

Neural complexity and structural connectivity

L. Barnett* and C. L. Buckley†

Department of Informatics, Centre for Computational Neuroscience and Robotics, School of Science and Technology,
University of Sussex, Brighton BN1 9QH, United Kingdom

S. Bullock‡

School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, United Kingdom

(Received 17 September 2008; revised manuscript received 13 March 2009; published 19 May 2009)

Tononi *et al.* [Proc. Natl. Acad. Sci. U.S.A. **91**, 5033 (1994)] proposed a measure of neural complexity based on mutual information between complementary subsystems of a given neural network, which has attracted much interest in the neuroscience community and beyond. We develop an approximation of the measure for a popular Gaussian model which, applied to a continuous-time process, elucidates the relationship between the complexity of a neural system and its structural connectivity. Moreover, the approximation is accurate for weakly coupled systems and computationally cheap, scaling polynomially with system size in contrast to the full complexity measure, which scales exponentially. We also discuss connectivity normalization and resolve some issues stemming from an ambiguity in the original Gaussian model.

DOI: [10.1103/PhysRevE.79.051914](https://doi.org/10.1103/PhysRevE.79.051914)

PACS number(s): 87.18.Sn, 87.19.lj, 87.19.lj, 89.75.Fb

I. INTRODUCTION

Over the last decade, considerable use has been made of a measure of neural complexity developed by Tononi *et al.* [1]. Their \mathcal{C}_N measure is motivated by the tension between local and holistic perspectives on neural systems. According to the authors, the hallmark of neural complexity is the *combination* of functional modularity and global unity of behavior across many anatomical scales. Their measure is an attempt to capture this theoretical position by employing calculations of mutual information shared between processes occurring on complementary subsystems of a neural network in order to identify networks that simultaneously support relatively segregated local behavior and relatively integrated global dynamics [2].

Although conceived within a specific neuroscience context, \mathcal{C}_N has received widespread attention across the behavioral and brain sciences, from fields as diverse as autonomous robotics [3], neural imaging [4,5], local dynamics of the mammalian brain [6], and the exploration of theories of sleep, consciousness, and schizophrenia [7]. Increasingly it is being seen as part of a family of statistical measures of complexity derived from information theory [8] with particular relevance within networks science [9–11].

Network complexity is typically characterized as a *structural* phenomenon, with complex networks occupying a regime between the totally ordered (as exemplified by lattice-like structures) and the totally disordered (as exemplified by Erdős-Rényi random graphs). \mathcal{C}_N by contrast, while suggesting a consonant order-disorder balance, is essentially a *dynamical* measure, defined in terms of a time-varying process running on a system of interconnected neural components. It thus seems natural and useful to inquire into the relationship

between the connectivity structure of a network and the dynamical complexity of processes running on that network [6,10]. Specifically, we might ask what type of connectivity structure is likely to host neural processes measuring high complexity under \mathcal{C}_N . Tononi *et al.* [1], for instance, demonstrate that network structures inspired by the properties of the cerebral cortex score highly, and that artificial networks optimized for high neural complexity share structural motifs with real cortical networks [12]. In this paper we attempt to establish a framework under which we may elucidate the structure-complexity relationship. We approach the problem via an analytic model which allows expression of an approximation to \mathcal{C}_N directly in terms of the underlying network connectivity structure. The approximation also turns out to be useful in its own right as a computationally cheap proxy for \mathcal{C}_N under the model assumptions.

It must be emphasized that \mathcal{C}_N is a *statistical* measure: it requires that, on the time scale under consideration, interactions between neural components may be approximated as *statistically stationary*. This may be viewed as a restriction on the scenarios for which it makes sense to apply the measure; thus while it would be inappropriate for application, say, to the short-term behavior of an embodied neural system reacting to temporally unpredictable external stimuli, it might be appropriate to, e.g., categorize neural states associated with long-term neurocognitive phenomena such as sleep states, epileptogenesis, or anesthetic action [13,14].

Various approaches can be taken in calculating the complexity measure. Here we are concerned with one popular analytic model that proceeds on the assumption that neural dynamics may be (approximately) characterized by a stationary multivariate stochastic process and that, furthermore, this can be taken to be Gaussian, enabling interactions between network components—and thence mutual information—to be expressed via a covariance matrix [1]. Some comments are perhaps in order regarding the Gaussian assumption. The model we consider lacks explicit extrinsic inputs; rather, neuronal activity is triggered by uncorrelated Gaussian noise [15], which we may think of, in the spirit of statistical me-

*l.c.barnett@sussex.ac.uk

†c.l.buckley@sussex.ac.uk

‡sgb@ecs.soton.ac.uk

chanics, as a rough approximation to the statistical structure of signals sampled from the environment. Implementations of Gaussian neural models as linear autoregressive processes (see, e.g., [1] and Sec. IV below) might be considered as approximations in the weakly coupled near-linear regime of the nonlinear dynamics present in more detailed neurocomputational models and, indeed, biological neural systems. While there is no dispute that linear stationary dynamics are likely to represent a poor model on the level of *individual* neurons, in the neuroscience literature a linear stationary Gaussian approximation is quite commonly invoked for large-scale systems and in particular *neural mass* models [16–21]. We remark that, from a practical point of view, the Gaussian assumption vastly simplifies estimation of $\mathcal{C}_{\mathcal{N}}$ from real or simulated data since reliable estimation of entropies from multivariate samples is highly nontrivial [20]. We should, however, be wary of inferring results of highly nonlinear systems from linear Gaussian models as featured in this and related studies.

We proceed to derive a computationally cheap approximation of $\mathcal{C}_{\mathcal{N}}$ for the weakly coupled Gaussian model in terms of the correlation matrix. We then address a problem in [1] regarding calculation of the covariance matrix (we stress that the complexity measure itself is not called into question but rather its operationalization for a particular model process). The issue is resolved by introduction of a continuous-time analog of the original, apparently discrete time, formulation. Our approximation is then reformulated for the continuous-time Gaussian model in terms of the associated connection matrix. We discuss how the resultant approximate measure elucidates the relationship between neural complexity and structural network connectivity, flagging *reciprocal* and (more generally) *cyclic* connectivity motifs as the principal contributors to neural complexity. We also indicate how moving to continuous time may in effect mitigate the impact of the problematic covariance matrix calculation. Issues regarding connectivity scaling and normalization are discussed. Finally, we illustrate our analysis by example of a simple ring lattice connectivity scheme, where connectivity strength decays with interneuron distance. The example ratifies the accuracy of our approximation and demonstrates, under suitable conditions, a complexity peak at intermediate decay values in line with the complexity peak found in [1] for an *ad hoc* covariance matrix with off-diagonal covariance decay. We have made available for download MATLAB® code implementing some of the analytical methods presented in this paper [22].

II. NEURAL COMPLEXITY

The scenario we address [1,12,23,24] is that of a system of n “neural components” (*nodes* for brevity) and a stationary multivariate stochastic process $\mathbf{X}(t) \equiv \{X_i(t) | i=1, \dots, n\}$, where $X_i(t)$ is to represent the activity at time t associated with the i th node. This process may be in discrete or continuous time. In [1] the authors introduced a *neural complexity* measure based on mutual information between subsystems of the given system. The idea behind the measure is that complex neural systems should be expected to exhibit a

balance between “integration” and “segregation” of neural subsystems. The measure is defined as follows: first the *integration* associated with the system is introduced as

$$\mathcal{I} \equiv \sum_{i=1}^n H_i - H, \quad (1)$$

where H denotes the entropy $H(\mathbf{X}(t))$ of the full joint process $\mathbf{X}(t)$ and H_i denotes the entropy $H(X_i(t))$ of the individual activation $X_i(t)$. Note that by stationarity these quantities and hence \mathcal{I} itself do not depend on time t . \mathcal{I} may be interpreted as a measure of the deviation from independence of the individual components of the system. Neural complexity is then defined to be

$$\mathcal{C}_{\mathcal{N}} \equiv \sum_{k=1}^{n-1} \left(\frac{k}{n} \mathcal{I} - \langle \mathcal{I} \rangle_k \right), \quad (2)$$

where $\langle \cdot \rangle_k$ denotes an average *over all subsystems of size k* ; more precisely, a k *subsystem* of the neural system may be specified by an ordered vector of unique indices: $\mathbf{i} = (i_1, \dots, i_k)$, with $1 \leq i_1 < i_2 < \dots < i_k \leq n$. If we then have some measurable quantity $Z(\mathbf{i})$, say, associated with k subsystem \mathbf{i} we denote by $\langle Z \rangle_k$, the average $\binom{n}{k}^{-1} \sum_{\mathbf{i}} Z(\mathbf{i})$ of Z where the sum ranges over all $\binom{n}{k}$ k -subsystem indices \mathbf{i} . Entropy and hence integration \mathcal{I} may be naturally restricted to k subsystems and are therefore such quantities. We note that $\langle \mathcal{I} \rangle_n = \mathcal{I}$ so that the $k=n$ term in Eq. (2) would be identically zero.

