Practical measures of integrated information for stationary systems

Adam B. Barrett and Anil K. Seth

Sackler Centre for Consciousness Science and *School of Informatics*, University of Sussex, Brighton, UK

adam.barrett@sussex.ac.uk, http://www.sussex.ac.uk/sackler

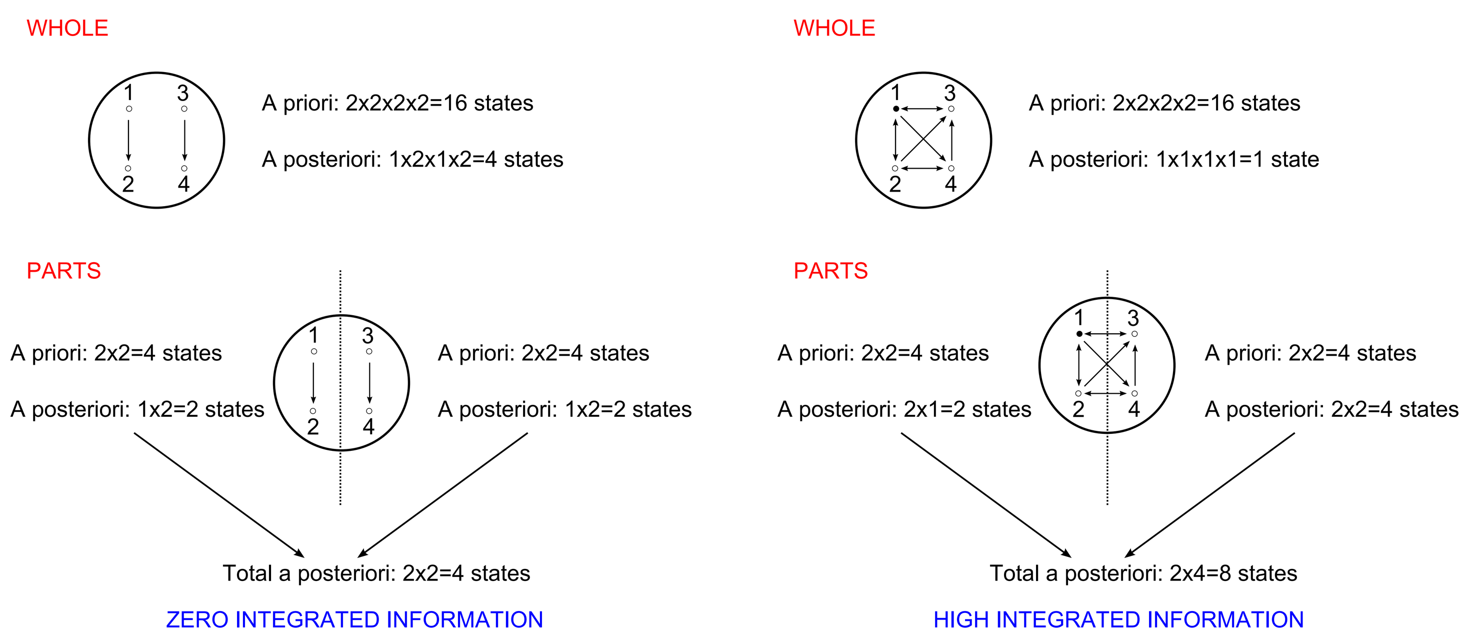
University of Sussex
Sackler Centre for Consciousness Science

Integrated information

Integrated information, Φ, quantifies the extent to which a system as a whole generates more information than the sum of its parts. Φ has been suggested to measure the quantity of consciousness generated by a system [1,2], and shares with similar approaches [3] an emphasis on measuring conjoined integration and segregation in a system's dynamics, reflecting basic properties of conscious phenomenology.

Φ_{2008}

The latest formulation, Φ_{2008}, measures the integrated information generated when a system transitions into a particular state [1,2]. Information is quantified as the reduction in entropy of a hypothetical past state that, *a priori*, had maximum entropy.



Measuring Φ_{2008} requires 2 **assumptions** that deny application to real neural systems. Namely, that the system be

1. **Memoryless:** The maximum entropy distribution can only be imposed as an initial condition. In the brain we only observe states with history. Hypothetical evolution of a state without history can only be characterized if the system has no memory.
2. **Discrete:** Continuous variables don't have a unique maximum entropy distribution.

Stationary Φ (Φ_{Stat}): definition

We propose a new measure, Φ_{Stat}, based on information generated by the current state about an *actual* past state. This enables application to any **stationary** time-series data.

