Interrelating anatomical, effective, and functional brain connectivity using propagators and neural field theory

P. A. Robinson[1,2]

[1]*School of Physics, University of Sydney, New South Wales 2006, Australia*

[2]*Brain Dynamics Center, Sydney Medical School - Western, University of Sydney, Westmead, New South Wales 2145, Australia*

(Received 27 September 2011; revised manuscript received 9 December 2011; published 17 January 2012)

It is shown how to compute effective and functional connection matrices (eCMs and fCMs) from anatomical CMs (aCMs) and corresponding strength-of-connection matrices (sCMs) using propagator methods in which neural interactions play the role of scatterings. This analysis demonstrates how network effects dress the bare propagators (the sCMs) to yield effective propagators (the eCMs) that can be used to compute the covariances customarily used to define fCMs. The results incorporate excitatory and inhibitory connections, multiple structures and populations, asymmetries, time delays, and measurement effects. They can also be postprocessed in the same manner as experimental measurements for direct comparison with data and thereby give insights into the role of coarse-graining, thresholding, and other effects in determining the structure of CMs. The spatiotemporal results show how to generalize CMs to include time delays and how natural network modes give rise to long-range coherence at resonant frequencies. The results are demonstrated using tractable analytic cases via neural field theory of cortical and corticothalamic systems. These also demonstrate close connections between the structure of CMs and proximity to critical points of the system, highlight the importance of indirect links between brain regions and raise the possibility of imaging specific levels of indirect connectivity. Aside from the results presented explicitly here, the expression of the connections among aCMs, sCMs, eCMs, and fCMs in terms of propagators opens the way for propagator theory to be further applied to analysis of connectivity.

DOI: 10.1103/PhysRevE.85.011912 PACS number(s): 87.10.−e, 87.18.Sn, 89.75.Fb, 87.61.Qr

I. INTRODUCTION

There is enormous interest in quantifying the static and dynamic connectivity of the brain to better understand how it processes inputs and performs tasks rapidly, how it maintains stability while preventing the spread of undesirable activity such as seizures, how it develops, and how it responds to damage [1–23]. These and related questions provide much of the motivation for the current human connectome project [10,17], for example. One approach to quantification is to determine anatomical connection matrices (aCMs, often called structural CMs), which summarize the known anatomical connectivity between pairs of brain regions. Increasingly, analyses include numbers of connections (e.g., as measured by numbers of axons) [1,2,10] and it is recognized that the corresponding strength-of-connection matrix (sCM) must be considered, which includes measures of synaptic strength, which can be time-varying [1]. The sCM is sometimes called a gain matrix [14,15] or an effective connectivity matrix but includes only direct effects of one region on another. Here we reserve the term effective CM (eCM) for the more general matrix that quantifies the neural effect of one region on another, whether or not they are directly connected [1,2,10,24]. In the case of direct connections only, the eCM reduces to the sCM, which is the main case that has been discussed in the literature. Functional magnetic resonance imaging (fMRI) provides information on functional connectivity, which is usually determined via correlations or covariances of activity between regions and is summarized in functional CMs (fCMs) [1,2,7,10,11,25–28].

Anatomical, strength-of-connection, effective, and functional CMs have many similarities, as illustrated in Fig. 1 for an aCM and corresponding fCM [11]; in particular, most of the strongest connections are common to both matrices. However, fCMs are fuller and many questions remain regarding how they are related [2,10,27–29], including (i) to what extent are the eCM and fCM determined by the aCM or sCM, and vice versa, especially as fCMs are usually defined to be covariance matrices of activity relative to baseline in different regions, and, thus, can have negative entries, and both eCMs and fCMs tend to have more nonzero entries than aCMs or sCMs? (ii) How does thresholding of connections (e.g., to eliminate those below a certain level) affect the appearance and relationships between CMs? (iii) How do spatial coarse-graining of brain regions, and temporal averaging of covariances, affect the CMs, especially with regard to their sparseness? (iv) How does one include inhibitory connections in CMs, which are usually restricted to summarizing relatively long-range excitatory connections? (v) How do local short-range connections more generally affect CMs? (vi) How are time delays and directionality in connections to be included in connectivity analysis, especially in order to reflect and determine functional causality? (vii) Is connectivity frequency dependent? (viii) What is the effect of brain stability and approach to instability (criticality) on CMs? Many of these issues are also relevant to other types of complex networks [21,22].

Approaches to analyzing structural and functional connectivity have multiplied in recent years, with numerous suggestions for how to proceed. Most methods involve systematic summary measures of overall network structure (clustering, path length, centrality, etc.) [2,5,10,29] or even *ad hoc* or phenomenological analyses of CMs. What is needed is a systematic theory that will enable aCMs, sCMs, eCMs, and fCMs measured by various means to be related both to each other, to summary measures, and to the underlying anatomy, physiology, neural activity, and geometry. Some efforts in this

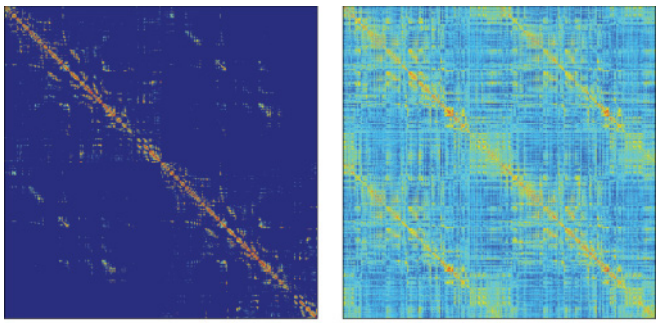

FIG. 1. (Color online) Examples of (a) anatomical and (b) functional connectivity matrices. Each of the 998 nodes corresponds to one row and one column. Entries in the matrix denote the strength of connections between nodes, with the strongest connections near the main diagonal and two other prominent diagonals that correspond to interhemispheric connections between homologous regions. Adapted from Ref. [11] with permission.

direction have been made, with initial success in relating key features of fCMs to underlying aCMs in specific examples [1,3–5,7,11,12,25–27,30], often using highly simplified models of neural dynamics. The core idea is that eCMs incorporate the combined effects of direct and indirect neural influences between regions, whereas sCMs include only neural influences via direct anatomical connections that are activated under the conditions of interest. A related point is that such effects as neuromodulation, long-term potentiation, and plasticity can change the strengths of specific anatomical connections dynamically so they do not contribute to functional connectivity under all conditions. Hence, eCMs and fCMs are expected to depend on the time scales over which they are observed, an effect that has been seen experimentally [2,7].

Here, methods from propagator (Green's function) theory are used to help unravel direct and indirect contributions to connectivity between brain regions, thereby enabling aCMs, sCMs, eCMs, and fCMs to be interrelated. Before proceeding to describe the approach it is essential to forestall possible misunderstanding by stressing that, despite mathematical links, the methods employed here *do not* correspond to any kind of quantum theory of the brain (except insofar as spikes can be viewed as quanta of activity) but to a systematic calculational approach that has its roots in long-proven field-theoretic methods of theoretical physics. These parallels are discussed in Secs. II and VII.

As demonstrated below, propagator methods prove to be well suited to disentangling direct and indirect influences between regions in a systematic way. They help to clarify network effects and enable inhibitory connections, multiple neural types, temporal delays, and measurement/postprocessing effects to be included in ways that do not depend on the specific models of interactions and neural dynamics chosen to evaluate the results in specific cases. In this approach, propagation of spikes between neurons, or spike-rate fields between neural populations in a mean-field approximation, is analogous to propagation of particles and the fields representing them, respectively, in quantum theory [31–36]. Likewise, spike-induced synaptic dynamics that couples neurons is analogous to scattering. Hence, sCMs correspond to direct (or *bare*) propagation between regions without interaction along the way, whereas eCMs incorporate all intermediate interactions (e.g., polysynaptic transmission in neuroscience and multiple scatterings in field theory) and, thus, correspond to *dressed* or *renormalized* propagators.

One approach to dealing with large assemblies of interacting neurons, such as those most relevant to the interactions between the macroscopic regions of interest used in determining CMs, is to replace individual spikes by a mean spike rate for each neuron or even to average this spike rate over many neurons to yield a population-based rate that is a function of position and time [34–71]. In such neural field theories (NFTs), mean firing rates, soma potentials, and other neural properties are averaged over spatial scales of tenths of a mm to yield a mean-field description of neural activity and related statistical theories. NFTs have been expressed in propagator form [40,42] and examples of these are, thus, used in Secs. V and VI to obtain analytical results. Perhaps most significantly, NFTs exhibit emergent activity properties that are not found in single neurons (e.g., collective modes and oscillations, eigenfunctions, resonances, and criticality).

The structure of this paper is as follows: Sections II and III describe a form of propagator theory sufficiently general to encompass both spike-based and neural field interactions between brain regions and show how sCMs, eCMs, and fCMs are interrelated. In Sec. IV we discuss how to include measurement and signal-processing effects in both theoretical and experimental CMs. Section V then writes the results in terms of neural field theory and Sec. VI illustrates how to apply them to specific systems and cases of CMs. This enables us to unravel structure-function relationships in systems that involve excitation, inhibition, asymmetry, temporal delays, resonances, propagating waves, and criticality.

II. GENERAL THEORY

In this section and the next two, we derive relationships between the propagation of neural activity and structural, effective, and functional CMs. We also clarify the connections between discrete matrix-based and continuous operator-based propagator formulations of neural dynamics and between purely spatial and spatiotemporal connectivities. Measurement and signal-processing issues are briefly outlined in Sec. IV.

A. Propagator formulation of neural interactions

We are interested in connectivity involving p distinct neural populations, each of which is potentially spatially distributed. We distinguish these populations by a subscript $a = 1,2,\ldots,p$. Each produces spikes at an instantaneous local rate $Q_a(\mathbf{r},t) = Q_a(x)$, where x denotes the vector consisting of position $\mathbf{r}$ and time t; here this is simply a notational convenience that carries no implications of relativistic effects, so we do not multiply t by a velocity to make the dimensions the same for all components, although this could be done without difficulty. We note that (i) either spiking or averaged (i.e., neural field) interactions between neurons can be incorporated using this approach, by viewing Q_a as a discretely (e.g., sum of δ functions) or continuously varying quantity, respectively [37,50,51]. In the latter case there is an implicit average over

a large number of neurons. (ii) When working with the cortex, which can be approximated as a folded two-dimensional (2D) sheet for many purposes, $\mathbf{r}$ can be treated as a 2D vector, an approach we use in Sec. III.

An important point here is that Q_a can be used to represent not just a mean firing rate but also a perturbation to that rate. This is because any rate or rate perturbation, however small, suffices, in principle, to establish a link between two neural regions—and, hence, is sufficient to trace a CM. Indeed, fCMs are usually computed by calculating correlations or covariances of small differences from baseline activity [1,2,10,11]. Hence, we concentrate on perturbations from baseline in the present work and restrict attention to linear approximations. Similar connectivity results can be expected to follow in nonlinear cases, but the dynamics linking incoming and outgoing activity at specific neural locations will be more complex. None of these points restricts fine-scale dynamics (e.g., generation of action potentials and pattern formation) to be linear.

Since spikes in neurons a are caused by neural inputs (from various afferent populations, which we label $b = 1, \ldots, p$ and external sources N_a), we can write

$$Q_a(\mathbf{r},t) = \sum_b \int\int \Lambda^{(0)}_{ab}(\mathbf{r},t,\mathbf{r}',t')Q_b(\mathbf{r}',t')d\mathbf{r}'dt' + N_a(\mathbf{r},t), \tag{1}$$

where $\Lambda^{(0)}_{ab}(\mathbf{r},t,\mathbf{r}',t') = 0$ for $t < t'$ to preserve causality, the integrals extend over all locations and times, and the sum is over all populations, including a. If position and/or time are discretized, as is usually the case in CM analyses, the integrals in Eq. (1) are replaced by sums over discrete values of x and t. The propagator $\Lambda^{(0)}_{ab}$ in Eq. (1) quantifies the activity evoked in population a at location $\mathbf{r},t$ by activity directly afferent from neurons of population b at $\mathbf{r}',t'$, as illustrated in Fig. 2, and includes (as we see in detail in Sec. III) both spatial and temporal variations of this connection strength in general. This formulation also permits inhibitory couplings to be included simply by allowing $\Lambda^{(0)}_{ab}$ to be negative if population b is inhibitory. It should be stressed that $\Lambda^{(0)}_{ab}$ is certainly not a δ function in space (neurons in a small region project to many others at various distances); nor is it a δ function in time, since signals between two small regions will take paths with a variety of different delays, due to variance of the detailed properties of the axons linking them, and synaptic and dendritic delays broaden signals in time.