It can be shown [1] that neural complexity may be expressed in the equivalent and somewhat more computationally tractable form

$$\mathcal{C}_{\mathcal{N}} \equiv \sum_{k=1}^{n-1} \left(\langle H \rangle_k - \frac{k}{n} H \right). \quad (3)$$

In the special case where the $\mathbf{X}(t)$ is multivariate Gaussian, the entropy H may be expressed simply in terms of the $n \times n$ covariance matrix $\Omega \equiv \overline{\mathbf{X}(t)^T \mathbf{X}(t)}$ [25], where the overbar represents an average over the statistical ensemble [26]. By stationarity Ω does not depend on time t and we have $H = \frac{1}{2} \ln([2\pi e]^n |\Omega|)$, where $|\cdot|$ denotes the determinant. Now it is easy to see that the $\mathbf{X}(t)$ restricted to any k subsystem \mathbf{i} will also be multivariate Gaussian and that the associated covariance matrix $\Omega(\mathbf{i})$ will be just the $k \times k$ submatrix of Ω defined by the indices i_1, \dots, i_k ; i.e., $\Omega(\mathbf{i})_{\alpha\beta} \equiv \Omega_{i_\alpha i_\beta}$. From Eq. (3) then, neural complexity may be expressed as [22]

$$\mathcal{C}_{\mathcal{N}} = \frac{1}{2} \sum_{k=1}^{n-1} \left(\langle \ln |\Omega| \rangle_k - \frac{k}{n} \ln |\Omega| \right). \quad (4)$$

For the remainder of this paper we assume that $\mathbf{X}(t)$ is multivariate Gaussian.

Let \mathcal{R} be the correlation matrix for the multivariate Gaussian system, with entries $r_{ij} \equiv \Omega_{ij} / \sqrt{\Omega_{ii} \Omega_{jj}}$. Then $|\Omega| = \Omega_{11} \dots \Omega_{nn} |\mathcal{R}|$ and it is straightforward to show (the log variance terms cancel out in the average over subsystems) that we can write $\mathcal{C}_{\mathcal{N}}$ in terms of \mathcal{R} as

$$\mathcal{C}_{\mathcal{N}} = \frac{1}{2} \sum_{k=1}^{n-1} \left(\langle \ln |\mathcal{R}| \rangle_k - \frac{k}{n} \ln |\mathcal{R}| \right). \quad (5)$$

III. APPROXIMATING NEURAL COMPLEXITY

There does not seem to be any convenient analytic expression for $\mathcal{C}_{\mathcal{N}}$ as defined by Eq. (4) and (5) and computing it numerically involves averaging over all 2^n possible subsystems. Thus it seems appropriate to ask whether a computationally cheaper *approximation* for $\mathcal{C}_{\mathcal{N}}$ might be found. We now derive such an approximation in terms of the correlation coefficients r_{ij} . Making use of the matrix identity $|\exp(A)| \equiv \exp[\text{trace}(A)]$ for any square matrix A and using the Taylor-series expansion for $\ln(1+x)$, we find

$$\ln |\mathcal{R}| = - \sum_{m=2}^{\infty} \frac{(-1)^m}{m} \text{trace}(\hat{\mathcal{R}}^m), \quad (6)$$

where we define $\hat{\mathcal{R}} \equiv \mathcal{R} - I$. Note that $\hat{\mathcal{R}}$ is symmetric and zero on the diagonal. Setting $\epsilon \equiv \|\hat{\mathcal{R}}\|$, with $\|\cdot\|$ as any *consistent matrix norm* [27], the m th term of the expansion is $\mathcal{O}(\epsilon^m)$. We note that the *spectral radius* of a matrix is the infimum of all *induced* matrix norms and that all induced norms are consistent [27]. If the spectral radius $\rho(\hat{\mathcal{R}})$ of $\hat{\mathcal{R}}$ is small, then—we might reasonably describe such a system as *weakly coupled* since correlations between neural components will be small—we can expect the first few terms of expansion (6) to furnish a good approximation for $\ln |\mathcal{R}|$.

We now use Eq. (6) to calculate the first two terms for the corresponding expansion of expression (5) for $\mathcal{C}_{\mathcal{N}}$. From Eqs. (B1) and (B2), respectively, we have

$$\langle \text{trace}(\hat{\mathcal{R}}^2) \rangle_k = \frac{k(k-1)}{n(n-1)} \text{trace}(\hat{\mathcal{R}}^2), \quad (7)$$

$$\langle \text{trace}(\hat{\mathcal{R}}^3) \rangle_k = \frac{k(k-1)(k-2)}{n(n-1)(n-2)} \text{trace}(\hat{\mathcal{R}}^3). \quad (8)$$

Using the identity

$$\sum_{k=1}^{n-1} k(k-1) \cdots (k-m+1) = \frac{n(n-1) \cdots (n-m)}{m+1}, \quad (9)$$

it then follows from Eq. (5) that

$$\mathcal{C}_{\mathcal{N}} = \mathcal{C}_{\mathcal{N}}^*(\mathcal{R}) + \mathcal{C}_{\mathcal{N}}^{**}(\mathcal{R}) + \mathcal{O}(\epsilon^4), \quad (10)$$

where [22]

$$\mathcal{C}_{\mathcal{N}}^*(\mathcal{R}) = \frac{n+1}{24} \text{trace}(\hat{\mathcal{R}}^2), \quad (11)$$

$$\mathcal{C}_{\mathcal{N}}^{**}(\mathcal{R}) = - \frac{n+1}{24} \text{trace}(\hat{\mathcal{R}}^3), \quad (12)$$

with $\mathcal{C}_{\mathcal{N}}^*(\mathcal{R}) = \mathcal{O}(\epsilon^2)$ and $\mathcal{C}_{\mathcal{N}}^{**}(\mathcal{R}) = \mathcal{O}(\epsilon^3)$. We can in principle calculate the approximation to successively higher orders although the mathematical labor involved quickly becomes prohibitive.

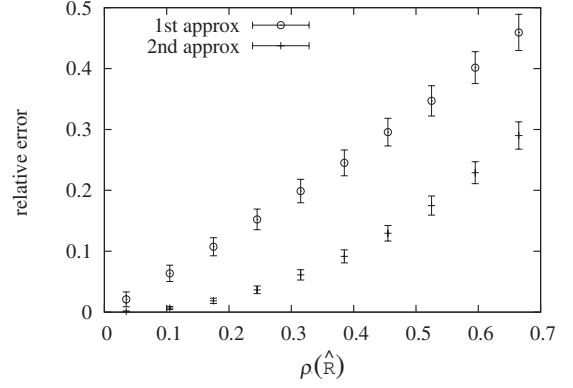


FIG. 1. Mean relative error of first and second approximations relative to exact neural complexity plotted against spectral radius of the dedagonalized correlation matrix. Error bars indicate one standard deviation.

Extensive experimentation with a variety of covariance matrices suggests that, at least if the coupling strength is not too high, Eq. (10) provides a good approximation to $\mathcal{C}_{\mathcal{N}}$. Even when the approximation error is large the *qualitative* behavior of $\mathcal{C}_{\mathcal{N}}$ appears to be well reflected. We note that the approximation is considerably cheaper to compute than the full measure $\mathcal{C}_{\mathcal{N}}$, scaling polynomially, as opposed to exponentially, with system size n .

To validate the accuracy of the approximation we ran the following experiment: 10^4 random covariance matrices of size 20×20 were generated and the exact neural complexity $\mathcal{C}_{\mathcal{N}}$ [Eqs. (2)–(5)] was calculated for each, along with the first- and second-order approximations $\mathcal{C}_{\mathcal{N}}^*(\mathcal{R})$ and $\mathcal{C}_{\mathcal{N}}^{**}(\mathcal{R}) + \mathcal{C}_{\mathcal{N}}^*(\mathcal{R})$, respectively [Eqs. (11) and (12)]. Covariance matrices were generated as $\Omega = (I - wC^\top)(I - wC)$ with off-diagonal elements of the square matrix C [28] drawn from a normal distribution $\mathcal{N}(0.8, 0.64)$ and on-diagonal elements from $\mathcal{N}(0, 0.04)$. The parameter w controls coupling strength, and was varied between 0 and 0.04. For each sampled covariance matrix we binned the relative errors $|\mathcal{C}_{\mathcal{N}} - \mathcal{C}_{\mathcal{N}}^*(\mathcal{R})|/|\mathcal{C}_{\mathcal{N}}|$ and $|\mathcal{C}_{\mathcal{N}} - \mathcal{C}_{\mathcal{N}}^*(\mathcal{R}) - \mathcal{C}_{\mathcal{N}}^{**}(\mathcal{R})|/|\mathcal{C}_{\mathcal{N}}|$ of the first- and second-order approximations with respect to the exact neural complexity, against the spectral radius $\rho(\hat{\mathcal{R}})$ of the dedagonalized correlation matrix $\hat{\mathcal{R}}$. Results are displayed in Fig. 1. Mean relative errors scale roughly linearly with $\rho(\hat{\mathcal{R}})$, confirming that $\rho(\hat{\mathcal{R}})$ provides a good measure of the accuracy of the approximations.

