Fig. S7f), which only involves auditory stimuli. Finally, in the motor task, wherein subjects are prompted to make physical movements, the dorsal attention (DAT) state mediates fluxes between disparate parts of the network (Fig. S7i), perhaps reflecting the role of the DAT system in directing goal-oriented attention \(^{50, 51}\). In this way, the brain’s non-equilibrium dynamics are not driven by a single underlying mechanism, but rather emerge from a complex pattern of fluxes that changes depending on the task. Examining the structural properties and cognitive neuroscientific interpretations of these flux networks is an important direction for future studies.

9 Testing the Markov assumption

In the main text, we employ a definition of entropy production that relies on the assumption that the time-series is Markovian; that is, that the state $x_t$ of the system at time $t$ depends only on the previous state $x_{t-1}$ at time $t-1$. Specifically, the entropy production of a Markov system is given by

$$S = \sum_{ij} P_{ij} \log \frac{P_{ij}}{P_{ji}},$$

(9)

where $P_{ij} = \text{Prob}[x_{t-1} = i, x_t = j]$ is the probability of observing the transition $i \to j$. For real time-series data, however, the dynamics may not be Markovian, and Eq. (9) is not exact. In general, the entropy production (per trial) is given by

$$S = \lim_{\ell \to \infty} \frac{1}{\ell} \sum_{i_1, \ldots, i_{\ell+1}} P_{i_1, \ldots, i_{\ell+1}} \log \frac{P_{i_1, \ldots, i_{\ell+1}}}{P_{i_{\ell+1}, \ldots, i_1}},$$

(10)
Flux networks reveal non-equilibrium dynamics unique to each cognitive task.

a. Coarse-grained brain states calculated using hierarchical clustering ($k = 8$), with surface plots indicating the z-scored activation of different brain regions. For each state, we calculate the cosine similarity between its high-amplitude (green) and low-amplitude (orange) components and seven pre-defined neural systems: default mode (DMN), frontoparietal (FPN), visual (VIS), somatomotor (SOM), dorsal attention (DAT), ventral attention (VAT), and limbic...
We label each state according to its largest high-amplitude cosine similarities. Flux networks illustrating the fluxes between the eight coarse-grained states at rest (b) and during seven cognitive tasks: emotional processing (c), working memory (d), social inference (e), language processing (f), relational matching (g), gambling (h), and motor execution (i). Edge weights indicate flux rates, and fluxes are only included if they are significant relative to the noise floor induced by the finite data length (one-sided t-test, \( p < 0.001 \)).

where \( P_{i_1,\ldots,i_{\ell+1}} = \text{Prob}[x_{t-\ell} = i_1, \ldots, x_t = i_{\ell+1}] \) is the probability of observing the sequence of states \( i_1, \ldots, i_{\ell+1} \). If the dynamics are Markovian, for example, then the limit converges for \( \ell = 1 \) and we recover Eq. (9). In general, one can approximate Eq. (10) by evaluating the function inside the limit for \( \ell \) as large as possible. In order to do so, however, one must estimate \( k^{\ell+1} \) different probabilities for a system with \( k \) states. Thus, given data limitations, it is often impractical to estimate the entropy production beyond the Markov approximation (\( \ell = 1 \)).

Here we demonstrate that the main conclusions about entropy production in the brain (summarized in Fig. 4 in the main text) do not depend qualitatively on the Markov approximation in Eq. (9). To do so, we consider the second-order approximation

\[
S \approx \frac{1}{2} \sum_{i,j,k} P_{ijk} \log \frac{P_{ijk}}{P_{kji}},
\]

which incorporates information about sequences of length three. As in the main text, we cluster the neural data using \( k = 8 \) coarse-grained states. Given that we are now required to estimate \( k^3 = 512 \) probabilities rather than just \( 8^2 = 64 \), there are inevitably entries in the sum in Eq. (11) that are infinite (i.e., those corresponding to reverse-time sequences \( k \to j \to i \) that are not observed in the time-series). As is common, we simply ignore these terms.