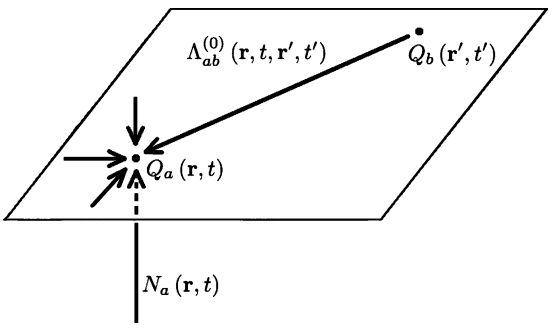

FIG. 2. Schematic of the terms in Eq. (1), showing contributions to activity $Q_a(\mathbf{r},t)$, including the local contribution $N_a(\mathbf{r},t)$ that results from external inputs, and the contribution that results from spikes propagating directly from $Q_b(\mathbf{r}',t')$.

By viewing the $Q_a(\mathbf{r},t)$ and $N_a(\mathbf{r},t)$ as elements of p-element column vectors, one element for each population, and the $\Lambda^{(0)}_{ab}(x,x')$ as elements of a $p \times p$ square matrix, we can write Eq. (1) in the more compact matrix notation

$$\mathbf{Q}(x) = \int \mathbf{\Lambda}^{(0)}(x,x')\mathbf{Q}(x')dx' + \mathbf{N}(x). \tag{2}$$

In this notation the column vector $\mathbf{Q}$ is a column matrix whose elements are the fields $Q_1(x), Q_2(x), \ldots, Q_p(x)$.

B. Matrix representation

If we discretize the space and time coordinates, we can replace the population label a and spatiotemporal coordinates x by a collective label m and the set b,x' by n. One then can make the elementwise identification

$$\Lambda^{(0)}_{ab}(x,x') = \Lambda^{(0)}_{mn}, \tag{3}$$

which puts $\mathbf{\Lambda}^{(0)}$ in the form of a CM that has been extended to include temporal delays, akin to the one introduced in Ref. [25]. The labeling of $\mathbf{Q}$ is correspondingly extended to designate not just population but also location and time, with $\mathbf{Q}$ becoming the column matrix with elements $Q_1(x_1), \ldots, Q_1(x_M), Q_2(x_1), \ldots, Q_p(x_M)$, where $M = n_r n_t$, and n_r and n_t are the numbers of points into which position and time are discretized, respectively.

In matrix notation (2) becomes simply

$$\mathbf{Q} = \mathbf{\Lambda}^{(0)}\mathbf{Q} + \mathbf{N}, \tag{4}$$

where the integrals over $\mathbf{r}'$ and t' have been replaced by the sums over elements that are inherent in matrix multiplication. Note that we do not normally show coordinates as arguments when we use this notation; instead labels designate matrix elements, as in the right side of Eq. (3). This multiplication can be made to approximate the integrations in Eq. (2) as accurately as one wishes by making the spatiotemporal divisions sufficiently fine. The price paid in obtaining Eq. (4) from Eq. (2) is that the matrices $\mathbf{Q}$ and $\mathbf{N}$ now have $P = pn_r n_t$ elements. Similarly, $\mathbf{\Lambda}^{(0)}$ is a $P \times P$ matrix.

From Eq. (4) one obtains the standard result for the response of a linear system to an external stimulus $\mathbf{N}$:

$$\mathbf{Q} = [\mathbf{I} - \mathbf{\Lambda}^{(0)}]^{-1}\mathbf{N}, \tag{5}$$

$$= \mathbf{T}\mathbf{N}, \tag{6}$$

$$= (\mathbf{I} + \mathbf{\Lambda})\mathbf{N}, \tag{7}$$

where $\mathbf{I}$ is the unit matrix and the superscript -1 denotes the matrix inverse. Equation (6) defines the transfer matrix $\mathbf{T}$ that links activity $\mathbf{Q}$ to inputs $\mathbf{N}$. This interpretation then sets the correct normalization for $\Lambda^{(0)}_{ab}$ in Eq. (1), which is an important point because experimental CMs are usually normalized in an arbitrary fashion. Equation (7) defines the matrix $\mathbf{\Lambda}$, which we identify as the eCM below. In the case of purely spatial neural couplings (i.e., neglecting time delays), Eq. (5) has been stated before [23]; the present result includes temporal aspects and places the result in a form that can be used to implement results from propagator theory. Significantly, activity is only linearly

stable if all the eigenvalues of $\mathbf{T}$ have negative real parts. Large-scale stability is required to avoid such outcomes as epileptic seizures [6,37,43,44]. We stress that this requirement does not rule out localized or transient instabilities and nonlinear dynamics.

Equations (5)–(7) enable the identification of sCMs and the calculation of their connection to eCMs and fCMs, along with other properties of neural activity and its correlations. In Eq. (7), $\mathbf{\Lambda}$ is the dressed propagator, while $\mathbf{\Lambda}^{(0)}$ is the bare propagator—i.e., $\mathbf{\Lambda}$ includes all propagation routes to x from other points x' (i.e., all network effects, including propagation within and between populations and via intermediate node), whereas $\mathbf{\Lambda}^{(0)}$ incorporates only direct propagation. We can, thus, identify $\mathbf{\Lambda}^{(0)}$ with the sCM and define $\mathbf{\Lambda}$ to be the eCM.

Equations (5)–(7) yield the eCM in terms of the sCM, while inversion gives

$$\mathbf{\Lambda}^{(0)} = \mathbf{I} - \mathbf{T}^{-1} = \mathbf{I} - (\mathbf{I} + \mathbf{\Lambda})^{-1}. \tag{8}$$

We can also expand Eq. (5) in powers of $\Lambda^{(0)}$ to give

$$\mathbf{Q} = [\mathbf{I} + \mathbf{\Lambda}^{(0)} + \{\mathbf{\Lambda}^{(0)}\}^2 + \{\mathbf{\Lambda}^{(0)}\}^3 + \cdots]\mathbf{N}, \tag{9}$$

which converges provided all the eigenvalues $\lambda_j^{(0)}$ of $\mathbf{\Lambda}^{(0)}$ satisfy $|\lambda_j^{(0)}| < 1$, which is guaranteed whenever the system is linearly stable. Comparison with Eq. (7) immediately yields

$$\mathbf{\Lambda} = \mathbf{\Lambda}^{(0)} + \{\mathbf{\Lambda}^{(0)}\}^2 + \{\mathbf{\Lambda}^{(0)}\}^3 + \cdots. \tag{10}$$

Conversely, one finds

$$\mathbf{\Lambda}^{(0)} = \mathbf{\Lambda} - \mathbf{\Lambda}^2 + \mathbf{\Lambda}^3 - \cdots, \tag{11}$$

which converges under the same conditions as Eq. (9).

Taking the terms on the right of Eq. (9) in order, this equation expresses activity $\mathbf{Q}$ as the sum of different interactions, with each term represented by a Feynman diagram, as shown in Fig. 3: the term $\mathbf{IN}$ expresses activity directly evoked by the external inputs, the term $\mathbf{\Lambda}^{(0)}\mathbf{N}$ incorporates activity evoked by direct connections from externally stimulated neurons, the term $[\mathbf{\Lambda}^{(0)}]^2\mathbf{N}$ represents activity evoked via one intermediate node (possibly in a different population), the term $[\mathbf{\Lambda}^{(0)}]^3\mathbf{N}$ represents cases with two intermediate nodes, and so on. These aspects are explicitly illustrated in Sec. V to make them more concrete and explore their effects.

The form of Eq. (9) can also be written as

$$\mathbf{Q} = [\mathbf{I} + \mathbf{\Lambda}^{(0)}(\mathbf{I} + \mathbf{\Lambda}^{(0)} + \{\mathbf{\Lambda}^{(0)}\}^2 + \cdots)]\mathbf{N}, \tag{12}$$

which yields Eq. (4) on identifying the contents of the parentheses (multiplied by $\mathbf{N}$) as being the value of $\mathbf{Q}$ at the location x'' of the last node before the final point (the matrix formulation automatically allows for all places and populations where that node might be). This equation thus states that propagation to x from x' can be viewed as the net effect of dressed propagation to all possible x'' from x' followed by bare propagation to x from those x'', as shown in Fig. 4. This is the exact analog of a standard result in quantum field theory [31–33].

C. Integral representation

Before proceeding further, it is worth noting that Eqs. (9)–(11) can also be written in integral form analogous to (1) and (2) by using the matrix-field formulation of Eq. (2). This replaces each matrix multiplication by a spatiotemporal integration and a sum over populations; e.g.,

$$\mathbf{\Lambda}^{(0)}\mathbf{N} \longleftrightarrow \sum_b \int \Lambda_{ab}^{(0)}(x,x')N_b(x')dx', \tag{13}$$

with multiple integrals involved in higher-order terms.

In the above analysis we saw that the propagator $\mathbf{\Lambda}^{(0)}$ plays the role of a spatiotemporal sCM. To obtain the commonly considered, purely spatial sCM, that measures influences of one point on another without regard to timing, one must integrate $\mathbf{\Lambda}^{(0)}$ over all possible values of t and t' to account for all influences that travel directly to x from x', regardless of timing [40]. This yields a purely spatial propagator:

$$\mathbf{\Lambda}^{(0)}(\mathbf{r},\mathbf{r}') = \int\int \mathbf{\Lambda}^{(0)}(\mathbf{r},t,\mathbf{r}',t')dtdt', \tag{14}$$

The left side represents the commonly considered purely spatial propagator, generalized to include multiple populations; an

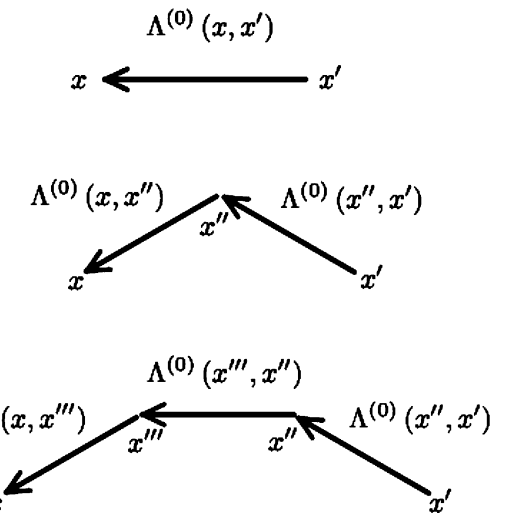

FIG. 3. Feynman diagrams showing (a) direct propagation, (b) propagation via one intermediate node, and (c) propagation via two intermediate nodes. The dressed propagator is a sum of these, and all higher-order, bare propagators. Note that there is no requirement that intermediate locations differ, so loops can occur.

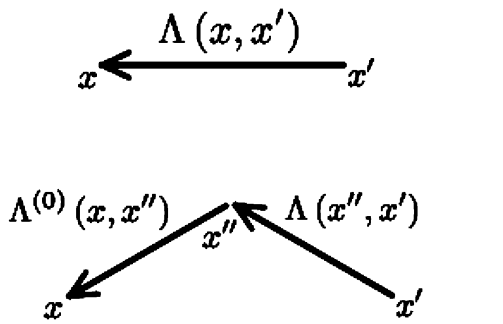

FIG. 4. Diagrammatic representation of Eq. (12), showing that dressed (i.e., effective) propagation to x from x' can be viewed as dressed propagation to x'' from x', followed by bare (i.e., direct) propagation to x from x'', where x'' is the location of the last interaction before x and all its possible locations are integrated over.

equivalent result holds for $\mathbf{\Lambda}$. Note that the dimensions of the spatial propagators differ from the spatiotemporal ones and that the integrals in Eq. (14) and similar expressions below can be extended to $\pm\infty$ without error because of the causality condition imposed in the definition of $\mathbf{\Lambda}^{(0)}$ at Eq. (1).

Equation (14) can be used to write down purely spatial analogs of the equations in the preceding sections. For example, the analog of Eq. (2) is

$$\mathbf{Q}(\mathbf{r}) = \int \mathbf{\Lambda}^{(0)}(\mathbf{r},\mathbf{r}')\mathbf{Q}(\mathbf{r}')d\mathbf{r}' + \mathbf{N}(\mathbf{r}). \tag{15}$$

D. Correlations, covariance, and coherence

A quantity often used to define the functional connectivity matrix is the covariance matrix of activity at spatial locations $\mathbf{r}$ and $\mathbf{r}'$ [1,2,10,11],

$$\mathbf{Cov}(\mathbf{r},\mathbf{r}') = \langle \mathbf{Q}(\mathbf{r},t)\mathbf{Q}^T(\mathbf{r}',t)\rangle, \tag{16}$$

where the average is over t or, equivalently if the system is ergodic, over realizations of the external inputs. The points $\mathbf{r}$ and $\mathbf{r}'$ are defined to be functionally connected if the covariance exceeds an arbitrarily imposed threshold, but difficulties surround the choice of threshold and the interpretation of negative values [2,10].