Figure 2 displays the approximations against exact neural complexity as a scatter plot. We see that the accuracy of the approximations degrade somewhat with the magnitude of neural complexity.

For systems that may reasonably be modeled as Gaussian, all that is required to calculate approximation (10) is a covariance matrix, which might, e.g., be bootstrap sampled from time-series data: we thus propose that the approximation be considered as a practicable and computationally cheap proxy for $\mathcal{C}_{\mathcal{N}}$. We remark that in [12,24] the authors introduce the simplified measure $\mathcal{C}(X)$, defined as the $k=n-1$ term of expression (2) for the full measure $\mathcal{C}_{\mathcal{N}}$ (see, e.g., Fig. 3c of [24]). From the above derivation it is clear that the corresponding $\mathcal{O}(\epsilon^2)$ and $\mathcal{O}(\epsilon^3)$ terms for the measure $\mathcal{C}(X)$ will

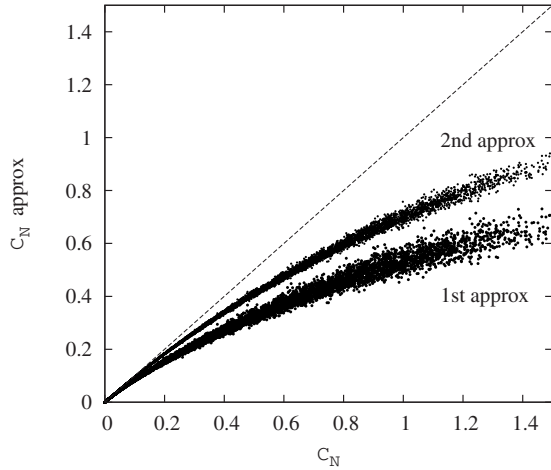


FIG. 2. Scatter plot of first approximation $C_N^*(\mathcal{R})$ and second approximation $C_N^*(\mathcal{R}) + C_N^{**}(\mathcal{R})$ [Eqs. (11) and (12)] against exact neural complexity C_N [Eqs. (2)–(5)] for randomly generated correlation matrices (see text for details). The closer a point is to the diagonal (dashed line), the better the approximation.

differ from those for C_N by constant factors depending just on system size n . This goes some way to justifying those authors’ intuition that $\mathcal{C}(X)$ —itself computationally cheaper than C_N —is a reasonable proxy for the full measure C_N .

IV. GAUSSIAN MODEL IN CONTINUOUS TIME

Tononi *et al.* [1] consider a model in which we are given an $n \times n$ connectivity matrix C , where C_{ij} is to be interpreted as the *weight* on the connection from node i to node j (i.e., node i is efferent while node j afferent). The neural process is described as linear and driven by uncorrelated Gaussian noise. We note that this model effectively reframes neural complexity as a “structural” property of neural *connectivity* rather than a “functional” property of neural *dynamics*.

The covariance matrix is then stated to be

$$\Omega = Q^\top Q, \quad (13)$$

where $Q \equiv (I - C)^{-1}$, based on an equation of the form [29]

$$\mathbf{X} = \mathbf{X} \cdot C + \mathbf{R}, \quad (14)$$

for the stationary process, where \mathbf{R} represents uncorrelated mean-zero unit-variance Gaussian noise. But there are problems with this derivation; it appears to be based on the erroneous assumption that stationarity of $\mathbf{X}(t)$ implies *equality* of the $\mathbf{X}(t)$ rather than the correct inference that the $\mathbf{X}(t)$ all have the same *distribution*. We consider the implications of this issue in Sec. VI.

In fact a correct derivation of the covariance matrix requires explicit specification for the Gaussian process $\mathbf{X}(t)$. In [1] it is not entirely clear whether the neural process is intended to be in discrete or continuous time although the form of Eq. (14) suggests the former; we return to this point in Sec. VI. Although it might be argued that in any case a discrete time model will be unacceptably unrealistic: real (biological) neural processes certainly cannot be said to occur in globally synchronized discrete time steps and even

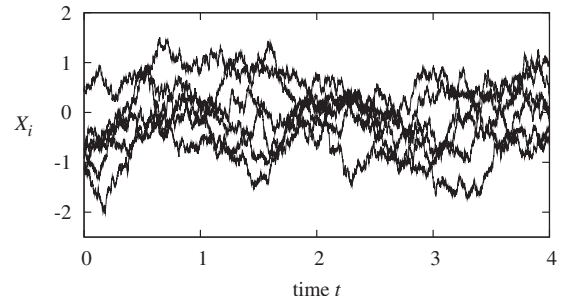


FIG. 3. An instantiation of the Ornstein-Uhlenbeck process [Eq. (15)]. See text for details.

computational models in nonbiological information processing applications are frequently simulated in continuous time. Below we introduce a continuous-time model and calculate the associated covariance matrix Ω , from which neural complexity C_N may be calculated via Eq. (4). In the section following we derive an approximation for C_N in terms of connectivity C .

We propose, as a continuous-time model in the spirit of the model presented in [1], the multivariate Ornstein-Uhlenbeck process [30,31]

$$d\mathbf{X}(t) = -\mathbf{X}(t) \cdot (I - C)dt + d\mathbf{W}(t), \quad (15)$$

where I is the identity matrix and $\mathbf{W}(t)$ a multivariate Wiener process with identity covariance matrix representing uncorrelated Gaussian noise. The process is multivariate Gaussian so that Eq. (4) applies. If we allow noise *levels* to differ per node, then we may recover an equivalent equation to Eq. (15) by a simple linear transformation of the connectivity matrix and a rescaling of activation levels. We note that Eq. (15) is entirely equivalent to the linearized version of a noisy neural network of Wilson-Cowan type [32] presented, e.g., in [33]. Alternatively, it may be viewed as a linearized noisy continuous time recurrent neural network [34]. While we view the impact of nonlinearity on neural dynamics as an important issue (cf. our remarks in Sec. I and see also [5]), it is not one that we address in this study.

Note that the *decay* (i.e., relaxation rate) of all nodes in Eq. (15) appears to be the same. This is arguably unrealistic; but it is moot whether the diagonal elements of C should properly be regarded as decay parameters, as opposed to *self-activations* (i.e., weights on self-connections). We remark that the issue of self-activation remains somewhat controversial in the computational neuroscience literature [35,36]. Now we could incorporate decay explicitly into Eq. (15); however, it complicates the mathematics considerably so we prefer to remain deliberately noncommittal on the question of decay vs self-activation (cf. Secs. VII and VIII).

Figure 3 illustrates an instantiation of process (15) with 20 variables (only six are plotted). The connectivity matrix C was constructed over a randomly generated directed graph with 20 vertices, 70 edges, and no self-connections. Weights C_{ij} , for $i \neq j$ were assigned to the connections at random from a normal distribution $\mathcal{N}(0.2, 0.09)$ and the on-diagonal weights C_{ii} from $\mathcal{N}(0, 0.0001)$. The process was simulated

using the exact method (Appendix A), sampled over 4 units of time t at intervals of $dt=0.001$.

We may approximate continuous process (15) by the discrete process

$$\mathbf{X}(t+dt) = \mathbf{X}(t) \cdot [I - (I - C)dt] + \mathbf{R}(t)\sqrt{dt}, \quad (16)$$

for a small time step dt , where $\mathbf{R}(t)$ is uncorrelated mean-zero unit-variance Gaussian noise. Setting for convenience

$$K \equiv I - (I - C)dt, \quad (17)$$

we have then

$$\begin{aligned} \mathbf{X}(t+dt)^\top \mathbf{X}(t+dt) &= K^\top \cdot \mathbf{X}(t)^\top \mathbf{X}(t) \cdot K + K^\top \cdot \mathbf{X}(t)^\top \mathbf{R}(t)\sqrt{dt} \\ &\quad + \mathbf{R}(t)^\top \mathbf{X}(t) \cdot K\sqrt{dt} + \mathbf{R}(t)^\top \mathbf{R}(t)dt. \end{aligned} \quad (18)$$

Averaging this equation over the ensemble *at a given time* t gives

$$\overline{\mathbf{X}(t+dt)^\top \mathbf{X}(t+dt)} = K^\top \cdot \overline{\mathbf{X}(t)^\top \mathbf{X}(t)} \cdot K + Idt, \quad (19)$$

noting that $\overline{\mathbf{R}(t)^\top \mathbf{R}(t)} \equiv I$ for all t , and that all other terms in the average vanish since $\mathbf{R}(t)$ is uncorrelated with $\mathbf{X}(t)$. At this point we invoke stationarity, as a consequence of which all (joint) moments of the process $\mathbf{X}(t)$ —and hence the covariance matrix $\Omega \equiv \overline{\mathbf{X}(t)^\top \mathbf{X}(t)}$ —do not vary with time so that the covariance matrix satisfies the relation

$$\Omega = K^\top \Omega K + Idt. \quad (20)$$

Substituting back the value of K from Eq. (17) and working to $\mathcal{O}(dt)$, we find that in the continuous limit $dt \rightarrow 0$

$$2\Omega = I + C^\top \Omega + \Omega C. \quad (21)$$

If C is *symmetric* we may solve Eq. (21) explicitly to get

$$\Omega = \frac{1}{2}(I - C)^{-1}. \quad (22)$$

Otherwise it may be solved in terms of the eigenvalues and eigenvectors of C (see, e.g., [33]). For our purposes, however, it suffices to note that we may expand Ω in the matrix power series [22]

$$\begin{aligned} 2\Omega &= \sum_{m=0}^{\infty} 2^{-m} \sum_{k=0}^m \binom{m}{k} (C^k)^\top C^{m-k} \\ &= I + \frac{1}{2}(C^\top + C) + \frac{1}{4}[(C^2)^\top + 2C^\top C + C^2] + \dots, \end{aligned} \quad (23)$$

insofar as it converges.