Across the different task settings, we find that the second-order entropy productions are nearly identical to the first-order (Markov) approximations presented in the main text (Fig. S8a). Moreover, the second-order entropy production remains significantly correlated with the frequency of physical responses in different tasks, with each response still inducing an additional \( 0.07 \pm 0.03 \) bits of produced entropy (Fig. S8b). Finally, in the working memory task, the second-order entropy
Fig. S8 | Second-order approximation of entropy production in the brain.

**a**, Second-order entropy production at rest and during seven cognitive tasks (dark bars), estimated using hierarchical clustering with \( k = 8 \) clusters. For comparison, we also include the first-order entropy productions from Fig. 4a in the main text (light bars). **b**, Second-order entropy production as a function of response rate for the tasks listed in panel **a** (dark points). Each response induces an average \( 0.07 \pm 0.03 \) bits of produced entropy (Pearson correlation \( r = 0.770, p = 0.026 \)). For comparison, we include the first-order entropy productions from Fig. 4b in the main text (light points). **c**, We find a significant difference in the second-order entropy production between low cognitive load and high cognitive load conditions in the working memory task (dark bars), where low and high loads represent 0-back and 2-back conditions, respectively (one-sided \( t \)-test, \( p < 0.001, t > 10, df = 198 \)). For comparison, we include the first-order entropy productions from Fig. 4c in the main text (light bars). Across all panels, second-order entropy productions (calculated using Eq. (11)) are divided by the fMRI
repetition time $\Delta t = 0.72$ s to compute an entropy production rate, and error bars reflect two standard deviations estimated using trajectory bootstrapping (see Methods in the main text).

production remains larger for high-load conditions than low-load conditions (Fig. S8c), suggesting that cognitive demands drive the brain away from equilibrium. Together, these results demonstrate that the brain’s entropy production is well-approximated by the Markov formulation used in the main text (Eq. (9)).

10 Varying the number of coarse-grained states

In Sec. 7, we presented methods for choosing the number of coarse-grained states $k$, concluding that $k = 8$ is an appropriate choice for our neural data. However, it is important to check that the entropy production results from the main text (summarized in Fig. 4) do not vary significantly with our choice of $k$. In Fig. S9a, we plot the estimated entropy production for each task setting (including rest) as a function of the number of coarse-grained states $k$. We find that the tasks maintain approximately the same ordering across all choices of $k$ considered, with the brain producing the least entropy during rest, the most entropy during the motor task, and the second most entropy during the gambling task. Furthermore, we find that the correlation between entropy production and physical response rate (Fig. 4b in the main text) remains significant for all $k \leq 8$ (that is, for all choices of $k$ for which we observe all transitions at least once in each task; Fig. S9a) as well as $k = 9, 11, 12, 13, \text{ and } 14$ (Fig. S9b). We remark that we do not study the case $k = 2$ because the entropy production is zero by definition for two-state systems (Fig. S9a). Finally, we confirm that the brain produces significantly more entropy during high-cognitive-load conditions than low-cognitive-load conditions in the working memory task (Fig. 4c in the main text) for all choices of $k$ considered (Fig. S9c). Together, these results demonstrate that the relationships between entropy production and physical and cognitive effort are robust to reasonable variation in the number of coarse-grained states $k$. 
Fig. S9 | Entropy production in the brain at different levels of coarse-graining.

a, Entropy production at rest and during seven cognitive tasks as a function of the number of clusters $k$ used in the hierarchical clustering. The raw entropy production (Eq. 9) is divided by the fMRI repetition time $\Delta t = 0.72$ s to compute an entropy production rate, and error bars reflect two standard deviations estimated using trajectory bootstrapping (see Methods in the main text).

b, Slope of the linear relationship between entropy production and physical response rate across tasks for different numbers of clusters $k$. Error bars represent one-standard-deviation confidence intervals of the slope and asterisks indicate the significance of
the Pearson correlation between entropy production and response rate. c, Difference between the entropy production during high-load and low-load conditions of the working memory task as a function of the number of clusters $k$. Error bars represent two standard deviations estimated using trajectory bootstrapping (see Methods in the main text), and the entropy production difference is significant across all values of $k$ (one-sided $t$-test, $p < 0.001$).