More generally, the unnormalized correlation matrix $\mathbf{C}$ is defined

$$\mathbf{C}(x,x') = \langle \mathbf{Q}(x)\mathbf{Q}^T(x')\rangle, \tag{17}$$

which strictly requires ergodicity and/or time-stationarity to permit the averaging to be done. In typical experimental circumstances, the averaging is done over multiple observations, whose representativity must then be assumed or tested.

The forms of Eqs. (16) and (17) are correct for $\mathbf{Q}$ viewed as a column vector but must be transposed if $\mathbf{Q}$ is viewed as the equivalent row vector, as is sometimes the case in the literature. We also stress that the population subscripts a and b are implicit in this notation. In time-stationary systems, $\mathbf{C}$ depends only on the difference $\tau = t - t'$, so

$$\mathbf{C}(\mathbf{r},\mathbf{r}',\tau) = \langle \mathbf{Q}(\mathbf{r},t+\tau)\mathbf{Q}^T(\mathbf{r}',t)\rangle, \tag{18}$$

with the average taken over t.

The covariance matrix can then be written

$$\mathbf{Cov}(\mathbf{r},\mathbf{r}') = \langle \mathbf{C}(\mathbf{r},t,\mathbf{r}',t)\rangle, \tag{19}$$

which is symmetric and, hence, contains no information on timing or the direction of causality between $\mathbf{r}$ and $\mathbf{r}'$.

From Eq. (6) one obtains

$$\mathbf{C}(x,x') = \left\langle \int\int \mathbf{T}(x,y)\mathbf{N}(y)\mathbf{N}^T(y')\mathbf{T}^T(x',y')dydy' \right\rangle, \tag{20}$$

$$= \int\int \mathbf{T}(x,y)\langle \mathbf{N}(y)\mathbf{N}^T(y')\rangle \mathbf{T}^T(x',y')dydy', \tag{21}$$

$$= \int \mathbf{T}(x,y)\mathbf{T}^T(x',y)dy. \tag{22}$$

In obtaining Eq. (21), $\mathbf{T}$ has been assumed to be a fixed quantity, while Eq. (22) further assumes that $\mathbf{N}(y)$ and $\mathbf{N}^T(y')$ have zero mean, are spatiotemporally uncorrelated, and have unit amplitude in (21); i.e.,

$$\langle \mathbf{N}(y)\mathbf{N}^T(y')\rangle = \mathbf{I}\delta(y-y'), \tag{23}$$

$$\mathbf{N}\mathbf{N}^T = \mathbf{I}, \tag{24}$$

where Eq. (23) is for the p-element coordinate-dependent form and Eq. (24) is for the P-element pure matrix form.

Another useful measure is the coherence between two points. This frequency-dependent quantity is defined as

$$\gamma^2(\mathbf{r},\mathbf{r}',\omega) = \frac{[C(\mathbf{r},\mathbf{r}',\omega)]^2}{C(\mathbf{r},\mathbf{r},\omega)C(\mathbf{r}',\mathbf{r}',\omega)}. \tag{25}$$

III. TRANSLATIONALLY INVARIANT CASES

An important family of special cases in which the above results take on significantly simpler forms involves propagators whose functional forms are translationally invariant in space and/or steady-state in time. It is important to note that translational invariance does *not* imply uniform all-to-all connectivity, still less activity that is constant; i.e., a system whose underlying properties are constant in space and time can support nonuniform activity.

While real brain propagators do not possess exact translational invariance because of spatial inhomogeneities and/or time variation, this is often a reasonable first approximation and it serves here to elucidate the physics and illustrate key links among sCMs, eCMs, and fCMs without the additional complications of the general case. The general case then can be investigated numerically, with its interpretation aided by insights obtained in more tractable situations. Analytic results also provide essential tests for verifying general numerical calculations.

In the spatiotemporally invariant case,

$$\mathbf{\Lambda}^{(0)}(x,x') = \mathbf{\Lambda}^{(0)}(x-x'), \tag{26}$$

and propagation depends only on the vector displacement between x and x'. In the case of time invariance of the propagator, one has

$$\mathbf{\Lambda}^{(0)}(x,x') = \mathbf{\Lambda}^{(0)}(\mathbf{r},\mathbf{r}',t-t'), \tag{27}$$

and, if there is translational invariance in space,

$$\mathbf{\Lambda}^{(0)}(x,x') = \mathbf{\Lambda}^{(0)}(\mathbf{r}-\mathbf{r}',t,t'). \tag{28}$$

In the case of full translational invariance of propagators, the integrals over the translationally invariant variables have the form of convolutions. These can be further analyzed by Fourier transforming them, which converts convolutions into products in Fourier space [72]. Equation (2) becomes

$$\mathbf{Q}(x) = \int \mathbf{\Lambda}^{(0)}(x-x')\mathbf{Q}(x')dx' + \mathbf{N}(x). \tag{29}$$

Similarly, the analog of Eq. (14) is

$$\mathbf{\Lambda}^{(0)}(\mathbf{r}-\mathbf{r}') = \int \mathbf{\Lambda}^{(0)}(\mathbf{r}-\mathbf{r}',t-t')d(t-t'), \tag{30}$$

while Eq. (17) becomes

$$\mathbf{C}(x-x') = \int \mathbf{T}(x-y)\mathbf{T}^T(x'-y)dy. \tag{31}$$

In Fourier space we write $k = (\mathbf{k},\omega)$, where $\mathbf{k}$ is the wave number and ω is the angular frequency and define the Fourier transform and its inverse by

$$F(k) = \int e^{-ik\cdot x} F(x) dx, \tag{32}$$

$$F(x) = \int e^{ik\cdot x} F(k) \frac{dk}{(2\pi)^{D+1}}, \tag{33}$$

where D is the spatial dimensionality and the dot product is defined by

$$k \cdot x = \mathbf{k}\cdot\mathbf{r} - \omega t. \tag{34}$$

In Fourier form, Eqs. (29)–(31) become

$$\mathbf{Q}(k) = \mathbf{\Lambda}^{(0)}(k)\mathbf{Q}(k) + \mathbf{N}(k), \tag{35}$$

$$\mathbf{\Lambda}^{(0)}(\mathbf{k}) = \mathbf{\Lambda}^{(0)}(\mathbf{k},\omega)\delta(\omega), \tag{36}$$

$$\mathbf{C}(k) = \mathbf{T}(k)\mathbf{T}^{\dagger}(k), \tag{37}$$

where the dagger denotes the Hermitian conjugate (i.e., the complex conjugate of the transposed matrix). Equations (33) and (37) imply

$$\mathbf{C}(x - x') = \int e^{ik\cdot(x-x')} \mathbf{T}(k)\mathbf{T}^{\dagger}(k) \frac{dk}{(2\pi)^{D+1}}, \tag{38}$$

which is the spatiotemporal version of the Wiener-Khinchtine theorem [73].

The cross-spectrum $\mathbf{C}(\mathbf{r},\mathbf{r}',\omega)$ is the temporal Fourier transform of $\mathbf{C}$ in the time-invariant case:

$$\mathbf{C}(\mathbf{r},\mathbf{r}',\omega) = \int e^{i\omega(t-t')} \mathbf{C}(\mathbf{r},\mathbf{r}',t-t')d(t-t'), \tag{39}$$

$$= \langle \mathbf{Q}(\mathbf{r},\omega)\mathbf{Q}^{\dagger}(\mathbf{r}',\omega)\rangle, \tag{40}$$

and represents the correlation between activity at frequency ω at locations $\mathbf{r}$ and $\mathbf{r}'$. When suitably normalized we recover the coherence $\boldsymbol{\gamma}(\mathbf{r},\mathbf{r}',\omega)$, which has unit entries for maximally correlated activity. In the translationally invariant case this becomes

$$\gamma(\mathbf{R},\omega) = C(\mathbf{R},\omega)/C(\mathbf{0},\omega). \tag{41}$$

A major advantage of translationally invariant systems is that the propagators only have to be parametrized by coordinate differences, not by the coordinates themselves, which means that the representations are lower dimensional and much more compact than in the general case.

IV. MEASUREMENT AND SIGNAL-PROCESSING EFFECTS

Most experimental estimates of fCMs are based on proxies for neural activity rather than activity itself. For example, in fMRI, covariances of perturbations relative to the baseline of the blood-oxygen-level-dependent (BOLD) signals at different points are often used to estimate functional connectivity [1,2,10,11,25]. Similar covariances in signals from other measurement modalities could also be used.

In fMRI, which is the most commonly used means of defining fCMs, the BOLD signal results from a complicated set of hemodynamic interactions driven by neural activity and is approximately the result of a linear spatiotemporal convolution of a complex hemodynamic response function with the neural activity itself [74–76]. In the case of this and other approximately linear functions of activity, including smoothing functions that may be applied in postprocessing, we can write a measured quantity $\mathbf{Z}$ as

$$\mathbf{Z}(x) = \int \mathbf{M}(x,x')\mathbf{Q}(x')dx', \tag{42}$$

$$= \int \mathbf{M}(x - x')\mathbf{Q}(x')dx', \tag{43}$$

where $\mathbf{M}$ incorporates both the physical processes that lead from $\mathbf{Q}$ to the measurable quantity, and linear postprocessing that yields the reported measurement $\mathbf{Z}$. Equation (43) holds in the translation-invariant case, which is an idealization, in which case we then find

$$\mathbf{Z}(k) = \mathbf{M}(k)\mathbf{Q}(k). \tag{44}$$

In the translationally invariant case a propagator for $\mathbf{Z}$ itself can be derived by Fourier transforming Eq. (43) and using Eq. (4), which yields

$$\mathbf{Z}(k) = \mathbf{M}(k)\mathbf{\Lambda}^{(0)}(k)\mathbf{M}^{-1}(k)\mathbf{Z}(k) + \mathbf{M}(k)\mathbf{N}(k). \tag{45}$$

Hence, the bare propagator for $\mathbf{Z}$ is

$$\mathbf{\Lambda}_Z^{(0)}(k) = \mathbf{M}(k)\mathbf{\Lambda}^{(0)}(k)\mathbf{M}^{-1}(k), \tag{46}$$

with

$$\mathbf{\Lambda}_Z^{(0)}(k) = \mathbf{\Lambda}^{(0)}(k), \tag{47}$$

if $\mathbf{M}$ commutes with $\mathbf{\Lambda}^{(0)}$. The eCM $\mathbf{\Lambda}_Z$ for $\mathbf{Z}$ is, thus, defined by analogy with the case for $\mathbf{Q}$. Figure 5 illustrates the correspondence between the propagators for the firing rates $\mathbf{Q}$ and those for the observed quantity $\mathbf{Z}$.

One can define the correlation matrix for $\mathbf{Z}$

$$\mathbf{C}_Z(x,x') = \langle \mathbf{Z}(x)\mathbf{Z}^T(x')\rangle, \tag{48}$$

$$= \iiint \mathbf{M}(x,y)\mathbf{T}(y,z) \times \mathbf{T}^T(y',z)\mathbf{M}^T(y',z)dydy'dz, \tag{49}$$

where Eq. (49) assumes Eq. (23). In the translationally invariant case, Eq. (48) yields

$$\mathbf{C}_Z(k) = \mathbf{Z}(k)\mathbf{Z}^{\dagger}(k), \tag{50}$$

$$= \mathbf{M}(k)\mathbf{T}(k)\mathbf{T}^{\dagger}(k)\mathbf{M}^{\dagger}(k), \tag{51}$$

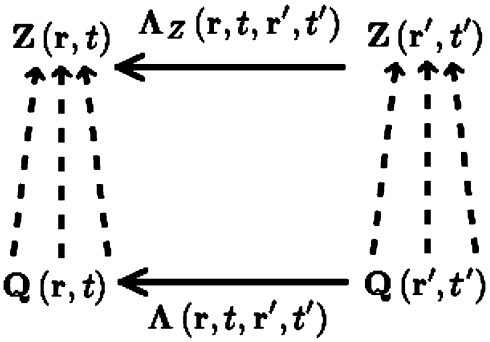

FIG. 5. Schematic of the correspondence between the propagators for the firing rates $\mathbf{Q}$ and those for the observed quantity $\mathbf{Z}$. The dashed arrows indicate the convolution over a spacetime region that yields $\mathbf{Z}$ from $\mathbf{Q}$.

which emphasizes that this class of measurement matrices **M** act as spatiotemporal filters.

The above results can be used to estimate the effects of coarse-graining CMs. Typically, coarse graining of measurements occurs because it is not experimentally possible to trace connections between arbitrarily small brain regions of interest (ROIs). Instead, the brain is parcellated into some number R of regions to form the basis of a purely spatial CM. As the spatial size of ROIs increases, the likelihood of pairwise connections between nodes rises and the CMs become fuller. We note that it is possible for a signal to pass through multiple synapses within a single node before progressing to another; however, this effect becomes less important as node sizes decrease.