For stationarity of Eq. (16), for fixed dt we need $|\mu| < 1$ for all eigenvalues μ of K . But the eigenvalues of K are precisely $1 - (1 - \lambda)dt$, where λ are the eigenvalues of C . We thus need $|1 - (1 - \lambda)dt| < 1$ which, in the continuous limit $dt \rightarrow 0$, yields the stationarity condition for process (15)

$$\Re(\lambda) < 1 \text{ for every eigenvalue } \lambda \text{ of } C. \quad (24)$$

We note that, however, Eq. (24) does not appear to guarantee convergence of series (23). However, it is not too difficult to show [37] that a sufficient condition for convergence is $\rho(C) < 1$, where $\rho(C)$ is the *spectral radius* [38] of C ; that is, the maximum of $|\lambda|$ over all eigenvalues λ of C . Note that $\rho(C) < 1$ implies Eq. (24).

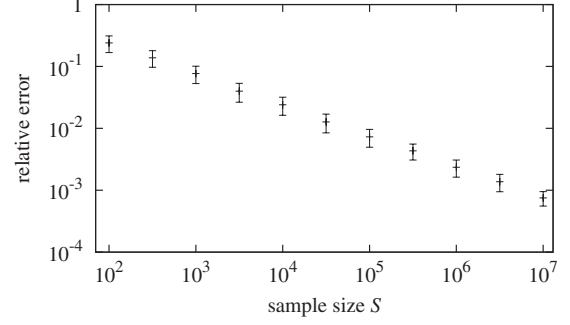


FIG. 4. Mean relative error of sample covariance matrices with respect to the value predicted by Eq. (21), plotted against sample size (log-log scale) for instantiations of the Ornstein-Uhlenbeck process [Eq. (15)] on connection matrices constructed over random graphs (see text for details). Error bars indicate one standard deviation.

Looking forward in time, process (15) decorrelates with exponential decay. Specifically, we may calculate that the lagged covariance is given by

$$\overline{\mathbf{X}(t)^\top \mathbf{X}(t+s)} = \Omega \exp[-(I - C)s]. \quad (25)$$

To verify analytical expression (21) for Ω we carried out the following experiment: connectivity matrices C were generated over 20-node directed random graphs with no self-connections and connection probability $p=0.3$ (disconnected graphs were rejected). Weights were then assigned to the (off-diagonal) connections at random from a normal distribution $\mathcal{N}(0.2, 0.09)$ and on-diagonal weights from a normal distribution $\mathcal{N}(0, 0.0001)$. Connectivity matrices with spectral radius $\rho(C) < 1$ were rejected in order to guarantee existence of a stationary solution and convergence of power series (23). For each connectivity matrix C , the stationary covariance matrix Ω was calculated via an iterative procedure [22] based on the series (23). 100 independent instantiations of the Ornstein-Uhlenbeck process were then simulated using the exact method for a number $10 \times S$ of time steps of size $dt=1$, with S as the sample size. The resulting time series (the process would first be run long enough to attain stationarity) was then bootstrap sampled S times and a covariance matrix was constructed from the sample. We then calculated the *relative error* $\|\Omega - \Omega^\dagger\|/\|\Omega\|$ for each instantiation, where Ω^\dagger is the sample covariance matrix and $\|\cdot\|$ denotes the *Frobenius norm* $\|M\| \equiv \text{trace}(M^\top M)$ on matrix space [27]. The relative errors were averaged over the 100 trial instantiations, and the mean and standard deviations were calculated. This procedure was repeated for a sequence of sample sizes S ; the results, plotted on a *log-log* scale are illustrated in Fig. 4. They indicate that sample stationary covariance matrices indeed approach the value predicted by Eq. (21), with sample relative error related to sample size by an inverse power law.

V. APPROXIMATING NEURAL COMPLEXITY FOR THE ORNSTEIN-UHLENBECK PROCESS

From the approximation formula (10) for \mathcal{C}_N we now derive a corresponding approximation—this time in terms of

the connectivity matrix C —for the Ornstein-Uhlenbeck process [Eq. (15)]. Let $\varepsilon \equiv \|C\|$ with $\|\cdot\|$ again as any consistent matrix norm. Writing

$$U \equiv \frac{1}{2}(C^\top + C), \quad (26)$$

$$V \equiv \frac{1}{4}[(C^\top)^\top + 2C^\top C + C^2], \quad (27)$$

we have from Eq. (23) that $2\Omega = I + U + V + \mathcal{O}(\varepsilon^3)$ with $U = \mathcal{O}(\varepsilon)$ and $V = \mathcal{O}(\varepsilon^2)$. We may then calculate that for $i \neq j$ the correlation coefficients are given by

$$r_{ij} = U_{ij} + W_{ij} + \mathcal{O}(\varepsilon^3), \quad (28)$$

where

$$W_{ij} \equiv V_{ij} - \frac{1}{2}(U_{ii} + U_{jj})U_{ij} \quad (29)$$

is $\mathcal{O}(\varepsilon^2)$. It is clear that ε will be of the same order of magnitude as the $\epsilon \equiv \|\hat{\mathcal{R}}\|$ of Sec. III. We then have

$$\text{trace}(\hat{\mathcal{R}}^2) = \sum_{i \neq j} U_{ij}^2 + 2 \sum_{i \neq j} U_{ij} W_{ij} + \mathcal{O}(\varepsilon^4), \quad (30)$$

$$\text{trace}(\hat{\mathcal{R}}^3) = \sum_{i \neq j \neq k} U_{ij} U_{ik} U_{jk} + \mathcal{O}(\varepsilon^4), \quad (31)$$

and from Eq. (10), gathering terms of the same order in ε , we may calculate

$$\mathcal{C}_{\mathcal{N}} = \mathcal{C}_{\mathcal{N}}^*(C) + \mathcal{C}_{\mathcal{N}}^{**}(C) + \mathcal{O}(\varepsilon^4), \quad (32)$$

where [22]

$$\mathcal{C}_{\mathcal{N}}^*(C) \equiv \frac{n+1}{24} \sum_{i \neq j} U_{ij}^2, \quad (33)$$

$$\mathcal{C}_{\mathcal{N}}^{**}(C) \equiv \frac{n+1}{24} \sum_{i \neq j} U_{ij} (2V_{ij} - U_{ij}^2), \quad (34)$$

with $\mathcal{C}_{\mathcal{N}}^*(C) = \mathcal{O}(\varepsilon^2)$ and $\mathcal{C}_{\mathcal{N}}^{**}(C) = \mathcal{O}(\varepsilon^3)$. As for the correlation-based approximation [Eq. (10)], we may expect that the spectral radius $\rho(C)$ of C will be an appropriate indicator of the accuracy of approximation (32).

The approximation may be expressed directly in terms of the connectivity matrix as a decomposition into sums over distinct indices:

$$\mathcal{C}_{\mathcal{N}}^*(C) = \frac{n+1}{48} \sum_{i \neq j} (C_{ij}^2 + C_{ij} C_{ji}), \quad (35)$$

$$\begin{aligned} \mathcal{C}_{\mathcal{N}}^{**}(C) &= \frac{n+1}{96} \sum_{i \neq j \neq k} (3C_{ij} C_{jk} C_{ik} + C_{ij} C_{jk} C_{ki}) \\ &+ \frac{n+1}{24} \sum_{i \neq j} C_{ii} (C_{ij}^2 + C_{ij} C_{ji}). \end{aligned} \quad (36)$$

Equations (35) and (36) have much to tell us about the relationship between dynamical complexity as measured by $\mathcal{C}_{\mathcal{N}}$ and structural connectivity as specified by the connection matrix C . First, they provide an insight into the *scaling* of neural complexity with connectivity (cf. Sec. VII). At a more

detailed level we may break down the equations as follows: the sum over C_{ij}^2 in Eq. (35) represents the overall connectivity strength (excluding self-connection). The $C_{ij} C_{ji}$ summation represents an average over *reciprocal connections*. The first two summations in Eq. (36) represent averages over two types of directed *three cycles* while the following two summations represent averages over connectivity strength and reciprocal connections modulated by self-connection. As for the correlation-based version [Eq. (10)], we could in principle calculate the approximation to successively higher order. The picture that emerges is that, as the coupling strength of the system [as measured by $\rho(C)$] is increased, successively higher-order cyclic motifs begin to exert an influence on neural complexity. In [39] we follow this line in more detail; there, we take Eqs. (35) and (36) as the starting point for a more thorough exploration of the *graph-theoretic* implications for neural complexity (cf. [10], although there the authors consider a discrete time process which we have argued may not be suitable for neural modeling; see also Sec. VI below).