- The **effective information** (EI) with respect to bipartition $\mathcal{B} = \{M^1, M^2\}$ is the information generated by the whole X minus the sum of that of the parts. For current state $\boldsymbol{x} = (\boldsymbol{m}^1, \boldsymbol{m}^2)$, the EI for timescale τ is:

$$\text{EI}[\boldsymbol{x}, \tau, \mathcal{B}] =: I(\boldsymbol{X}_{t-\tau}|\boldsymbol{X}_t = \boldsymbol{x}) - \sum_{k=1}^{2} I(\boldsymbol{M}^k_{t-\tau}|\boldsymbol{M}^k_t = \boldsymbol{m}^k) \,. \qquad (1)$$

- The **minimum information bipartition** (MIB) minimizes normalized EI:

$$\mathcal{B}^{\text{MIB}}(\boldsymbol{x}, \tau) =: \arg_{\mathcal{B}} \min \left\{ \frac{\text{EI}[\boldsymbol{x}, \tau, \mathcal{B}]}{K(\mathcal{B})} \right\} , \qquad (2)$$

$$K(\{M^1, M^2\}) =: \min_k [H(\boldsymbol{M}^k_t)] \,. \qquad (3)$$

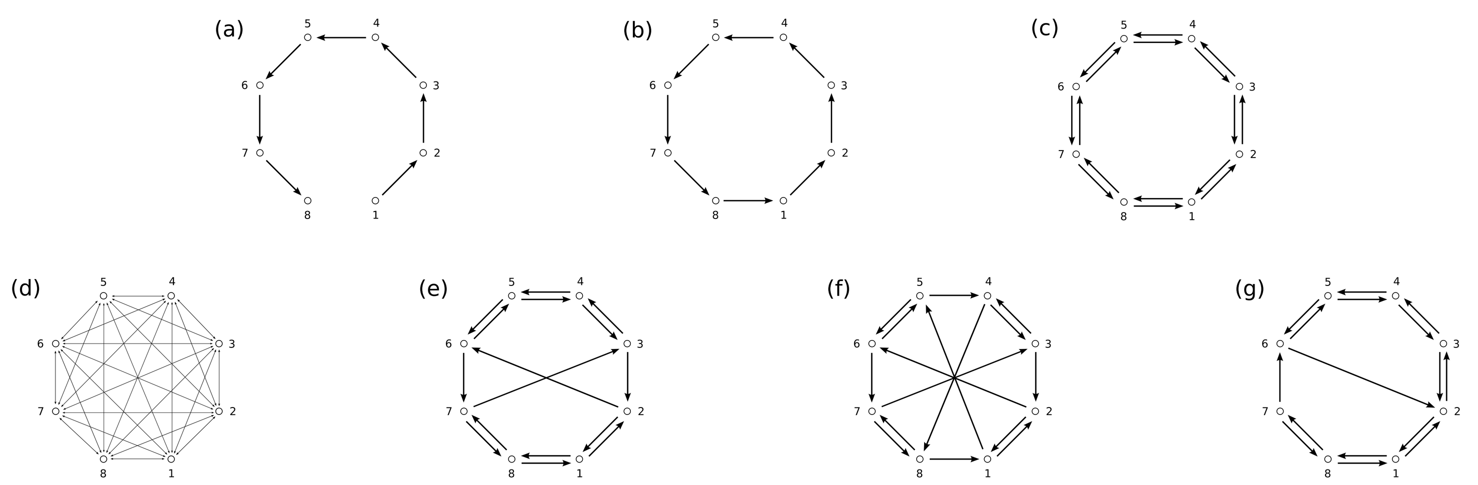
Normalization is necessary because sub-systems that are almost as large as the whole generate almost as much information as the whole; we require a bias towards approximately even bipartitions.

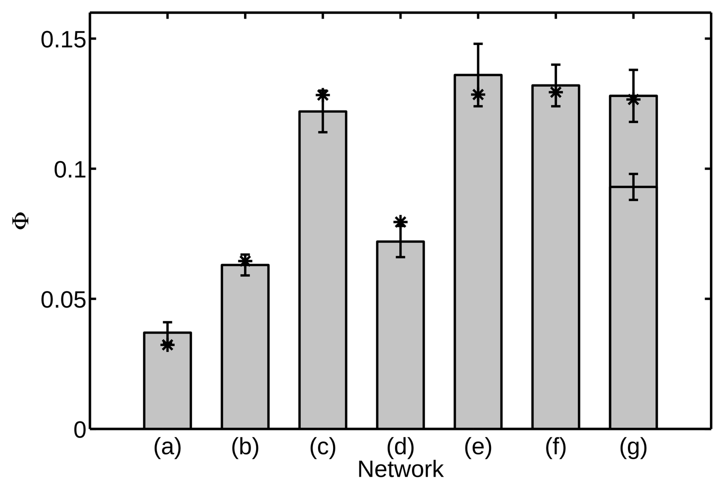
- Finally, the **integrated information** is the EI with respect to the MIB:

$$\Phi_{\text{Stat}}[\boldsymbol{x}, \tau] =: \text{EI}[\boldsymbol{x}, \tau, \mathcal{B}^{\text{MIB}}(\boldsymbol{x}, \tau)] \,. \qquad (4)$$