Coarse graining can be written in the form of Eq. (41), at least approximately, by making M a spatial kernel function with characteristic width equal to the size of an ROI. Similarly, the effects of time averaging in determining CMs can be explored by using **Z** to model a temporal averaging window, which, thus, can be pictured as a "temporal ROI." This is relevant to the question of task-related changes to functional connectivity, which are critical to specific responses but tend to average out over long periods of time [2,7,11]. Spatiotemporal ROIs can also be defined and are actually implicit in all existing ROIs, since these all involve both spatial and temporal measurement scales. Spatial and temporal averaging restricts attention to low-k and low-ω components of the fCM—in other words, to large-scale, long-lasting features.

A step that is often taken in estimating CMs is to threshold the raw fCM to remove all values below a certain level [this is not of the form of Eq. (42) or Eq. (43) but can be handled separately] [1,2,10]. Alternatively, information on strengths of connections is omitted and the CM is simply converted to binary form, with entries equal to 1 (an above-threshold connection exists) or 0 (no above-threshold connection). These give similar results, but it is increasingly being recognized that connection-strength information should be retained if possible [1,2,10].

The outline of measurement and signal-processing issues presented here is far from comprehensive but serves to indicate how such effects can be incorporated into CM analysis. We illustrate these points in Secs. V and VI.

V. NEURAL FIELD FORMULATION

In this section we use NFT to write the above results in forms that can be applied to a variety of situations in Sec. VI to provide concrete illustrations of the key results. To achieve this we use a specific model of the corticothalamic system that has been successfully applied to explain many aspects of brain activity [8,37,43,44,52–54]. In addition, use of translationally invariant propagators in the analysis permits us to use analytically tractable forms and, thus, to obtain insights that would be difficult to achieve in the general case. In Sec. V A we briefly review the model used, which has been introduced and discussed in detail in the publications just cited. Section III B then expresses the key results from Secs. II–IV in this form. We note that this is just one example of NFT and that one can also use propagators derived from other variants; e.g., Refs. [34–36,38,39,45–49,55–71], although not all of these include the axonal time delays that are essential to calculate full correlation matrices, for example.

A. Model system

We illustrate the key aspects of the analysis in Sec. II using an established model of the corticothalamic system, schematically illustrated in Fig. 6 [42–44,52,53]. This system includes cortical excitatory (e) and inhibitory (i) populations, along with populations representing thalamic specific relay (s) nuclei and the thalamic reticular (r) nucleus; external inputs (n) are also included. It has sufficient flexibility to model corticothalamic systems using physiologically realistic parameters, while also enabling specialized limiting cases to be examined to elucidate more fundamental issues, which is the approach taken here. We approximate the dynamics using NFT, in which the properties of many neurons are averaged over to obtain equations of motion for mean activity as a function of position and time [34–49,77]. This is likely a good approximation to cortical networks, which have up to 10^{11} nodes when viewed as networks of neurons. Note that the cortex can be viewed approximately as a two-dimensional sheet, due to its thinness, while the coordinates in the thalamus are related one to one to those in the cortex via the primary thalamocortical mapping. This means that distances between thalamic points are actually expressed in terms of distances between the corresponding points on the cortex, so a single coordinate can be used to refer to all structures [42]. Physical distances (and corresponding velocities) on the thalamus are about 0.1 times the one-to-one mapped cortical distances **r** [42].

It has been previously established that the system in Fig. 6 has a steady-state solution with low population firing rates, so long as the input is not too strong [40,43,52–54]. Here we consider small perturbations from this fixed point, and all quantities are viewed as perturbations.

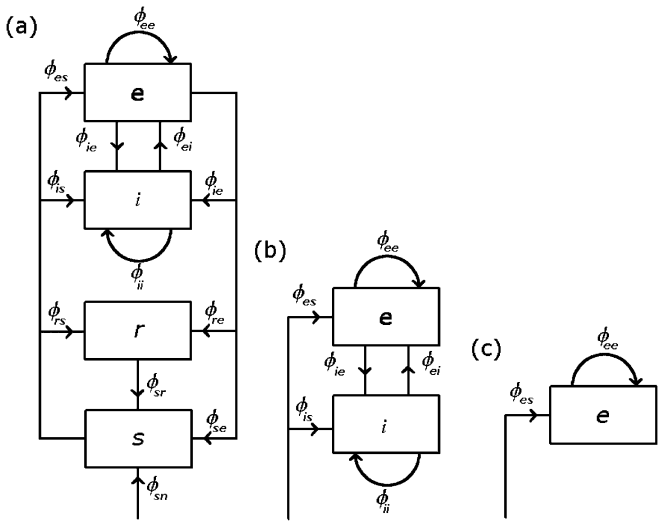

FIG. 6. Schematic of the model systems considered. (a) Corticothalamic system, showing the excitatory (e), inhibitory (i), reticular nucleus (r), and specific nuclei (s) populations, inputs n, and the fields that travel among them. (b) The cortical subsystem. (c) Purely excitatory subsystem.

Perturbations Q_a of mean firing rate can be approximated as being linearly related to perturbations of the mean soma voltage V_a, with

$$Q_a(\mathbf{r},t) = \rho_a V_a(\mathbf{r},t), \tag{52}$$

where ρ_a is the slope of the full nonlinear response curve that relates Q_a to V_a, evaluated at the fixed point; ρ_a is assumed constant here.

The perturbation V_a to mean soma potential is a sum of contributions V_{ab} from various other populations b

$$V_a(\mathbf{r},t) = \sum_b V_{ab}(\mathbf{r},t). \tag{53}$$

The subpotential V_{ab} responds to the incoming field of afferent activity ϕ_{ab} from neurons b according to

$$D_{ab}V_{ab}(\mathbf{r},t) = \nu_{ab}\phi_{ab}(\mathbf{r},t-\tau_{ab}), \tag{54}$$

$$D_{ab} = \frac{1}{\alpha\beta}\frac{d^2}{dt^2} + \left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\frac{d}{dt} + 1, \tag{55}$$

where α and β are characteristic rates of soma voltage decay and rise, respectively, which we treat as independent of a and b here, although this assumption can be easily relaxed [42]. In Eq. (54) $\nu_{ab} = N_{ab}s_{ab}$, s_{ab} is the synaptic strength expressed as the time-integrated soma voltage response to a unit input, and τ_{ab} is any discrete time delay required for signals to reach population a from b and is always zero if $a = b$ (for $a = b$ time delays are accounted for via the internal propagation of ϕ_{aa}, as discussed below). Quantities such as velocities, ranges, and time delays are interpreted as mean or typical values in NFT, not as extremal ones.

We can also write Eq. (54) in temporal propagator form as

$$V_{ab}(\mathbf{r},t) = \nu_{ab}\int_{-\infty}^{t} K_{ab}(t-t')\phi_{ab}(\mathbf{r}',t'-\tau_{ab})dt', \tag{56}$$

$$K_{ab}(u) = \alpha\beta[e^{-\alpha u} - e^{-\beta u}]/(\beta-\alpha), \tag{57}$$

for $u \geqslant 0$, with $K_{ab}(u) = 0$ for $u < 0$, and with $\beta \gg \alpha$ without loss of generality. In Fourier space, we have

$$K_{ab}(\omega) = (1-i\omega/\alpha)^{-1}(1-i\omega/\beta)^{-1} = L(\omega), \tag{58}$$

where Eq. (58) defines $L(\omega)$.

The activity fields ϕ_{ab} derive from the source Q_b and propagate according to the propagator equation

$$\phi_{ab}(\mathbf{r},t) = \iint \Gamma_{ab}(\mathbf{r}-\mathbf{r}',t-t')Q_b(\mathbf{r}',t')d\mathbf{r}'dt', \tag{59}$$

whence

$$\phi_{ab}(\mathbf{k},\omega) = \Gamma_{ab}(\mathbf{k},\omega)Q_b(\mathbf{k},\omega), \tag{60}$$

where Γ_{ab} is the mean propagator for axons projecting to $\mathbf{r},t$ from $\mathbf{r}',t'$ and we have assumed translation invariance for simplicity, although this is not mandatory in NFT [42].

Using (52)–(57) we can write

$$Q_a(\mathbf{r},t) = \sum_b \iint \Lambda^{(0)}_{ab}(\mathbf{r}-\mathbf{r}',t-t'-\tau_{ab}) \times Q_b(\mathbf{r}',t')d\mathbf{r}'dt', \tag{61}$$

with

$$\Lambda^{(0)}_{ab}(\mathbf{r}-\mathbf{r}',t-t') = G_{ab}\int K_{ab}(t-t'') \times \Gamma_{ab}(\mathbf{r}-\mathbf{r}',t''-t')dt'', \tag{62}$$

$$\Lambda^{(0)}_{ab}(\mathbf{k},\omega) = G_{ab}L(\omega)\Gamma_{ab}(\mathbf{k},\omega)e^{i\omega\tau_{ab}}, \tag{63}$$

$$G_{ab} = \rho_a\nu_{ab} = \rho_a N_{ab}s_{ab}, \tag{64}$$

where the gain G_{ab} is the average number of extra spikes produced in neurons of type a per extra incoming spike from neurons of type b. Note that Eq. (63) clarifies how the sCM $\mathbf{\Lambda}^{(0)}$ is related to the underlying aCM, which we identify with $\mathbf{\Gamma}$, with $\mathbf{\Lambda}^{(0)}$ including the effects of both time delays and strengths (i.e., gains) of connections.

A number of forms of Γ_{ab} have been proposed, including integral forms that are either local or nonlocal, either of which can be employed in the analysis developed above, although some nonlocal propagators do not retain the axonal delays required to calculate correlations, for example. A particular class of axonal propagator corresponds to differential equations (typically damped wave equations) for the fields ϕ_{ab} [37–40,45,46,48,49]. Using one such form [40,42], Eq. (59) can be written in differential form as

$$\mathcal{D}_{ab}\phi_{ab}(\mathbf{r},t) = Q_b(\mathbf{r},t), \tag{65}$$

$$\mathcal{D}_{ab} = \frac{1}{\gamma_{ab}^2}\frac{\partial^2}{\partial t^2} + \frac{2}{\gamma_{ab}}\frac{\partial}{\partial t} + 1 - r_{ab}^2\nabla^2, \tag{66}$$

where r_{ab} is the characteristic range of axons projecting to neurons a from b, v_{ab} is the mean axonal velocity, and $\gamma_{ab} = v_{ab}/r_{ab}$ is an effective damping rate that reflects the rate at which spikes traverse the finite range of axons. Equations (60) and (66) imply

$$\Gamma_{ab}(\mathbf{k},\omega) = \frac{1}{(1-i\omega/\gamma_{ab})^2 + k^2r_{ab}^2}, \tag{67}$$

$$= \frac{1}{\mathcal{D}_{ab}(\mathbf{k},\omega)}. \tag{68}$$

Note that when we write $\mathbf{k}$ and ω, or $\mathbf{r}$ and t, separately (as here and below), rather than in space-time notation [cf., Sec. II A, where $k = (\mathbf{k},\omega)$], we use the standard notations $k = |\mathbf{k}|$ and $r = |\mathbf{r}|$ for simplicity.

In the approximately 2D cortex, the propagator [Eq. (67)] has the coordinate-space form [40]

$$\Gamma_{ab}(\mathbf{R},\tau) = \frac{v_{ab}e^{-\gamma_{ab}\tau}\Theta(v_{ab}\tau - R)}{2\pi r_{ab}^2(v_{ab}^2\tau^2 - R^2)^{1/2}}, \tag{69}$$

where the Heaviside function $\Theta(\tau)$ enforces causality. Thus, the effect of a δ-function stimulus at the origin dissipates at a rate γ_{ab} while propagating outward behind a front that moves at velocity v_{ab}. Mean-field dissipation corresponds to the reduction in amplitude that occurs when signals terminate at the ends of axons, which are approximately exponentially distributed in range.

In 1D, one finds

$$\Gamma_{ab}(X,\tau) = \frac{e^{-\gamma_{ab}\tau}}{2r_{ab}\tau}[\Theta(X+v_{ab}\tau) - \Theta(X-v_{ab}\tau)]. \tag{70}$$

Here we write the spatial coordinate difference as X to distinguish it from the spatiotemporal coordinate x used in earlier sections.