VI. IMPACT OF THE COVARIANCE MATRIX CALCULATION ERROR

If we expand the erroneous formula (13) of [1] for the covariance matrix as a power series in C analogous to Eq. (23), we obtain

$$\Omega = I + C^\top + C + (C^\top)^\top + C^\top C + C^2 + \dots \quad (37)$$

This yields a corresponding first-order approximation exactly twice that of our value of $\mathcal{C}_{\mathcal{N}}^*(C)$ (the corresponding second-order approximation will differ more markedly). Thus, somewhat fortuitously, results based on the erroneous calculation will be, at least to a first approximation, consonant with those obtained under our derivation. The error might, on this basis, be said to have had—notwithstanding its quite extensive propagation through the literature [3,6,12]—a reduced impact.

We remarked earlier that it is not clear from [1] exactly what form of Gaussian process was intended there. If, as their formula (14) suggests, it was the *discrete time* multivariate Gaussian AR(1) process

$$\mathbf{X}(t+1) = \mathbf{X}(t) \cdot C + \mathbf{R}(t), \quad (38)$$

then a parallel analysis to that in Secs. IV and V above yields the equation $\Omega = C^\top \Omega C + I$ [cf. Eq. (20)] for the corresponding covariance matrix which leads to the series expansion

$$\Omega = I + C^\top C + (C^\top)^\top C^2 + \dots, \quad (39)$$

and the approximation $\mathcal{C}_{\mathcal{N}} = \mathcal{C}_{\mathcal{N}}^*(C) + \mathcal{O}(\varepsilon^6)$, where

$$\begin{aligned} \mathcal{C}_{\mathcal{N}}^*(C) &= \frac{n+1}{24} \sum_{i \neq j \neq k \neq l} C_{ki} C_{kj} C_{li} C_{lj} \\ &+ \frac{n+1}{24} \sum_{i \neq j \neq k} (4C_{ii} C_{ij} C_{ki} C_{kj} + C_{ki}^2 C_{kj}^2) \\ &+ \frac{n+1}{12} \sum_{i \neq j} (C_{ii}^2 C_{ji}^2 + C_{ii} C_{jj} C_{ij} C_{ji}) \end{aligned} \quad (40)$$

is $\mathcal{O}(\epsilon^4)$ [cf. our Eqs. (35) and (36)]. This is evidently a very different measure from our continuous-time version and does not appear to chime with the intuition of the authors of [1]. In particular, it suggests a strong role for self-connections (rather than reciprocal), and does not, we think, reflect the ideas of the authors of [1] and subsequent papers. We believe that, in contrast, our reformulation of the Gaussian model in continuous time remains true to the spirit of the intuitions underpinning the measure.

VII. CONNECTIVITY NORMALIZATION, AND TIME/ACTIVATION RESCALING

It is clear that neural complexity $\mathcal{C}_{\mathcal{N}}$ will be sensitive to *scaling* of the connection matrix, which relates to the coupling strength of the associated process; that is, if we consider a new connection matrix $\tilde{C} \equiv wC$, where $w > 0$ is a scalar, then the covariance matrix $\tilde{\Omega}$ —and hence neural complexity—depends, in highly nonlinear fashion, on w . This raises some awkward issues as to how we should compare neural complexity across different neural systems: should we, for instance, be prepared to label one system more complex than another if they differ merely in the scaling of the connection weights? This suggests that we might wish to impose some parsimonious “standard” scaling of connectivity; i.e., some form of *normalization* of the connectivity matrix C .

Now we note that connectivity scaling cannot be arbitrary; as mentioned previously, we require condition (24) to guarantee stationarity for Eq. (15). As the infimum of all induced matrix norms, the spectral radius $\rho(C)$ seems a reasonable candidate measure of overall connectivity strength. We shall refer to normalization of C by $w/\rho(C)$, where $0 < w < 1$ is a given scale parameter, as *spectral normalization* [40]. Spectral normalization then guarantees stationarity and also convergence of the covariance matrix series expansion [Eq. (23)] from which our neural complexity approximations derive. From previous remarks and empirical evidence (cf. Secs. III and VIII), we expect that under spectral normalization the scale parameter w will furnish a good predictor of the accuracy of our approximation. Other approaches to normalization include use of a matrix norm such as the Frobenius norm as a measure of overall connectivity strength [4].

But this is not the end of the story. Ultimately, how we choose to normalize should be motivated by physical considerations. We are entitled to ask, for instance, whether spectral normalization makes any physical sense in the context of modeling some “real” (biological) neural system. Tononi *et al.* [1] adopted a different approach. They imposed the constraint that the sum of the afferent synaptic weights per node be equal to a constant value w ; that is, they demand $\sum_j C_{ij} = w$ for every node i ; we refer to this as *Tononi normalization*. Their (unstated) motivation may stem from evidence suggesting that in biological neural systems weights on afferent connections tend to adjust so that, for example, a neuron with few incoming connections will tend to weight those inputs more heavily than a neuron with many incoming connections [41,42]. This approach, however, raises its own set of issues: how are we to deal with *inhibitory* connections

(i.e., negative weights)? We note that under Tononi normalization the scale factor w is always an eigenvalue of C . If all weights are non-negative then it is also the eigenvalue with largest modulus so that $\rho(C) \equiv w$ and stationarity is guaranteed. However, this may not hold true in the presence of inhibitory connections and it is unclear how we should then proceed.

A further subtlety is the following: any transformation of the form

$$C \rightarrow \beta C + (1 - \beta)I, \quad (41)$$

with $\beta > 0$ simply multiplies the covariance matrix by the factor $1/\beta$ and thus leaves $\mathcal{C}_{\mathcal{N}}$ invariant. The transformation can be effected by rescaling time according to $t \rightarrow t/\beta$ and then rescaling activations by $\mathbf{X} \rightarrow \mathbf{X}/\sqrt{\beta}$. But the transformation does *not* leave the approximations $\mathcal{C}_{\mathcal{N}}^*(C)$ and $\mathcal{C}_{\mathcal{N}}^{**}(C)$ invariant; rather, they transform as

$$\mathcal{C}_{\mathcal{N}}^*(C) \rightarrow \beta^2 \mathcal{C}_{\mathcal{N}}^*(C), \quad (42)$$

$$\mathcal{C}_{\mathcal{N}}^{**}(C) \rightarrow \beta^3 \mathcal{C}_{\mathcal{N}}^{**}(C) + 2\beta^2(1 - \beta)\mathcal{C}_{\mathcal{N}}^*(C). \quad (43)$$

So transformation (41) is arbitrary; *any* $\beta > 0$ will do. In practice, if β is too large then the spectral radius of C blows up and series (23) fails to converge. This suggests that for any connectivity matrix C there should be some optimal value of β for which $\mathcal{C}_{\mathcal{N}}^*(C)$ [$\mathcal{C}_{\mathcal{N}}^*(C) + \mathcal{C}_{\mathcal{N}}^{**}(C)$] furnishes the *best* approximation for $\mathcal{C}_{\mathcal{N}}$. There is one “natural” β value: precisely that β which yields a transformed C with *zero trace*, namely,

$$\hat{\beta} \equiv \frac{1}{1 - \text{trace}(C)/n}. \quad (44)$$

Since $\text{trace}(C) = \sum_i \lambda_i < n$, where λ_i are the eigenvalues of C and (by assumption) $\Re(\lambda) < 1$ for all eigenvalues [43], it follows that $\hat{\beta} > 0$ always. We may view this *detracing*—i.e., application of Eq. (41) with $\beta = \hat{\beta}$ —as a normalization of the “average decay” of process (15). We have found in practice that detracing appears quite generally to yield the best approximation $\mathcal{C}_{\mathcal{N}}^*(C)$ [$\mathcal{C}_{\mathcal{N}}^*(C) + \mathcal{C}_{\mathcal{N}}^{**}(C)$] for $\mathcal{C}_{\mathcal{N}}$. So far we have not been able to justify this analytically.

As regards neural complexity, we shall consider time/activation rescaling—and detracing in particular—as a purely mathematical device. We take the view that C represents the “true” connectivity and hence that any normalization should be performed *before* rather than after such rescaling.