11 Robustness to head motion and signal variance

In the main text, we showed that the brain’s entropy production is significantly correlated with the frequency of physical responses (Fig. 4b) and increases during periods of cognitive exertion (Fig. 4c). Here, we show that the effects of physical and cognitive effort on entropy production cannot be explained by head movement within the scanner (a common confound in fMRI studies) nor variance in the neural time-series. To quantify head movement, for each time point in every time-series, we compute the spatial standard deviation of the difference between the current image and the previous image. This quantity, known as DVARS, is a common measure of head movement in fMRI data. Importantly, we find that entropy production is not significantly correlated with the average DVARS within each task (Fig. S10a), thereby demonstrating that the relationship between entropy production and physical response rate is not simply due to the confound of subject head movement within the scanner. Additionally, we find that entropy production is not significantly correlated with the variance of the neural data within each task (Fig. S10b). This final result establishes that our entropy production estimates are not simply driven by variations in the amount of noise in the neural data across different tasks.

12 Data processing

The resting, emotional processing, working memory, social inference, language processing, relational matching, gambling, and motor execution fMRI scans are from the S1200 Human Connectome Project release. Brains were normalized to fslr32k via the MSM-AII registration with 100 regions. CompCor, with five principal components from the ventricles and white matter masks, was used to regress out nuisance signals from the time series. In addition, the 12 detrended motion
Fig. S10 | Entropy production in the brain cannot be explained by head movement nor signal variance.

Entropy production versus the average DVARS (a) and the variance of the neural time-series (b) at rest and during seven cognitive tasks. Across both panels, entropy productions are estimated using hierarchical clustering with $k = 8$ clusters and are divided by the fMRI repetition time $\Delta t = 0.72$ s to compute entropy production rates. Error bars reflect two standard deviations estimated using trajectory bootstrapping (see Methods in the main text).

estimates provided by the Human Connectome Project were regressed out from the regional time series. The mean global signal was removed and then time series were band-pass filtered from 0.009 to 0.08 Hz. Then, frames with greater than 0.2 mm frame-wise displacement or a derivative root mean square (DVARS) above 75 were removed as outliers. We filtered out sessions composed of greater than 50 percent outlier frames, and we only analyzed data from subjects that had all scans remaining after this filtering, leaving 590 individuals. The processing pipeline used here has previously been suggested to be ideal for removing false relations between neural dynamics and behavior. Finally, for each subject and each scan, we only analyze the first 176 time points, corresponding to the length of the shortest task (emotional processing); this truncation controls for the possibility of data size affecting comparisons across tasks.
References

1. Schrödinger, E. *What is life? The physical aspect of the living cell and mind* (Cambridge University Press Cambridge, 1944).

2. Brangwynne, C. P., Koenderink, G. H., MacKintosh, F. C. & Weitz, D. A. Cytoplasmic diffusion: Molecular motors mix it up. *J. Cell Biol.* **183**, 583–587 (2008).

3. Yin, H., Artsimovitch, I., Landick, R. & Gelles, J. Nonequilibrium mechanism of transcription termination from observations of single rna polymerase molecules. *Proc. Natl. Acad. Sci.* **96**, 13124–13129 (1999).

4. Huang, K. C., Meir, Y. & Wingreen, N. S. Dynamic structures in Escherichia coli: Spontaneous formation of mine rings and mind polar zones. *Proc. Natl. Acad. Sci.* **100**, 12724–12728 (2003).

5. Mehta, P. & Schwab, D. J. Energetic costs of cellular computation. *Proc. Natl. Acad. Sci.* **109**, 17978–17982 (2012).

6. Lan, G., Sartori, P., Neumann, S., Sourjik, V. & Tu, Y. The energy–speed–accuracy trade-off in sensory adaptation. *Nat. Phys.* **8**, 422 (2012).