For some purposes it is useful to have a simpler form than that of Eq. (69) or Eq. (70), including for some of the illustrations in later sections. Hence, in 1D we introduce the asymmetric form

$$\Gamma_{ab}(X,\tau) = \tfrac{1}{2}e^{-\gamma_{ab}\tau}[(1+\eta)\delta(X - v_{ab}\tau) + (1-\eta)\delta(X + v_{ab}\tau)], \tag{71}$$

with $-1 \leqslant \eta \leqslant 1$, which corresponds to two dissipative δ functions propagating away from the origin, with strengths $(1 \pm \eta)/2$. The case $\eta = 0$ corresponds to the symmetric propagator introduced in Ref. [38]. Fourier transformation of Ref. [71] yields

$$\Gamma_{ab}(k,\omega) = \frac{(1 - i\omega/\gamma_{ab}) - ikr_{ab}\eta}{(1 - i\omega/\gamma_{ab})^2 + k^2 r_{ab}^2}. \tag{72}$$

The 1D differential equation that corresponds to Ref. [72] is

$$\mathcal{D}_{ab}\phi_{ab}(x,t) = \left(1 + \frac{\partial}{\partial t} - \eta r_{ab}\frac{\partial}{\partial x}\right) Q(x,t), \tag{73}$$

with $\nabla^2 = \partial^2/\partial x^2$ in Eq. (66) in this case. It is important to stress that asymmetry is not equivalent to lack of translational invariance: The asymmetry can be the same at every point.

It should be noted that propagators such as Eqs. (67) and (72) have recently been described as corresponding to a long-wavelength limit of an integral formulation akin to Eq. (59) or Eq. (62) [58,61] and, hence, as being less general. However, this is a misinterpretation: the present propagator Γ_{ab} and the one in the cited references both approximate the actual brain connectivity and either can be derived exactly from the other in the Fourier domain if $kr_{ab} \to 0$ is valid for all relevant wave numbers—so neither propagator is more fundamental and neither should be considered a "long-wavelength approximation." Indeed, Ref. [58] noted that PDE forms can be derived without such an approximation, in the process rediscovering a previous expression for the propagator in terms of a difference of two Macdonald functions (modified Bessel functions of the second kind) [40,72]. In fact, the main difference between the two propagators is that they approach zero at slightly different rates as $kr_{ab} \to \infty$, a difference that cannot be distinguished on the basis of current connectivity data. Both can be, and have been, successfully used down to scales of order 1 mm [63–65,70,71].

B. NFT evaluation of connection matrices

The sCM $\mathbf{\Lambda}^{(0)}$ has elements defined by Eq. (63) with the accompanying definitions of $L(\omega)$ and $\Gamma_{ab}(\mathbf{k},\omega)$. From this it is immediately possible to calculate $\mathbf{\Lambda}$, correlations, and covariances via the matrix manipulations discussed in previous sections. To obtain more insight into the physical effects involved, including measurement issues, we focus on cases relevant to electroencephalographic (EEG) and fMRI measurements, especially as it is the latter that are most often used to define fCMs. The overall sCM for the system shown in Fig. 6(a) is

$$\mathbf{\Lambda}^{(0)} = \begin{pmatrix} \Lambda_{ee}^{(0)} & \Lambda_{ei}^{(0)} & 0 & \Lambda_{es}^{(0)} \\ \Lambda_{ie}^{(0)} & \Lambda_{ii}^{(0)} & 0 & \Lambda_{is}^{(0)} \\ \Lambda_{re}^{(0)} & 0 & 0 & \Lambda_{rs}^{(0)} \\ \Lambda_{se}^{(0)} & 0 & \Lambda_{sr}^{(0)} & 0 \end{pmatrix}, \tag{74}$$

which interrelates the following quantities from Eq. (2):

$$\mathbf{Q} = \begin{pmatrix} Q_e \\ Q_i \\ Q_r \\ Q_s \end{pmatrix}, \tag{75}$$

$$\mathbf{N} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ N_s \end{pmatrix}, \tag{76}$$

where Eq. (76) highlights that external stimuli enter the system via the relay nuclei.

The system in Fig. 6(a) has been previously studied [42–44]. It has been found that $r_{ab} \approx 0$ is a reasonable approximation for $b = i,r,s$, implying that the corresponding $\gamma_{ab} \approx \infty$, the *local interaction approximation* [40,42], in which $\Gamma_{ab}(\mathbf{k},\omega) = 1$. Moreover, since the average number of synapses linking cortical neurons of types a and b is closely proportional to the numbers of neurons involved [40,48,78], and synaptic properties (e.g., excitatory or inhibitory) depend on the afferent neuron, we can make the *random connectivity approximation* that $\Lambda_{ib} = \Lambda_{eb}$ for $b = e,i,s$. These approximations yield

$$\Lambda_{ee}^{(0)} = \Lambda_{ie}^{(0)} = \frac{G_{ee}L}{(1 - i\omega/\gamma_{ee})^2 + k^2 r_{ee}^2}, \tag{77}$$

$$= J_{ee}\Gamma_{ee}, \tag{78}$$

$$\Lambda_{ei}^{(0)} = \Lambda_{ii}^{(0)} = G_{ei}L = J_{ei}, \tag{79}$$

$$\Lambda_{es}^{(0)} = \Lambda_{is}^{(0)} = G_{es}Le^{i\omega\tau_{es}} = J_{es}, \tag{80}$$

$$\Lambda_{re}^{(0)} = G_{re}Le^{i\omega\tau_{re}} = J_{re}, \tag{81}$$

$$\Lambda_{se}^{(0)} = G_{se}Le^{i\omega\tau_{se}} = J_{se}, \tag{82}$$

$$\Lambda_{rs}^{(0)} = G_{rs}L = J_{rs}, \tag{83}$$

$$\Lambda_{sr}^{(0)} = G_{sr}L = J_{sr}, \tag{84}$$

$$\Lambda_{sn}^{(0)} = G_{sn}L = J_{sn}, \tag{85}$$

where we have introduced the compact notation [42]

$$J_{ab} = G_{ab}Le^{i\omega\tau_{ab}}, \tag{86}$$

and have noted that the only nonzero discrete interpopulation delays are τ_{es}, τ_{se}, and τ_{re}, with intrapopulation delays included via the γ_{aa} [42,43,52]. In some previous work, the notation $N_s = \Lambda_{\rm sn}\phi_{\rm sn}$ was used [42], with $\phi_{\rm sn}$ being the incoming signal from outside the corticothalamic system, as shown in Fig. 6(a).

C. EEG and fMRI measurements

Both EEG and fMRI measurements primarily detect signals resulting indirectly from spikes afferent at synapses that cause synaptic currents due to transient opening of ion channels.

In the case of EEG, synaptic currents close to form current dipoles that are strongest for cortical pyramidal cells,

particularly as these are more numerous than inhibitory interneurons and tend to be aligned so the dipoles reinforce one another instead of canceling [39,79–81]. EEG then detects resulting voltage changes at the scalp after spatial filtering by the effects of volume conduction in the intervening material.

The primary contribution is fMRI is the metabolic load required to restore the resting voltage difference after synaptic ion channels have been opened by afferent spikes [82,83]; again pyramidal cells dominate and dominate in driving the astrocytic signaling that mediates the hemodynamic response [82–84] that is ultimately detected via the BOLD signal in fMRI [74,75,84–86].

Given the points made in the previous paragraph, we choose the signal

$$z(\mathbf{r},t) = \phi_{ee}(\mathbf{r},t), \tag{87}$$

to trace our sCM, eCM, and fCM rather than tracking the full matrices. This retains spatial information, but only in a single field rather than the 10 internal fields in Fig. 6(a), but ignores effects that link ϕ_{ee} to the final measurements, which depend on the type of measurement done. The corresponding measurement matrix is

$$\mathbf{M}(\mathbf{r},t) = \mathrm{diag}[\Gamma_{ee}(\mathbf{r},t),\epsilon,\epsilon,\epsilon], \tag{88}$$

where ϵ is an arbitrarily small nonzero quantity that preserves the invertibility of $\mathbf{M}$. We, thus, find $\mathbf{\Lambda}_Z^{(0)} = \mathbf{\Lambda}^{(0)}$, as in Eq. (47), and then take the limit $\epsilon \to 0$ from this point onward.

In Fourier space, we have

$$\phi_{ee}(\mathbf{k},\omega) = \Gamma_{ee}(\mathbf{k},\omega)T_{es}(\mathbf{k},\omega)N_s(\mathbf{k},\omega), \tag{89}$$

since the only external input is N_s in Fig. 6(a).

From Eqs. (5)–(7) and (74)–(86) we can now calculate the eCM Λ_{es} in Fourier space, noting that $\delta_{es}(\mathbf{k},\omega) = 0$:

$$T_{es}(\mathbf{k},\omega) = \delta_{es}(\mathbf{k},\omega) + \Lambda_{es}(\mathbf{k},\omega) = \Lambda_{es}(\mathbf{k},\omega), \tag{90}$$

$$= [\mathbf{I} - \mathbf{\Lambda}^{(0)}(\mathbf{k},\omega)]^{-1}_{es}, \tag{91}$$

$$= \frac{\Lambda^{(0)}_{es}}{\det[\mathbf{I} - \mathbf{\Lambda}^{(0)}(\mathbf{k},\omega)]}, \tag{92}$$

$$\det[\mathbf{I} - \mathbf{\Lambda}^{(0)}] = \left[1 - \Lambda^{(0)}_{ee} - \Lambda^{(0)}_{ei}\right]\left[1 - \Lambda^{(0)}_{sr}\Lambda^{(0)}_{rs}\right] - \Lambda^{(0)}_{es}\left[\Lambda^{(0)}_{se} + \Lambda^{(0)}_{sr}\Lambda^{(0)}_{re}\right]. \tag{93}$$

Using Eqs. (90)–(93), we can write $\phi_{ee} = T_{Zes}N_s$, with the Fourier space form

$$T_{Zes}(\mathbf{k},\omega) = \frac{A(\omega)}{k^2r_{ee}^2 + p^2r_{ee}^2}, \tag{94}$$

$$A(\omega) = J_{es}/\Delta, \tag{95}$$

$$p^2r_{ee}^2 = (1 - i\omega/\gamma_{ee})^2 - J_{ee}(1 - J_{sr}J_{rs})/\Delta, \tag{96}$$

$$\Delta = (1 - J_{ei})(1 - J_{sr}J_{rs}) - J_{es}(J_{se} + J_{sr}J_{re}). \tag{97}$$

D. EEG and fMRI propagators, correlations, covariance, and coherence

The form of Eq. (94) of T_{ee} is particularly convenient for studying the correlation and covariance properties of the system. First, we note that the spatial transfer function is obtained from Eq. (47), which gives

$$\Lambda_Z(\mathbf{k},\omega) = T_{Zes}(\mathbf{k},\omega). \tag{98}$$

By inverse Fourier transforming the result [Eq. (36)] in space, and using Eq. (98), we obtain the spatial eCM

$$\Lambda_Z(x) = \frac{G_{es}e^{-p_0|x|}}{2r_{ee}^2p_0\Delta_0}, \tag{99}$$

$$\Lambda_Z(\mathbf{R}) = \frac{G_{es}K_0(p_0R)}{2\pi r_{ee}^2\Delta_0}, \tag{100}$$

where p_0 and Δ_0 are the values of p and Δ at $\omega = 0$. Averaging over any small range of position (as is unavoidable experimentally), while allowing for the implicit $2\pi R$ factor mentioned earlier, removes the singularity in Eq. (100) at $R = 0$, so it and related singularities below do not cause any practical difficulties.

Using the results of Ref. [41] for transfer functions of this general functional form we can immediately write

$$C_Z(\mathbf{k},\omega) = \left|\frac{A(\omega)}{k^2r_{ee}^2 + p^2r_{ee}^2}\right|^2, \tag{101}$$

$$C_Z(x,\omega) = \frac{|A(\omega)|^2e^{-p_r|x|}}{4|p|^2}\left[\frac{\cos(p_ix)}{p_r} + \frac{\sin(p_i|x|)}{p_i}\right], \tag{102}$$

$$C_Z(\mathbf{R},\omega) = \frac{\pi|A(\omega)|^2}{p_ip_r}\mathrm{Im}[K_0(p^*R)], \tag{103}$$

where Eqs. (102) and (103) apply in 1D and 2D, respectively; p_i and p_r are the real and imaginary parts of p, respectively; the asterisk in Eq. (103) denotes the complex conjugate; and K_0 is a Macdonald function (modified Bessel function of the second kind) [72]. The Fourier transforms required to find $C_Z(R,T)$ from Eqs. (102) or (103) can be evaluated analytically only for certain limiting cases of the system in Fig. 6 and are usually most easily evaluated by numerical means.

The coherence function can be calculated directly from Eq. (103) using the definition of Eq. (41). This gives

$$\gamma(x,\omega) = e^{-p_r|x|}\left[\cos(p_i|x|) + \frac{p_r\sin(p_i|x|)}{p_r}\right], \tag{104}$$

$$\gamma(\mathbf{R},\omega) = \frac{\mathrm{Im}K_0(p^*R)}{\mathrm{Arg}\,p}, \tag{105}$$

in 1D and 2D, respectively, where Arg is the complex argument; Eq. (105) is equivalent to a previous result [41] and is unity at $R = 0$.