VIII. EXAMPLE: EXPONENTIAL CONNECTIVITY DECAY

In [1] the authors find a complexity peak at intermediate “covariance decay” levels for covariance matrices of Toeplitz form [38]. Their covariance matrices are constructed in *ad hoc* fashion with Gaussian off-diagonal decay rather than from connection matrices. Here, as a case study, we echo this construction—now explicitly for the continuous-time process [Eq. (15)]—using circulant Toeplitz connection matrices with exponential connectivity decay.

Consider an $n \times n$ connection matrix given (prenormalization) by

$$C_{ij} = c_{|i-j|}, \quad (45)$$

where $c_k = a^{\min(k, n-k)}$ for $k=1, \dots, n-1$ with $0 < a < 1$ as a decay parameter. Thus (excitatory) connection strength diminishes with distance between nodes, which may be thought of as arranged on a circular lattice. The constant diagonal term $c_0 \equiv s$ may be thought of as either self-activation or as a node activation decay parameter (cf. the remarks at the beginning of Sec. IV). It is set independently of connectivity decay a and is permitted to be negative (inhibitory). Previous studies [1,23] suggest that under this scenario neural complexity should be expected to peak at some intermediate value of a between zero and one: for $a \rightarrow 0$ nodes become isolated and the network maximally segregated, whereas for $a \rightarrow 1$ nodes become uniformly connected and the network thus maximally integrated.

We consider the large system limit $n \rightarrow \infty$. Since C is circulant, its eigenvalues $\lambda_m, m=0, 1, \dots, n-1$ may be calculated explicitly [38]. Let $\rho \equiv e^{-2\pi i/n}$, an n th root of unity. Then the eigenvalues, which are all real, may be calculated as

$$\lambda_m = \sum_{k=0}^{n-1} c_k \rho^{mk} \approx s + \frac{2a(\cos \theta_m - a)}{1 - 2a \cos \theta_m + a^2}, \quad (46)$$

where $\theta_m \equiv 2\pi m/n$. For large n we approximate θ_m as a continuous real parameter $0 \leq \theta < 2\pi$ and the λ_m as a function,

$$\lambda(\theta) \equiv s + \frac{2a(\cos \theta - a)}{1 - 2a \cos \theta + a^2}, \quad (47)$$

of θ . It is straightforward to calculate that $\lambda(\theta)$ takes on extremal values at $\theta=0, \pi$ and we have (in the large n limit)

$$\lambda(0) = s + \frac{2a}{1-a}, \quad (48)$$

$$\lambda(\pi) = s - \frac{2a}{1+a}. \quad (49)$$

We always have $\lambda(0) \geq \lambda(\pi)$, so that $\lambda(0)$ is the largest (real) eigenvalue. To identify the eigenvalue with largest absolute value—i.e., the spectral radius $\rho(C)$ —we observe that

$$|\lambda(0)| \geq |\lambda(\pi)| \Leftrightarrow s \geq -\frac{2a^2}{1-a^2}. \quad (50)$$

Noting that since C is symmetric $U \equiv C$ and $V \equiv C^2$, we may calculate from Eq. (33) [Eq. (34)] that in the large system limit $n \rightarrow \infty$, without normalization or scaling

$$\mathcal{C}_{\mathcal{N}}^*(C) \approx \frac{n(n+1)}{12} \frac{a^2}{1-a^2}, \quad (51)$$

$$\mathcal{C}_{\mathcal{N}}^{**}(C) \approx \frac{n(n+1)}{12} \frac{a^2}{1-a^2} \left(2s + 3 \frac{a^2}{1-a^2} \right). \quad (52)$$

$\mathcal{C}_{\mathcal{N}}^*(C)$ must then be multiplied by $(w/\nu)^2$ and $\mathcal{C}_{\mathcal{N}}^{**}(C)$ by $(w/\nu)^3$, where ν is the normalization factor and w the scale

parameter. To detract, we apply transformations (42) and (43) with $\beta \equiv \hat{\beta} = [1 - (w/\nu)s]^{-1}$.

First we consider the case where there is no normalization (i.e., $\nu=1$) but we have a scale parameter $w > 0$. Stationarity condition (24) demands that $\lambda(0) < 1$. This implies $s < 1/w$, in which case we require $0 < a < (\frac{1}{w} - s) / (2 + \frac{1}{w} - s)$. Within this range we see from Eqs. (42) and (43) that for fixed s in the large system limit, for both the first and second approximations there is no complexity peak: neural complexity simply increases monotonically with increasing a . Simulation reveals that this holds as well for finite system size and for the full measure $\mathcal{C}_{\mathcal{N}}$. This is unsurprising: increasing a just boosts the overall connectivity strength.

For spectral normalization, in the large system limit the normalization factor $\nu \equiv \rho(C)$ switches from $|\lambda(0)|$ to $|\lambda(\pi)|$ at the value of a for which the inequality [Eq. (50)] fails. Stationarity condition (24) is automatically satisfied for all parameter values. Figure 5 compares the first approximation $\mathcal{C}_{\mathcal{N}}^*(C)$ under spectral normalization, as calculated from Eq. (33), as well as the large system limit [Eq. (51)] with the exact value of $\mathcal{C}_{\mathcal{N}}$ calculated from Eq. (4), with covariance matrix calculated according to Eq. (22). Figure 6 compares the second approximation $\mathcal{C}_{\mathcal{N}}^*(C) + \mathcal{C}_{\mathcal{N}}^{**}(C)$ as calculated from Eqs. (33) and (34) as well as the large system limit [Eqs. (51) and (52)] with the exact value of $\mathcal{C}_{\mathcal{N}}$. In both cases detracting was applied after normalization. We see that the first approximation is reasonably accurate, the second approximation more so, and both are qualitatively similar to the full measure $\mathcal{C}_{\mathcal{N}}$. The large system limit approximations are generally less accurate for larger values of a . Further simulation confirms that—as might be expected—the accuracy of all approximations deteriorates with increasing scale factor w ; although qualitatively, the approximations continue to reflect the full measure well.

We see that as $a \rightarrow 1$, in the large system limit $\mathcal{C}_{\mathcal{N}}^*(C)$ tends linearly to zero with a , independently of s . If $s \neq 0$, we may calculate that as $a \rightarrow 0$, $\mathcal{C}_{\mathcal{N}}^*(C) \rightarrow 0$ quadratically with a . For $s > 0$ we see the expected complexity peak at an intermediate value of a . For $s < 0$ complexity rises to a peak at the “kink” $a = \sqrt{-s/(2-s)}$, where the spectral radius $\rho(C)$ transitions discontinuously from $|\lambda(0)|$ to $|\lambda(\pi)|$ according to Eq. (50), and then falls away again.

For $s=0$, somewhat surprisingly, we see that there is no intermediate complexity peak; neural complexity simply decreases monotonically from the finite value $\frac{1}{48}n(n+1)w^4$ at $a \rightarrow 0$. Effectively, as $s \rightarrow 0$ from either the positive or negative side, the complexity peak shifts toward $a=0$. However, as alluded to at the beginning of Sec. IV, if we take the view that the diagonal elements of C represent node activation decay parameters, then we could expect that in a more realistic scenario these will vary somewhat among nodes. In this case, simulations indicate that any inhomogeneity in node decay produces a similar effect to a nonzero fixed s parameter as in our analysis; that is, we do indeed (with normalization) see a complexity peak. See, e.g., Fig. 6(a), where for the “ s variable” plot, diagonal elements of C were drawn at

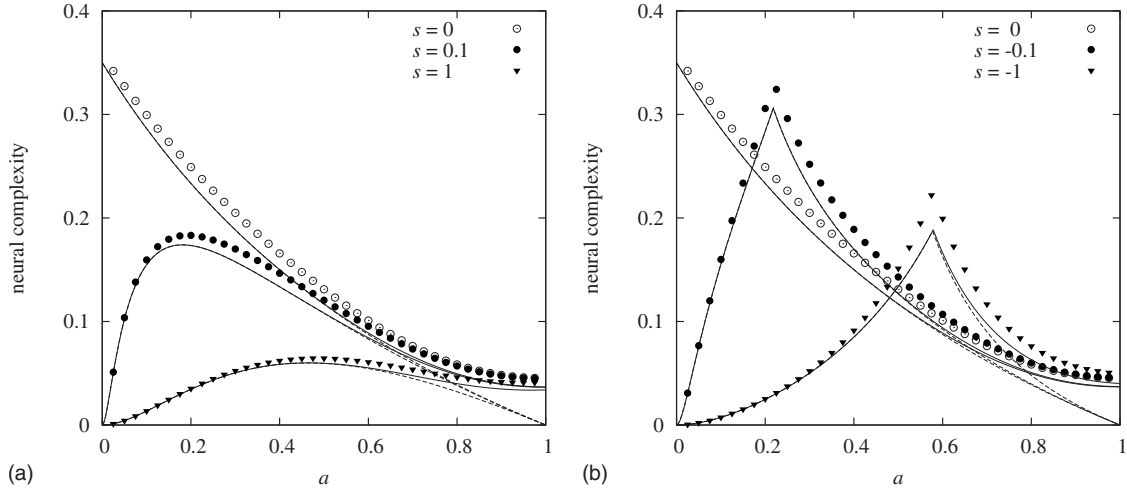


FIG. 5. First approximation: neural complexity under spectral normalization for a ring lattice with exponential connectivity decay, plotted against decay a for several values of the parameter s . Symbols give the exact value C_N of Eq. (4), with covariance matrix calculated according to Eq. (22). Solid lines plot the first approximation $C_N^*(C)$ as calculated from Eq. (35) while dashed lines plot the large system limit [Eq. (51)]. For the approximations, detracing was applied after normalization. Parameters: $n=20$, $w=0.2$. (a) $s \geq 0$ (b) $s \leq 0$.

random from a normal distribution with zero mean and small variance.