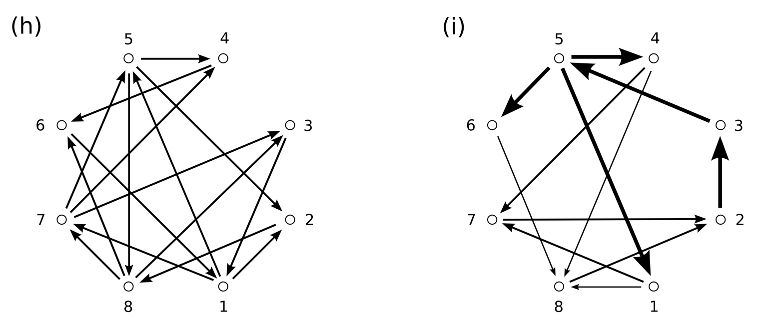
Φ_{Stat} in practice

Φ_{Stat} can be computed numerically from **time-series** and analytically given a generative model. For Gaussian systems, only covariance matrices are required [4], making it extremely easy to apply in practice.





Φ_{Stat} computed analytically and from simulated time-series for example networks. Simulations typically give accurate results, but network (g) exhibits instability.



Optimized networks with (h) 2 equal afferents per node, $\Phi_{\text{Stat}} = 0.25$; (i) all afferents to a given node equal, $\Phi_{\text{Stat}} = 0.30$. (Constant total afferent.)

Auto-regressive Φ (ARΦ)

For non-Gaussian systems we introduce an alternative measure, ARΦ, based on how well the present **predicts** the past, but only to the extent that predictions based on the whole outstrip predictions based on parts independently. We consider the regression

$$\boldsymbol{X}_{t-\tau} = A \cdot \boldsymbol{X}_t + \boldsymbol{\epsilon}_t \,, \qquad (5)$$

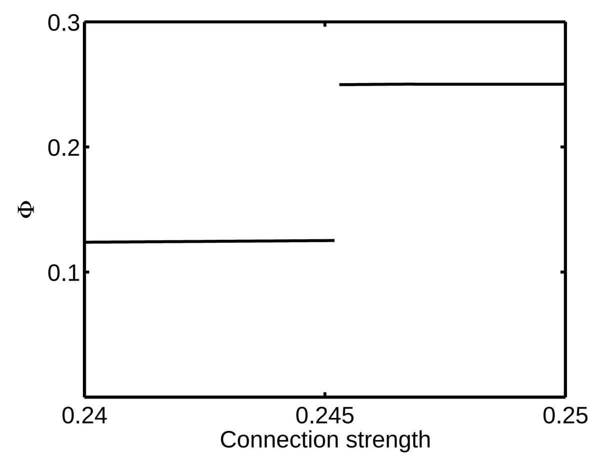
and define predictive power (PP) as

$$\text{PP}(\boldsymbol{X}_{t-\tau}|\boldsymbol{X}_t) =: \log|\text{COV}(\boldsymbol{X}_{t-\tau})| - \log|\text{COV}(\boldsymbol{\epsilon}_t)| \,. \qquad (6)$$

(For a part M, substitute M for X.) Then construct ARΦ analogously to Φ_{Stat}, replacing I with PP.

- ARΦ is equivalent to Φ_{Stat} for Gaussian systems [4], whereas for non-Gaussian systems it provides a pragmatic alternative to Φ_{Stat}.
- ARΦ uses linear regression of the present to **predict the past**. It thus invites comparison with **causal density** [3], which uses regression of the present to **predict the future**.

Instability



Varying connection 6 → 1 in network (h) results in a discontinuity in Φ_{Stat}. This arises from using **normalized** EI to find the MIB but **non-normalized** EI to determine Φ.

Conclusion

- Our measures provide new opportunities for empirically testing the relationship between integrated information and consciousness.
- ARΦ (and Φ_{Stat} for Gaussian systems) are state-independent. If they correlate with consciousness, then conscious level (i) is constant during each stationary epoch in brain activity; (ii) changes when functional connectivity changes.
- Φ_{Stat} and ARΦ use the **stationary** distribution and measure a **'process'**. By contrast, Φ_{2008} uses the **maximum entropy** distribution and measures a **'capacity'**.
- The normalization-instability challenges ascription of physical meaning to Φ.

Acknowledgments: AKS is supported by EPSRC Leadership Fellowship EP/G007543/1, which also supports the work of ABB. Support is also gratefully acknowledged from the Dr. Mortimer and Theresa Sackler Foundation.

References: [1] D. Balduzzi, G. Tononi *PLoS Comput Biol* 4(6) (2008). [2] G. Tononi *Biol Bull* 215(3) (2008). [3] A.K. Seth *Cog Comput* 1(1) (2009). [4] L. Barnett, A.B. Barrett, A.K. Seth *Phys Rev Lett* 103 (2009).