VI. ILLUSTRATIVE APPLICATIONS

In this section we apply the NFT results of Sec. V to study systems of increasing generality in order to explore specific physical effects and thereby elucidate the theoretical results. In particular, we explore the case of a single excitatory population, the purely cortical case of excitatory and inhibitory neurons, and situations with loop-induced resonances; asymmetric cases such as Eq. (70) can equally be treated using the same methods. In all cases, we use $z = \phi_{ee}$ to trace connection matrices. In future, general cases can be

analyzed with the benefit of interpretations that draw on these results.

A. Excitatory population only

The simplest subcase of the system shown in Fig. 6(a) is the one shown in Fig. 6(c), which we use to illustrate some of the key aspects of our analysis. This corresponds to setting $J_{es} = J_{sn} = 1$, $J_{ie} = J_{re} = J_{se} = J_{sr} = J_{rs} = 0$ and deleting terms corresponding to the ei, ie, is, and ii connections, which removes references to these in Eqs. (77)–(86).

Equation (67) shows $\Gamma_{ee}(\mathbf{k},\omega)$, with Eqs. (69) and (70) giving the corresponding coordinate space forms in 2D and 1D, respectively. The spatial aCM is obtained by integrating these expressions over time or, more simply, by Fourier transforming the relevant analog of (36), which gives

$$\Gamma_{ee}(x) = \frac{e^{-|x|/r_{ee}}}{2r_{ee}}, \tag{106}$$

$$\Gamma_{ee}(\mathbf{R}) = \frac{K_0(R/r_{ee})}{2\pi r_{ee}^2}, \tag{107}$$

in 1D and 2D, respectively, and the sCM is obtained by multiplying these by G_{ee}. Equation (76) enables the corresponding results to be calculated for the sCM $\Lambda_{es}^{(0)}$, which equals Γ_{ee} here because we have set $J_{es} = 1$.

The eCM is given by Eq. (98) with $A(\omega) = 1$ and

$$p^2 r_{ee}^2 = (1 - i\omega/\gamma_{ee})^2 - G_{ee}L. \tag{108}$$

From Eqs. (99) and (106), we obtain the spatial eCM

$$\Lambda_Z(x) = \frac{1}{2r_{ee}(1-G_{ee})^{1/2}} \exp\left[-\frac{|x|(1-G_{ee})^{1/2}}{r_{ee}}\right]. \tag{109}$$

Figure 7(a) shows the 1D eCM $\Lambda_Z(x)$ for various values of the gain G_{ee}, with $\Lambda_Z^{(0)}$ corresponding to the case $G_{ee} = 0$. We see that the scale length of Λ_Z diverges as $G_{ee} \to 1$, in accord with Eq. (109). The point $G_{ee} = 1$ has previously been determined to correspond to a saddle-node bifurcation in this system [40]. This divergence occurs because an average of G_{ee} action potentials are regenerated for each action potential that is produced. Accordingly, the spatial integral of Eq. (108) is $(1 - G_{ee})^{-1} = 1 + G_{ee} + G_{ee}^2 + \cdots$, which diverges at the critical point. As $G_{ee} \to 0$ the sCM and eCM become equal, which is in accordance with the picture in Sec. II, and in Fig. 3 in particular, when the strength of scattering is zero.

The divergence of the correlation length at the critical point implies that the system becomes closely coupled over increasingly long ranges and that influences can, thus, cascade throughout it. This accords with the results of a number of authors on avalanches of neural activity and other signs of criticality (self-organized or otherwise) in neural systems [12,16,18–20], but we stress that criticality is inherent at all spectral resonances for the reasons discussed here and in other contexts [41]: The existence of a resonance implies low damping, so coherent waves can propagate long distances and build up to high amplitudes.

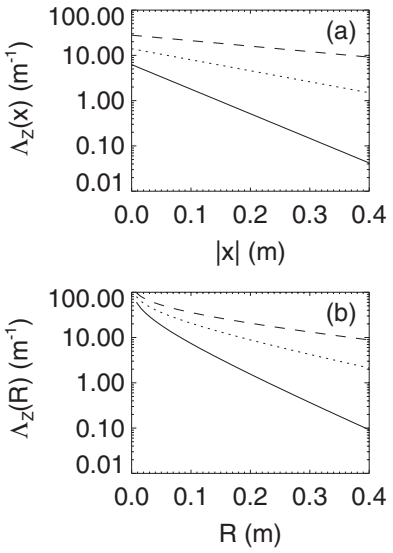

FIG. 7. Effective CMs Λ_Z for the system of Fig. 6(c) with $r_{ee} = 0.08$ m, a typical value in the cortex [39,42,52], $G_{es} = 1$, $G_{ee} = 0,0.8.0.95$ (from bottom to top in each frame), and $\alpha = \beta = 1000\ \mathrm{s}^{-1}$ to minimize complexities arising from synaptodendritic low-pass filtering in this example. The sCM $\Lambda_Z^{(0)}$ corresponds to the case $G_{ee} = 0$. (a) 1D vs. $|x|$. (b) 2D vs. R.

Figure 7(b) shows the corresponding propagators for the 2D case, from Eqs. (100) and (108). Here the scale length diverges in the same way as in 1D, with

$$\Lambda_Z(\mathbf{R}) = \frac{K_0[R(1-G_{ee})^{1/2}/r_{ee}]}{2\pi r_{ee}^2}. \tag{110}$$

The amplitude of this result has a weakly divergent (logarithmic) dependence on $1 - G_{ee}$ as $G_{ee} \to 1$, while the spatial integral of Eq. (110) again diverges as $(1 - G_{ee})^{-1}$ for the same reason as in 1D. Although Λ_Z diverges logarithmically as $R \to 0$, as seen in Fig. 7(b), this must be multiplied by a factor $2\pi R$ to find the differential number of connections per unit R, which has no singularity.

Figure 8 shows the fCM given by normalized values of the covariance $\mathrm{Cov}_Z(x) = C_Z(x, T = 0)$ and $\mathrm{Cov}_Z(\mathbf{R}) = C_Z(\mathbf{R}, T = 0)$ for the 1D and 2D cases, respectively. The covariance is evaluated by integrating $C_Z(x,\omega)/2\pi$ and $C_Z(\mathbf{R},\omega)/2\pi$ over ω. These integrals are dominated by resonant frequencies ω_{res} where Eqs. (102) or (103) are large; i.e., points where $|pr_{ee}| \ll 1$. In the present case, $\omega_{\mathrm{res}} = 0$ and near this frequency p^2 can be approximated by a complex-valued linear function of ω that passes close to the origin, with a minimum modulus proportional to $1 - G_{ee}$. Concentrating on the dominant spatial dependence, we, thus, find

$$\mathrm{Cov}_Z(x) \propto \exp[-|x|(1-G_{ee})^{1/2}/r_{ee}], \tag{111}$$

$$\mathrm{Cov}_Z(\mathbf{R}) \propto K_1\left[\frac{R(1-G_{ee})^{1/2}}{r_{ee}}\right], \tag{112}$$

where K_1 is a Macdonald function [72]. In both cases, the correlation length diverges near the critical point, as $(1 - G_{ee})^{-1/2}$, as for the underlying propagators. The amplitude of the unnormalized correlations also increases as $G_{ee} \to 1$, but the functional form is not simple and the integrals cannot be

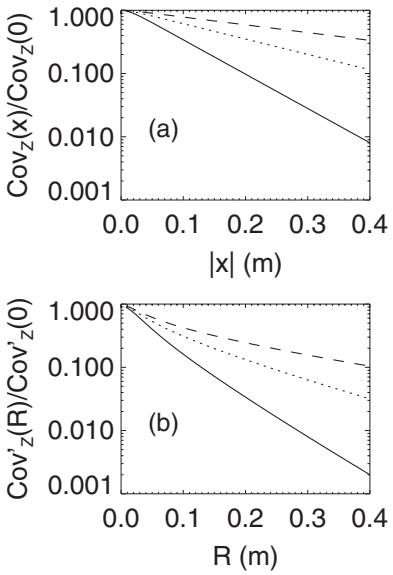

FIG. 8. Spatial functional CM Cov_Z for the system of Fig. 6(c), evaluated numerically for $r_{ee} = 0.08$ m, $\gamma_{ee} = 125$ s^{-1}, $\alpha = \beta = 1000$ s^{-1}, and $G_{es} = 1$, with $G_{ee} = 0{,}0.8.0.95$ from bottom to top in each frame. (a) 1D $\mathrm{Cov}_Z(x)/\mathrm{Cov}_Z(0)$ vs. $|x|$. (b) 2D $\mathrm{Cov}'_Z(\mathbf{R})/\mathrm{Cov}'_Z(0)$ vs. R, where Cov'_Z is an average of Cov_Z over a few mm, as would occur in experimental measurements, which removes the singularity at $R = 0$.

evaluated analytically in closed form. The coherence function also has the same divergent correlation length r_{ee}/p_0 as the critical point is approached.

Figure 9 shows the dependence of normalized forms of the correlation function $C_Z(0,T)$ versus T for the same system as in Figs. 7 and 8; i.e., the temporal correlation at a given spatial

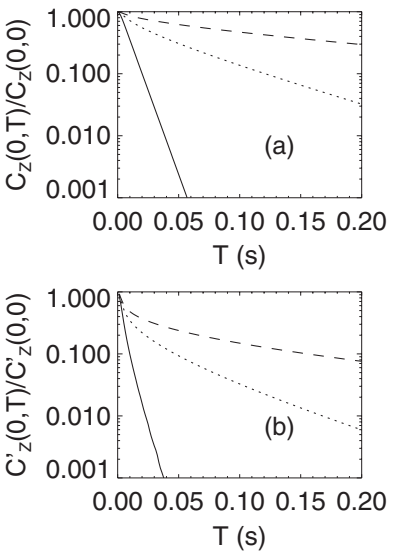

FIG. 9. Normalized correlation function C_Z for the system of Fig. 6(c), evaluated numerically for $r_{ee} = 0.08$ m, $\gamma_{ee} = 125$ s^{-1}, $\alpha = \beta = 1000\,\mathrm{s}^{-1}$, and $G_{es} = 1$, with $G_{ee} = 0{,}0.8.0.95$, from bottom to top in each frame. (a) 1D $C_Z(0,T)/C_Z(0,0)$ vs. $|x|$. (b) 2D $C'_Z(0,T)/C'_Z(0,0)$ vs. R, where C'_Z is an average of C_Z over a few mm, as would occur in experimental measurements, which removes the singularity at $R = 0$.

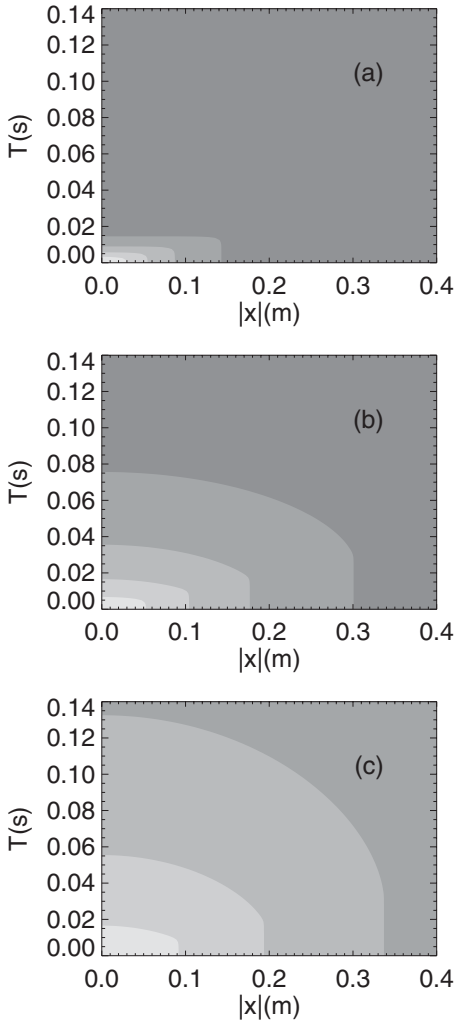

FIG. 10. 1D spatiotemporal functional CMs $C_Z(x,T)/C_Z(0,0)$ vs. $|x|$ and T for the system of Fig. 6(c) with the same parameters as Figs. 7–9 and light colors indicating high values, with contours in steps of 0.2. (a) sCM (eCM with $G_{ee} = 0$). (b) eCM, $G_{ee} = 0.8$; (c) eCM, $G_{ee} = 0.95$.

point. There is a rapid fall off with T in both 1D and 2D, with a slope that decreases rapidly in magnitude as $G_{ee} \to 1$, again reflecting the existence of the critical point. Analytic results for $G_{ee} = 0$ show that the correlation time is the lesser of α^{-1} and γ_{ee}^{-1}, which is confirmed by the results in Fig. 9 and similar calculations for other parameters.