Two major points emerge from this exercise: first, it is clear that some form of normalization will be required to produce the characteristic peak in complexity at intermediate connectivity decay. Second, some variation in the node decay parameters—emerges as a crucial factor. Both factors are strongly reinforced in further research by the authors [39]. We note with interest that in the *ad hoc* covariance decay model of [1], while the issue of self-activation/node decay is obscured, it was apparently deemed necessary to add noise to the diagonal of the covariance matrices.

IX. DISCUSSION

The relationship between the dynamic complexity of a neural system as measured by C_N and the structural connectivity of the network underlying that system might hitherto be regarded as somewhat opaque, in the sense that it may not be easily discerned from the actual definition of the measure. In this study, we have developed what we consider to be a viable analytical framework that goes some way toward elucidating this relationship. Our methodology, appropriate to weakly coupled complex neural systems for which a stationary Gaussian approximation represents a valid approach, proceeds as follows: first, we derive an asymptotic series

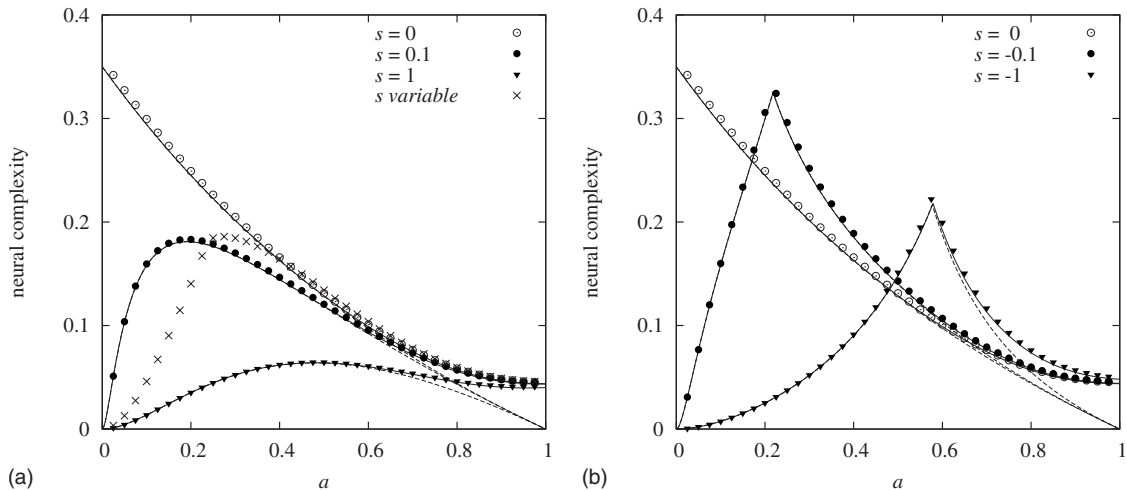


FIG. 6. Second approximation: neural complexity under spectral normalization for a ring lattice with exponential connectivity decay, plotted against decay a for several values of the parameter s (for the s variable plot diagonal elements of C were drawn at random from a normal distribution with zero mean and standard deviation=0.2). Symbols give the exact value C_N of Eq. (4), with covariance matrix calculated according to Eq. (22). Solid lines plot the second approximation $C_N^*(C) + C_N^{**}(C)$ as calculated from [Eqs. (35) and (36)] while dashed lines plot the large system limit [Eqs. (51) and (52)]. For the approximations, detracing was applied after normalization. Parameters: $n=20$, $w=0.2$. (a) $s \geq 0$ (b) $s \leq 0$.

expansion for \mathcal{C}_N in terms of the correlation matrix (and hence the coupling strength) of the process. The first few terms of the expansion are calculated explicitly (in principle the expansion could be extended to higher-order terms). We then introduce, via the continuous-time stochastic process of Sec. IV, a principled basis—in the spirit of statistical mechanics—for a Gaussian model of neural dynamics. The model establishes an explicit link between the structural connectivity of the system as represented by the connection matrix C , and its functional dynamics as represented by the stochastic process $\mathbf{X}(t)$. Reframing our approximation for \mathcal{C}_N in terms of the connection matrix then yields insight into the structure-complexity relationship.

The picture that emerges of this relationship foregrounds the role of cyclic connectivity motifs which, in functional terms, translate into *recurrent* dynamics. With increasing coupling strength, higher-order cycles become successively more significant. Thus for very weakly coupled systems, reciprocal connectivity dominates complexity. As coupling strength increases, three cycles begin to contribute significantly, and so on. Although there is still much work to be done; it remains challenging to impute from the approximation precisely what types of connectivity structure might give rise to high neural complexity under given (physically motivated) constraints. There are suggestions in the literature that, for instance, high complexity may be associated with *small world* network topologies [11,44–47] (but see also [48]) or with modular hierarchical structures [46,49]. We propose that our analytical approach may clarify these issues; this line is pursued in [39] and further work in progress where, taking the results presented here as a starting point, the dependence of neural complexity on the graph topology of a putative network underlying a neural system is analyzed in detail.

Other points of interest—which emerge clearly from the worked example of Sec. VIII—include the crucial role of connectivity scaling/normalization, and the significance (and indeed interpretation) of self-activation and activation decay variance. As a spinoff from our approach, we have already noted that the approximation to \mathcal{C}_N in terms of the correlation matrix furnishes a practical and computationally cheap proxy for the full measure. The approximation is applicable to time-series data, and is not only accurate for weakly coupled systems, but even at higher coupling strengths appears to reflect well the qualitative behavior of \mathcal{C}_N .

Regarding the problematic covariance matrix calculation of [1], our view is that while results obtained under that formulation might be viewed as suspect within the assumptive framework of a discrete neural process, they are nonetheless largely validated by reinterpretation within the framework of an appropriate continuous-time process.

We have expressed caveats over the ambit of stationary linear Gaussian models as deployed in this and other studies. Much work is required—not least empirical—to identify more precisely the limitations of the approach and, ultimately, to develop viable analytical techniques for situations where such models turn out to be inappropriate.

Finally, we are hopeful that the mathematical tools developed in this study might have wider application to

information-theoretic complexity analyses; we are actively pursuing several such applications in our current research.

ACKNOWLEDGMENTS

The authors would like to thank Anil Seth, Annamaria Cucinotta, and Srinandan Dasmahapatra for useful discussions. This study was supported by the Spatially Embedded Complex Systems Engineering (SECSE) project, EPSRC-GB (U.K.) Grant No. EP/C51632X/1.

APPENDIX A: EXACT SIMULATION OF THE MULTIVARIATE ORNSTEIN-UHLENBECK PROCESS

Process (15) may be simulated *exactly* [rather than, e.g., by naive iteration of Eq. (16) with a small finite time step dt] as follows: we may calculate that, *conditional on* $\mathbf{X}(t)=\mathbf{x}$, the expected value of $\mathbf{X}(t+s)$ is given by $\mathbf{x} \exp[-(I-C)s]$ while the covariance matrix $M(s)$ satisfies the matrix differential equation

$$\frac{dM}{ds} = I + C^T M + M C - 2M, \quad (\text{A1})$$

with initial condition $M(0)=0$ [comparing with Eq. (24), we may check that as $s \rightarrow \infty M(s)$ tends toward the stationary covariance matrix Ω]. For a *univariate* Ornstein-Uhlenbeck process this equation may be solved explicitly [50]; in the multivariate case it may be solved numerically. We can also solve it explicitly in the case that C is symmetric, as

$$M(s) = \frac{1}{2}(I-C)^{-1}\{I - \exp[-2(I-C)s]\}. \quad (\text{A2})$$

Now since we know that $\mathbf{X}(t+s)$ is always multivariate Gaussian, and we know that its mean and covariance matrix conditional on $\mathbf{X}(t)=\mathbf{x}$, given a past state of the process we thus know the precise distribution of the state at an arbitrary time s in the future. Specifically, given a previous value \mathbf{x}_t for the state of the process, we may generate a subsequent state \mathbf{x}_{t+s} as a deviate drawn from a multivariate normal distribution with mean $\mathbf{x}_t \exp[-(I-C)s]$ and covariance matrix $M(s)$. This may be achieved by a Cholesky decomposition [27] of $M(s)$ applied to uncorrelated unit-variance Gaussian noise (note that s need not be small). MATLAB® code implementing the simulation may be found in [22].