7. e Silva, M. S. *et al.* Active multistage coarsening of actin networks driven by myosin motors. *Proc. Natl. Acad. Sci.* **108**, 9408–9413 (2011).

8. Stuhrmann, B., e Silva, M. S., Depken, M., MacKintosh, F. C. & Koenderink, G. H. Nonequilibrium fluctuations of a remodeling in vitro cytoskeleton. *Phys. Rev. E* **86**, 020901 (2012).

9. Harris, J. J., Jolivet, R. & Attwell, D. Synaptic energy use and supply. *Neuron* **75**, 762–777 (2012).

10. Lynn, C. W. & Bassett, D. S. The physics of brain network structure, function and control. *Nat. Rev. Phys.* **1**, 318 (2019).

11. Gnesotto, F., Mura, F., Gladrow, J. & Broedersz, C. Broken detailed balance and nonequilibrium dynamics in living systems: A review. *Rep. Prog. Phys.* **81**, 066601 (2018).
12. Carnot, S. *Reflexions sur la puissance motrice du feu* (Bachelier, Paris, France, 1824).

13. Esposito, M. Stochastic thermodynamics under coarse graining. *Phys. Rev. E* **85**, 041125 (2012).

14. Egolf, D. A. Equilibrium regained: From nonequilibrium chaos to statistical mechanics. *Science* **287**, 101–104 (2000).

15. Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948).

16. Norberg, K. & Siejö, B. Cerebral metabolism in hypoxic hypoxia. ii. citric acid cycle intermediates and associated amino acids. *Brain Res.* **86**, 45–54 (1975).

17. Du, F. *et al.* Tightly coupled brain activity and cerebral ATP metabolic rate. *Proc. Natl. Acad. Sci.* **105**, 6409–6414 (2008).

18. Erečińska, M. & Silver, I. A. Atp and brain function. *J. Cereb. Blood Flow Metab.* **9**, 2–19 (1989).

19. Van Essen, D. C. *et al.* The WU-Minn Human Connectome Project: An overview. *Neuroimage* **80**, 62–79 (2013).

20. Thomas Yeo, B. *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).

21. Cunningham, J. P. & Byron, M. Y. Dimensionality reduction for large-scale neural recordings. *Nat. Neurosci.* **17**, 1500 (2014).

22. Battle, C. *et al.* Broken detailed balance at mesoscopic scales in active biological systems. *Science* **352**, 604–607 (2016).

23. Zia, R. & Schmittmann, B. Probability currents as principal characteristics in the statistical mechanics of non-equilibrium steady states. *J. Stat. Mech.* **2007**, P07012 (2007).
24. Newman, M. & Barkema, G. *Monte Carlo methods in statistical physics* (Oxford University Press, New York, USA, 1999).

25. Sherrington, D. & Kirkpatrick, S. Solvable model of a spin-glass. *Phys. Rev. Lett.* **35**, 1792 (1975).

26. Seifert, U. Entropy production along a stochastic trajectory and an integral fluctuation theorem. *Phys. Rev. Lett.* **95**, 040602 (2005).

27. Roldán, É. & Parrondo, J. M. Estimating dissipation from single stationary trajectories. *Phys. Rev. Lett.* **105**, 150607 (2010).

28. Lamrous, S. & Taileb, M. Divisive hierarchical k-means. In *CIMCA*, 18–18 (IEEE, 2006).

29. Barch, D. M. *et al.* Function in the human connectome: Task-fMRI and individual differences in behavior. *Neuroimage* **80**, 169–189 (2013).

30. Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S. & Turner, R. Movement-related effects in fmri time-series. *Magn. Reson. Med.* **35**, 346–355 (1996).

31. Castellano, C., Fortunato, S. & Loreto, V. Statistical physics of social dynamics. *Rev. Mod. Phys.* **81**, 591 (2009).

32. Palva, J. M. *et al.* Neuronal long-range temporal correlations and avalanche dynamics are correlated with behavioral scaling laws. *Proc. Natl. Acad. Sci.* **110**, 3585–3590 (2013).