In Figs. 10 and 11 we show normalized forms of the spatiotemporal fCM $C_Z(x,T)$ and $C_Z(\mathbf{R},T)$ in 1D and 2D, respectively, for several values of G_{ee}. In both cases, the zone of high correlation expands rapidly as $G_{ee} \to 1$, in accord with the results in Figs. 8 and 9, whose curves are cuts along the coordinate axes of the frames in Figs. 10 and 11, respectively.

Connectivities are most often visualized by dividing the cortex into nodes and then plotting the results for the spatial propagators and correlations in the form of CMs, whose axes are labeled by node number, as in Fig. 1 [1–3,5,7–12,16,25]. Figure 12 shows the spatial sCM, eCM, and fCM in this format for the 1D propagator (99) and various G_{ee}. The strongest connections are short-range, corresponding to the

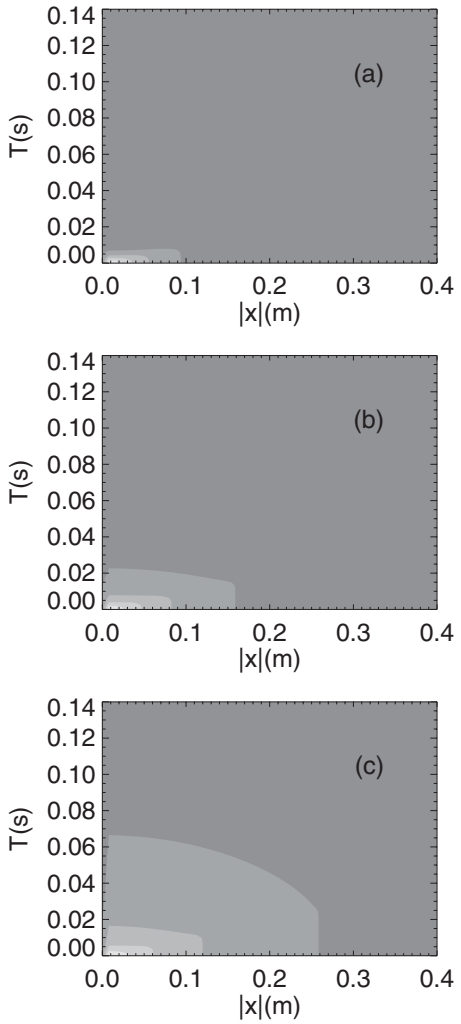

FIG. 11. 2D spatiotemporal functional CMs $C'_Z(\mathbf{R},T)/C'_Z(0,0)$ vs. R and T for the system of Fig. 6(c) with the same parameters as Figs. 7–9 and light colors indicating high values, with contours in steps of 0.2, and where C'_Z is an average of C_Z over a few mm, as would occur in experimental measurements, which removes the singularity at $R = 0$. (a) sCM (eCM with $G_{ee} = 0$). (b) eCM, $G_{ee} = 0.8$; (c) eCM, $G_{ee} = 0.95$.

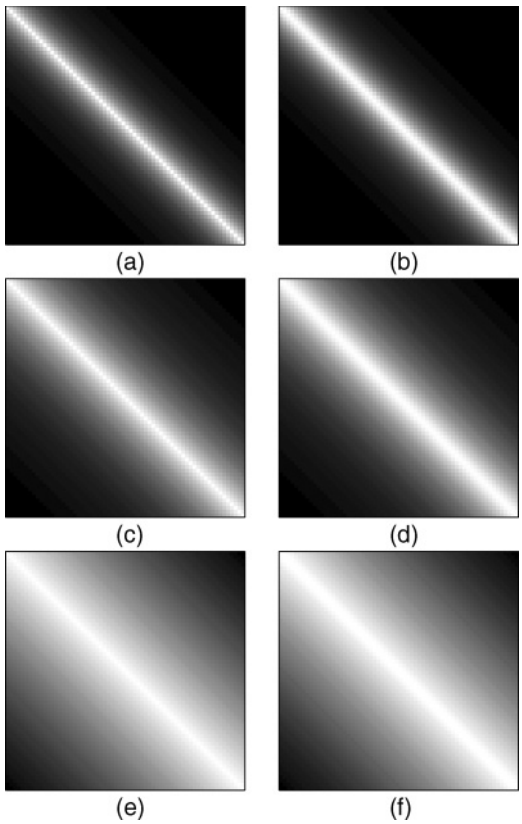

FIG. 12. 1D spatial sCMs, eCMs, and fCMs from Figs. 7(a) and 8(a), displayed in the connection matrix format of Fig. 1, with light colors indicating strong connections. A linear system of length 0.8 m with open boundaries has been divided into 64 nodes. The left column shows eCMs (the sCM is the topmost frame), while the right column shows fCMs. The rows have $G_{ee} = 0,0.8,0.99$, from top to bottom.

main diagonal, and fall off rapidly with distance, as was seen in earlier figures. As G_{ee} increases, the eCM and fCM increasingly fill with larger entries.

When node numbers are used to label 2D CMs, it is impossible to give all adjacent nodes consecutive labels, and the appearance of the resulting CM depends sensitively on the actual labeling used [9]. Recently, a labeling was introduced that preferentially assigns nearby labels to adjacent nodes [9] via an iterative algorithm. In this approach, the 2D spatial array of nodes is divided in half along each dimension, giving four equal subnetworks. Each subnetwork is then assigned a random sequential quarter of node labels available for the whole network. These steps are then repeated for each of the subnetworks and iterated until each node has a unique label. The stages for a 64 node network are shown in Fig. 13 [9].

When the labeling described in the previous paragraph is applied to a 2D system with 64 nodes, we find the CMs seen in Fig. 14. These CMs have the characteristic appearance of approximate block diagonality seen in many published CMs [cf. Fig. 1(b)] and giving a strong impression of nested levels of modules that is often taken to be a sign of modularity and/or hierarchy in experimental CMs. However, as discussed recently [9], this is an illusion: The present system has no hierarchy, and no modularity beyond that implicit in the fact that nearby nodes are more likely to be connected than distant ones, due to the rapid fall-off in distance seen in Eqs. (99) and (100).

The top left and bottom right frames of Fig. 14 are the ones that correspond most closely to figures such as Fig. 1 [which, however, has 998 nodes rather than the 64 in Fig. 14, giving a finer structure]. The fCM only fills with large entries very close to the critical point. This raises the possibility that proximity to the critical point may be measurable via comparisons between sCMs and fCMs, which would be analogous to previously suggested spectral measures of proximity to criticality [40].

B. Excitatory and inhibitory populations

The approach introduced here has the advantage that it covers multiple neural populations, including inhibitory ones. If we generalize the analysis of Sec. VI A to include both populations in Fig. 6(b), this corresponds to setting $G_{es} = J_{sn} = 1, G_{se} = G_{re} = G_{sr} = G_{rs} = 0$ in Fig. 6(a). The results of Sec. VI A remain qualitatively unchanged, except that $A(\omega)$

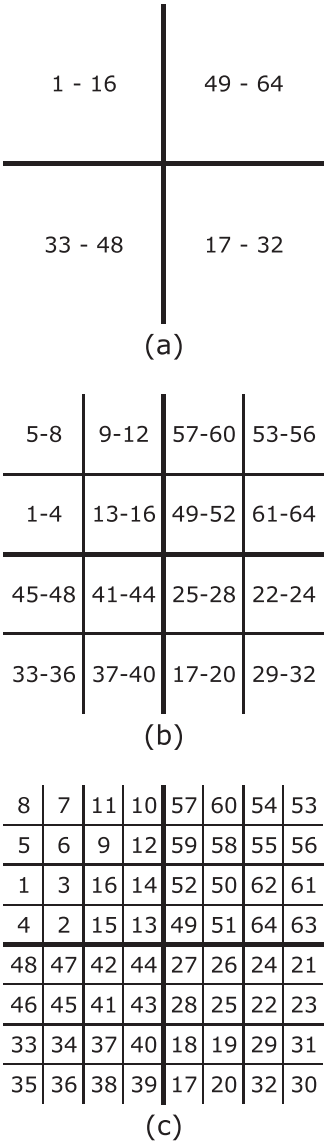

FIG. 13. 2D 64-node labeling method [9]. (a) First subdivision and random assignment of 16 consecutive labels to each subdivision. (b) Second subdivision and random assignment of 4 consecutive labels to each subdivision. (c) Final subdivision with unique labels assigned randomly to each node.

and the term $G_{ee}L$ in Eq. (108) are both divided by the factor $1 - G_{ei}L$. This leads to

$$\Lambda_Z(x) = \frac{1}{2r_{ee}\Delta_0}\left(1 - \frac{G_{ee}}{1-G_{ei}}\right)^{-1/2} \times \exp\left[-\frac{|x|}{r_{ee}}\left(1 - \frac{G_{ee}}{1-G_{ei}}\right)^{1/2}\right], \quad (113)$$

$$\Lambda_Z(\mathbf{R}) = \frac{1}{2\pi r_{ee}^2\Delta_0} K_0\left[\frac{R}{r_{ee}}\left(1 - \frac{G_{ee}}{1-G_{ei}}\right)^{1/2}\right], \quad (114)$$

$$\Delta_0 = 1 - G_{ei}, \quad (115)$$

in 1D and 2D, respectively. These results can be summarized by noting that the quantity $G_{ee}\Lambda_Z$ is the propagator for all the

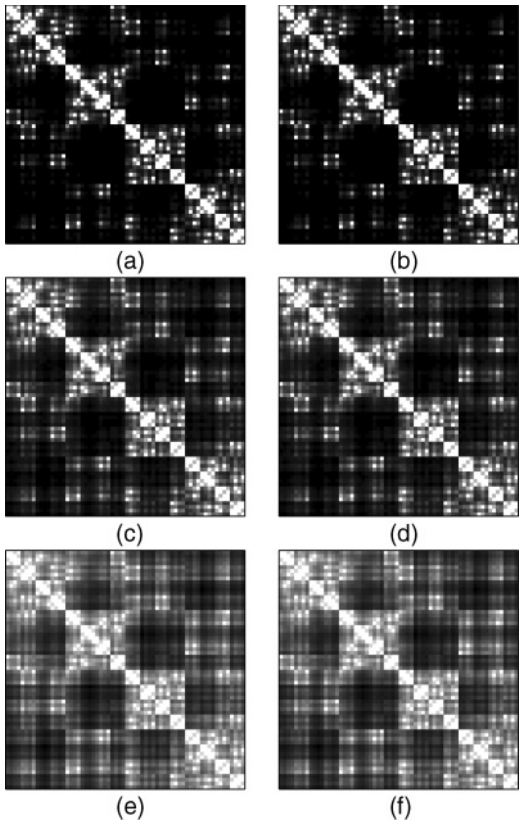

FIG. 14. 2D spatial sCMs, eCMs, and fCMs from Figs. 7(b) and 8(b), displayed in the connection matrix format of Fig. 1. A square system of linear size 0.64 m has been divided into 8×8 nodes and relabeled using the method illustrated in Fig. 12. The left column shows eCMs (the sCM is the topmost frame), while the right column shows fCMs. The rows have $G_{ee} = 0, 0.8, 0.99$, from top to bottom. This figure should be compared with the top-left or bottom-right quarter of Fig. 1(b), which contain intrahemispheric connections.

G_{ee} spikes produced per spike afferent on the e population in Sec. VI A and that this is modified by the replacement $G_{ee} \to G_{ee}/(1 - G_{ei})$ throughout. This replacement corresponds to a renormalization of the propagator due to the inhibitory effects of the i population. Correspondingly, the results shown in Eqs. (111) and (112), and the resulting correlation length, are modified in the same way. There is still a saddle-node bifurcation at $p_0 = 0$, but this now occurs at $G_{ee}/(1 - G_{ei}) = 1$ or, equivalently, $G_{ee} + G_{ei} = 1$ [40]. This allows large, nearly balanced gains to coexist, while maintaining stability, since G_{ei} is negative. Correlation lengths and times diverge at the critical point in similar ways to those discussed in Sec. VI A.

C. CT system: Effects of resonances

For the system in Fig. 6(a), with $J_{es} = 1$, the results (98)–(105) apply, with

$$A(0) = 1/\Delta_0, \quad (116)$$

$$p_0^2 r_{ee}^2 = 1 - G_{ee}(1 - G_{sr}G_{rs})/\Delta_0, \quad (117)$$

$$\Delta_0 = (1 - G_{ei})(1 - G_{sr}G_{rs}) - G_{es}(G_{se} + G_{sr}G_{re})e^{i\omega t_0}. \quad (118)$$

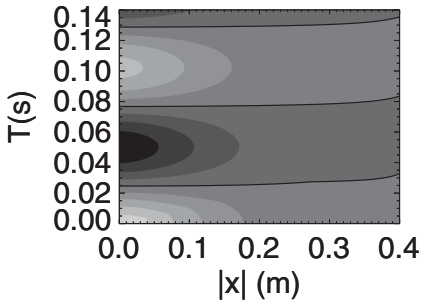

FIG. 15. 1D spatiotemporal functional CMs $C(x,T)/C(0,0)$ vs. $|x|$ and T for the system of Fig. 6(a) with the parameters $\alpha = \beta = 60\ \mathrm{s}^{-1}$, $\gamma_{ee} = 125\ \mathrm{s}^{-1}$, $r_{ee} = 0.08$ m, $t_0 = 0.08$ s, $G_{ee} = G_{ei} = G_{re} = G_{se} = 0$, $G_{es} = 1.0$, $G_{sr} = -1.85$, $G_{rs} = 1.85$. Light colors indicate high values and contours where $C(x,T) = 0$ are marked by solid curves.