APPENDIX B: AVERAGING THE TRACE OVER SUBSYSTEMS

Let M be a symmetric $n \times n$ matrix that is zero on the diagonal. Below, in summing over k -subsystem indices \mathbf{i} , for mathematical convenience we count all permutations of the i_α so that the number of indices (still distinct but now unordered) is $\frac{n!}{(n-k)!}$ rather than $\frac{n!}{k!(n-k)!}$. We have then

$$\begin{aligned} \frac{n!}{(n-k)!} \langle \text{trace}(M^2) \rangle_k &= \sum_{\mathbf{i}} \text{trace}[M(\mathbf{i})^2] \\ &= \sum_{i_1 \neq \dots \neq i_k} \sum_{\alpha, \beta} M_{i_\alpha i_\beta}^2 \\ &= \sum_{i_1 \neq \dots \neq i_k} \sum_{\alpha \neq \beta} M_{i_\alpha i_\beta}^2 \end{aligned}$$

$$\begin{aligned}
&= k(k-1) \sum_{i_1 \neq \dots \neq i_k} M_{i_1 i_2}^2 \\
&= k(k-1) \frac{(n-2)!}{(n-k)!} \sum_{i \neq j} M_{ij}^2 \\
&= k(k-1) \frac{(n-2)!}{(n-k)!} \text{trace}(M^2),
\end{aligned}$$

so that

$$\langle \text{trace}(M^2) \rangle_k = \frac{k(k-1)}{n(n-1)} \text{trace}(M^2). \quad (\text{B1})$$

Next, we have

$$\begin{aligned}
\frac{n!}{(n-k)!} \langle \text{trace}(M^3) \rangle_k &= \sum_i \text{trace}[M(i)^3] \\
&= \sum_{i_1 \neq \dots \neq i_k} \sum_{\alpha, \beta, \gamma} M_{i_\alpha i_\beta} M_{i_\beta i_\gamma} M_{i_\gamma i_\alpha} \\
&= \sum_{i_1 \neq \dots \neq i_k} \sum_{\alpha \neq \beta \neq \gamma} M_{i_\alpha i_\beta} M_{i_\beta i_\gamma} M_{i_\gamma i_\alpha} \\
&= k(k-1)(k-2) \sum_{i_1 \neq \dots \neq i_k} M_{i_1 i_2} M_{i_2 i_3} M_{i_3 i_1} \\
&= k(k-1)(k-2) \frac{(n-3)!}{(n-k)!} \sum_{i \neq j \neq l} M_{ij} M_{jl} M_{li} \\
&= k(k-1)(k-2) \frac{(n-3)!}{(n-k)!} \text{trace}(M^3),
\end{aligned}$$

so that

$$\langle \text{trace}(M^3) \rangle_k = \frac{k(k-1)(k-2)}{n(n-1)(n-2)} \text{trace}(M^3). \quad (\text{B2})$$

-
- [1] G. Tononi, O. Sporns, and G. M. Edelman, Proc. Natl. Acad. Sci. U.S.A. **91**, 5033 (1994).
- [2] G. Tononi, G. M. Edelman, and O. Sporns, Trends Cogn. Sci. **2**, 474 (1998).
- [3] A. K. Seth and G. M. Edelman, Adapt. Behav. **12**, 5 (2004).
- [4] K. J. Friston, G. Tononi, O. Sporns, and G. M. Edelman, Hum. Brain Mapp. **3**, 302 (1995).
- [5] C. J. Stam, Clin. Neurophysiol. **116**, 2266 (2005).
- [6] O. Sporns, G. Tononi, and G. M. Edelman, Neural Networks **13**, 909 (2000).
- [7] G. Tononi and G. M. Edelman, Science **282**, 1846 (1998).
- [8] A. K. Seth, E. M. Izhikevich, G. N. Reeke, and G. M. Edelman, Proc. Natl. Acad. Sci. U.S.A. **103**, 10799 (2006).
- [9] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
- [10] M. De Lucia, M. Bottaccio, M. Montuori, and L. Pietronero, Phys. Rev. E **71**, 016114 (2005).
- [11] M. P. Shanahan, Phys. Rev. E **78**, 041924 (2008).
- [12] O. Sporns, G. Tononi, and G. M. Edelman, Cereb. Cortex **10**, 127 (2000).
- [13] A. K. Seth, B. J. Baars, and G. M. Edelman, Conscious Cogn **14**, 119 (2005).
- [14] B. J. Baars, T. Ramsøy, and S. Laureys, Trends Neurosci. **26**, 671 (2003).
- [15] We remark that the measure has been extended to incorporate explicit extrinsic inputs [12].
- [16] H. R. Wilson and J. D. Cowan, Biophys. J. **12**, 1 (1972).
- [17] P. L. Nunez, Math. Biosci. **21**, 279 (1974).
- [18] W. J. Freeman, Electroencephalogr. Clin. Neurophysiol. **34**, 9 (1978).
- [19] A. R. McIntosh and F. Gonzalez-Lima, Hum. Brain Mapp. **2**, 2 (1994).
- [20] O. Sporns and G. Tononi, Complexity **7**, 28 (2002).
- [21] P. L. Nunez and R. Srinivasan, Clin. Neurophysiol. **117**, 2424 (2006).
- [22] L. Barnett, C. L. Buckley, and S. Bullock, MATLAB® routines for neural complexity analysis (2009), http://www.secse.net/ncomp/ncomp_tools.zip.
- [23] G. Tononi, O. Sporns, and G. M. Edelman, Proc. Natl. Acad. Sci. U.S.A. **93**, 3422 (1996).
- [24] G. Tononi, O. Sporns, and G. M. Edelman, Proc. Natl. Acad. Sci. U.S.A. **96**, 3257 (1999).
- [25] In this paper we consider $\mathbf{X}(t)$ as a row vector. The “ \top ” symbol denotes matrix transpose.
- [26] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley-Interscience, New York, 1991).
- [27] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, New York, 1985).
- [28] We note that any positive-definite matrix can be a covariance matrix and that, furthermore, any positive-definite matrix may be written as the *Gram matrix* $X^\top X$ of a set of vectors X [27].
- [29] The C and \mathbf{X} appear to be incorrectly reversed on the right-hand side of the equivalent equation in [1].
- [30] G. E. Uhlenbeck and L. S. Ornstein, Phys. Rev. **36**, 823 (1930).
- [31] B. K. Øksendal, *Stochastic Differential Equations: An Introduction With Applications* (Springer, Berlin, 1998).
- [32] H. R. Wilson and J. D. Cowan, Kybernetika **13**, 55 (1973).
- [33] R. F. Galán, PLoS One **3**, e2148 (2008).
- [34] R. D. Beer, Adapt. Behav. **3**, 469 (1995).
- [35] R. D. Beer, Neural Comput. **18**, 3009 (2006).
- [36] W. C. Schieve, A. R. Bulsara, and G. M. Davis, Phys. Rev. A **43**, 2613 (1991).
- [37] We may use Gelfand’s theorem with any *consistent matrix norm* [38].
- [38] R. M. Gray, *Toeplitz and Circulant Matrices: A Review* (Now, Boston, 2006).
- [39] L. Barnett, C. L. Buckley, and S. Bullock, *On a Graph Theoretic Interpretation of Neural Complexity* (unpublished).
- [40] Note that this is not the same as normalization by the *spectral norm* of C , which is defined as the square root of the maximum eigenvalue of $C^\top C$ [27].

- [41] G. Turrigiano, *Curr. Opin. Neurobiol.* **17**, 318 (2007).
- [42] I. Rabinowitch and I. Segev, *Trends Neurosci.* **31**, 377 (2008).
- [43] Note that for *complex* eigenvalues this still follows since complex eigenvalues always come in conjugate pairs.
- [44] D. J. Watts and S. H. Strogatz, *Nature (London)* **393**, 440 (1998).
- [45] O. Sporns, G. Tononi, and G. M. Edelman, *Behav. Brain Res.* **135**, 69 (2002).
- [46] O. Sporns, *BioSystems* **85**, 55 (2006).
- [47] C. C. Hilgetag, G. A. Burns, M. A. O'Neil, J. W. Scannel, and M. P. Young, *Philos. Trans. R. Soc. London, Ser. B* **355**, 91 (2000).
- [48] C. L. Buckley and S. Bullock, in *Proceedings of the 9th European Conference on Artificial Life*, edited by F. Almeida e Costa (Springer, Berlin, 2007).; URL: <http://eprints.ecs.soton.ac.uk/14282/>.
- [49] C. C. Hilgetag, M. A. O'Neil, and M. P. Young, *Philos. Trans. R. Soc. London, Ser. B* **355**, 71 (2000).
- [50] D. T. Gillespie, *Phys. Rev. E* **54**, 2084 (1996).