33. Koenderink, G. H. *et al.* An active biopolymer network controlled by molecular motors. *Proc. Natl. Acad. Sci.* **106**, 15192–15197 (2009).

34. Van Aelst, L. & DSouza-Schorey, C. Rho gtpases and signaling networks. *Genes Dev.* **11**, 2295–2322 (1997).

35. Cornblath, E. J. *et al.* Temporal sequences of brain activity at rest are constrained by white matter structure and modulated by cognitive demands. *Communications Biology, in press*. 

37
36. Liu, X. & Duyn, J. H. Time-varying functional network information extracted from brief instances of spontaneous brain activity. *Proc. Natl. Acad. Sci.* **110**, 4392–4397 (2013).

37. Raichle, M. E. Behind the scenes of functional brain imaging: A historical and physiological perspective. *Proc. Natl. Acad. Sci.* **95**, 765–772 (1998).

38. Dworkin, J. D. *et al.* The extent and drivers of gender imbalance in neuroscience reference lists. *arXiv preprint arXiv:2001.01002* (2020).

39. Caplar, N., Tacchella, S. & Birrer, S. Quantitative evaluation of gender bias in astronomical publications from citation counts. *Nat. Astron.* **1**, 1–5 (2017).

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

41. Evans, D. J., Cohen, E. G. D. & Morriss, G. P. Probability of second law violations in shearing steady states. *Phys. Rev. Lett.* **71**, 2401 (1993).

42. Dieterich, E., Camunas-Soler, J., Ribezzi-Crivellari, M., Seifert, U. & Ritort, F. Single-molecule measurement of the effective temperature in non-equilibrium steady states. *Nat. Phys.* **11**, 971–977 (2015).

43. Jain, A. K., Murty, M. N. & Flynn, P. J. Data clustering: A review. *ACM Comput. Surv.* **31**, 264–323 (1999).

44. Friston, K., Frith, C., Liddle, P. & Frackowiak, R. Functional connectivity: The principal-component analysis of large (PET) data sets. *J. Cereb. Blood Flow Metab.* **13**, 5–14 (1993).

45. Shine, J. M. *et al.* Human cognition involves the dynamic integration of neural activity and neuromodulatory systems. *Nat. Neurosci.* **22**, 289 (2019).

46. Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. & Farmer, J. D. Testing for nonlinearity in time series: the method of surrogate data. *Physica D* **58**, 77–94 (1992).
47. Gomez-Marín, A., Parrondo, J. M. & Van den Broeck, C. Lower bounds on dissipation upon coarse graining. *Phys. Rev. E* **78**, 011107 (2008).

48. Shiraishi, N. & Sagawa, T. Fluctuation theorem for partially masked nonequilibrium dynamics. *Phys. Rev. E* **91**, 012130 (2015).

49. Polettini, M. & Esposito, M. Effective thermodynamics for a marginal observer. *Phys. Rev. Lett.* **119**, 240601 (2017).

50. Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L. & Raichle, M. E. Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proc. Natl. Acad. Sci.* **103**, 10046–10051 (2006).

51. Vossel, S., Geng, J. J. & Fink, G. R. Dorsal and ventral attention systems: distinct neural circuits but collaborative roles. *Neuroscientist* **20**, 150–159 (2014).

52. Roldán, É. & Parrondo, J. M. Entropy production and Kullback-Leibler divergence between stationary trajectories of discrete systems. *Phys. Rev. E* **85**, 031129 (2012).

53. Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L. & Petersen, S. E. Spurious but systematic correlations in functional connectivity mri networks arise from subject motion. *Neuroimage* **59**, 2142–2154 (2012).

54. Schaefer, A. et al. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri. *Cereb. Cortex* **28**, 3095–3114 (2018).

55. Siegel, J. S. *et al.* Data quality influences observed links between functional connectivity and behavior. *Cereb. Cortex* **27**, 4492–4502 (2017).