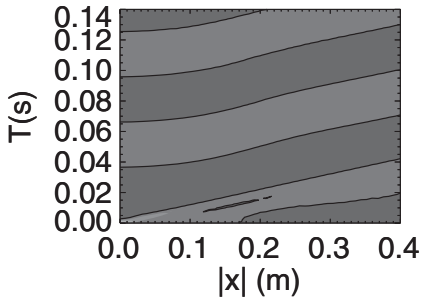

FIG. 16. 1D spatiotemporal functional CMs $C(x,T)/C(0,0)$ vs. $|x|$ and T for the system of Fig. 6(a) with the parameters $\alpha = 80\ \mathrm{s}^{-1}$, $\beta = 800\ \mathrm{s}^{-1}$, $\gamma_{ee} = 116\ \mathrm{s}^{-1}$, $r_{ee} = 0.086$ m, $t_0 = 0.085$ s, $G_{ee} = 6.8$, $G_{ei} = -8.1$ $G_{re} = 1.0$, $G_{se} = 2.5$, $G_{es} = 1.7$, $G_{sr} = -1.9$, $G_{rs} = 0.8$, which are the same as those in Ref. [52], where typical human corticothalamic values were inferred for cases with strong alpha resonances. Light colors indicate high values and contours where $C(x,T) = 0$ are marked by solid curves.

The most significant difference from previous sections is the appearance of the oscillatory term involving the factor $e^{i\omega t_0}$, with $t_0 = \tau_{es} + \tau_{se} = \tau_{es} + \tau_{re}$ being the total loop delay, which produces resonances as it moves in and out of phase with the lead term in Eq. (118). Positive static feedback corresponds to the prefactor of $e^{i\omega t_0}$ being positive, in which case there are resonances when $\omega t_0 \approx 2\pi m$, with $m = 0,1,2,\ldots$ [43,44,54]. The system continues to exhibit a saddle-node bifurcation at $\omega = 0$ when $p_0 = 0$, and the correlation length diverges as r_{ee}/p_0 as $p_0 \to 0$.

Resonances occur when $|p(\omega)r_{ee}| \ll 1$ at nonzero ω. To probe their effect, we first examine a case of an intrathalamic resonance that occurs in the loop that reciprocally links the r and s nuclei in Fig. 6(a). This loop has a resonance at

$$\omega = \sqrt{\alpha\beta}, \tag{119}$$

which becomes linearly unstable when

$$G_{sr}G_{rs} = -(\alpha+\beta)^2/\alpha\beta, \tag{120}$$

[42,43,53]. On tuning the system close to this instability, we find the 1D covariance seen in Fig. 15, which is analogous to that in Fig. 10(c). Now, strong temporal modulations of the covariance are seen, with a period consistent with the frequency shown in Eq. (119) and that extend to increasingly large T as the critical point is approached. Negative values of covariance occur because the oscillations of the least stable frequency are out of phase when observed at locations where they are half a cycle apart, which occurs with a period in accord with Eq. (119).

Another type of resonance occurs in the corticothalamic system in Fig. 6(a) when the positive feedback loop from e to s and back is dominant, giving rise to a resonance at $\omega \approx 2\pi/t_0$. This feedback induces a system resonance, which lies at approximately 10 Hz when α and β are large [43,54]. Paralleling the discussion that leads to Eqs. (111) and (112), we expect divergent correlation length and time as this resonance becomes sharper and approaches a subcritical Hopf bifurcation [44]. Moreover, the corticothalamic waves that correspond to this resonance have a propagation velocity close to v_{ee}, which should yield a normalized correlation that varies in both space and time. Figure 16 shows just this behavior, with small-$|x|$ correlations that oscillate vs. T, similarly to those in Fig. 14, plus large-$|x|$ oscillations that propagate at approximately 10 m $\mathrm{s}^{-1} = v_{ee}$ [87], as indicated by the sloping contours (the period is about 0.06 s because of the combined effects of both the alpha frequency and its harmonic). Even at $T = 0$ both positive and negative values of the covariance are possible, a phenomenon seen in experimental fCMs.

VII. SUMMARY AND DISCUSSION

This paper introduces an approach based on propagators (Green's functions) to facilitate the analysis of anatomical, effective, and functional connection matrices (aCMs, sCMs, eCMs, and fCMs) in brain networks. The methods of neural field theory are then applied to evaluate and explore the results in illustrative examples that highlight the main findings. Some of the key results obtained are as follows:

(i) Anatomical, strength-of-connection, effective, and functional CMs have been defined in terms of propagators that incorporate both spatial and temporal aspects of connectivity: aCMs and sCMs correspond to bare propagators that carry signals directly between points along direct anatomical connections, with sCMs including the strength of connectivity; eCMs correspond to dressed propagators that include all neural influences of one point on another, whether direct or via intermediate nodes; and fCMs are defined in terms of normalized correlations between points. Very low levels of activity, or perturbations in activity, suffice to define connections between points, so linear theory can be used to probe the relationships between CMs (nonlinear interactions are, of course, also possible, but are not considered here). Either spiking or rate-based measures of activity work in such definitions.

(ii) The definitions in (i) are sufficiently general to enable multiple effects to be included whose incorporation has hitherto been problematic: excitation and inhibition, multiple cortical and noncortical neural populations, loops, time delays, directionality, measurement and postprocessing effects, and matrix and continuum formulations that can handle approximations to CMs or CMs determined from experimental data.

(iii) Use of the definitions in (i), plus the machinery of propagator theory, enables sCMs and eCMs to be interrelated

and fCMs to be calculated from the other two. Notably, eCMs can be expanded in terms of powers of sCMs, which elucidates the roles of indirect paths involving various numbers of intermediate nodes and may permit interpretation of brain imaging in terms of these effects. The relationship shown in Eq. (6) of sCMs to the transfer function enables a canonical normalization of the sCM $\Lambda^{(0)}$ to be defined and to replace the arbitrary normalizations used in the literature to date.

(iv) The importance of indirect connections is highlighted, particularly as these give rise to network effects, including feedback loops and resonances, some of which involve multiple structures (e.g., corticothalamic loops) and delays. This emphasizes the need to include temporal aspects and noncortical structures in analyses of brain connectivity and provides the tools to do so.

(v) It is shown how measurement and postprocessing effects can be incorporated. These include unavoidable distortions of neural activity measures in the indirect signals that are actually recorded, e.g., the EEG which results after dipole currents are generated and resulting potential changes are modified by volume conduction [39,81], and fMRI which measures the BOLD signal that arises after complex spatiotemporal hemodynamics [74–76,86]. Postprocessing such as spatial and/or temporal averaging, thresholding, and other steps can also be included. It is anticipated that further consideration of these aspects will help to clarify the connections between CMs determined at different spatial and temporal resolutions, including the effects of transient network modulations under task conditions.

(vi) The main results have been re-expressed in terms of NFT in order to treat continuum approximations to neural systems. NFT approaches also enable analytic insights to be obtained into a range of phenomena, including critical points and other aspects of propagator behavior. Continuum approximations become increasingly appropriate as the number of network nodes increases beyond the tens or hundreds typically used to date, especially when one considers that the brain actually has 10 to 100 billion nodes at the neural level.

(vii) Neural field theory has been used to compute the CMs for exemplar model single-population, cortical, and corticothalamic systems. These examples serve to illustrate the methods and key physical effects, but propagators from other forms of NFT and other physical systems can equally well be used in the present analysis. Notably, critical points exist at previously discovered saddle-node and Hopf bifurcations of the dynamics of this system. At these points, the system is marginally stable, correlation lengths and times diverge, and eCMs and fCMs increasingly fill with non-negligible entries as the whole system becomes highly correlated prior to a transition to a nonlinearly synchronized state [40,43,44]. This filling of the matrices accords with observations [2,10,11]. When delay loops exist, resonances occur that ultimately give rise to Hopf bifurcations [43,44]. These also produce long-range correlations of activity at the resonant frequency as the critical bifurcation point is approached. Such correlations can have either sign, leading to negative entries in the fCM, and can be either propagating or nonpropagating, depending on the system modes that underlie them. The present work potentially enables the proximity to critical points to be determined quantitatively from fCMs.

(viii) The coherence matrix has also been computed and found to diverge near critical points in a similar manner to the eCM and fCM. This is relevant to recent observations of critical behavior at a wide variety of frequencies in brain networks [13,16].

(ix) NFT results for completely uniform systems have also been displayed in the standard matrix format that is widely used to display spatial CMs. In the 2D case, this highlights the recently noted fact that CMs can have a strong visual appearance of hierarchy and/or modularity, even when they are completely homogeneous and contain no spatial nonuniformities aside from a rapid decrease in neural connection strength with distance that has the same functional form at every point [9].

(x) The propagator-based approach provides a number of natural analogies with other field theories of physics, particularly in quantum contexts. These analogies can lead to useful insights in the neural case, some of which are discussed in Sec. II, but must not be taken too far. In particular, the present work does not constitute a quantum theory of the brain.

Some of the analogies with field theories are summarized in Table I. Other aspects have been investigated in connection with field theories of the statistics of two- and three-state spiking-neuron models, which we do not pursue here [34–36,88].

Overall, the work here answers many of the questions raised in the first paragraphs of the Introduction or provides tools to help address them. It might be argued that the present approach is quite complex and that one could successfully employ simpler approaches tailored to specific situations. However, the problem of brain connectivity is a difficult one, and many obstacles have been identified as lying in the path of previous methods. Thus, while simpler approaches may well work in restricted situations, the present framework

TABLE I. Some parallels between standard quantum theories of interactions of photons with atoms and the neural propagator theory developed here.

Standard	Neural
Atom	Node
Atomic state (discrete neural population state levels)	(Continuous voltage)
Bare vacuum	Neural silence
Dressed vacuum	Steady-state firing
Photon	Spike
Antiparticle	Inhibitory spike
Stimulated emission	Spike due to incoming spike
Absorption	Incoming spike does not evoke a spike
Bare propagator	sCM $\Lambda^{(0)}$
Dressed propagator	eCM Λ
Scattering	Neural interaction
Multiple scattering	Polysnaptic propagation
Charge	Gain G_{ab}
Renormalized charge	Renormalized gain, e.g., $G_{ee}/(1-G_{ei})$

provides at least some of the means to interrelate such techniques, determine their regimes of validity, and provide improved methods (including advances on the present work) in future.

Some directions for future investigations based on the groundwork laid here include applications to experimental CMs, currently underway, investigation of measurement and postprocessing effects in more detail, calculation of other network measures from the theory, and design of alternative methods that exploit the propagator formulation to achieve improved and novel imaging outcomes, such as imaging of specific levels of polysnaptic connectivity. Similar approaches to other complex networks may also be envisaged, with implications for multiple applications in physics and other fields [21,22].

ACKNOWLEDGMENTS

I thank O. Sporns for permission to use the material in Fig. 1; D. Gummersall and J. Henderson for assistance with graphics; and K. Friston, J. Henderson, S. Sarkar, R. Abeysuriya, and G. Pandejee for stimulating discussions and comments on the manuscript. The Australian Research Council and Westmead Millennium Foundation supported this work.

[1] K. J. Friston, Brain Conn. **1**, 13 (2011).
[2] E. Bullmore and O. Sporns, Nat. Rev. Neurosci. **10**, 186 (2009).
[3] M. Rubinov, O. Sporns, C. van Leeuwen, and M. Breakspear, BMC Neurosci. **10**, 55 (2009).
[4] O. Sporns, D. R. Chialvo, M. Kaiser, and C. C. Hilgetag, Trends. Cogn. Sci. **9**, 1364 (2004).
[5] O. Sporns, G. Tononi, and G. M. Edelman, Cereb. Cortex **10**, 127 (2000).
[6] M. Kaiser and C. C. Hilgetag, Front. Neuroinform. **4**, 8 (2010).
[7] C. J. Honey, R. Kötter, M. Breakspear, and O. Sporns, Proc. Natl. Acad. Sci. USA **104**, 10240 (2007).
[8] P. A. Robinson, J. A. Henderson, E. Matar, P. Riley, and R. T. Gray, Phys. Rev. Lett. **103**, 108104 (2009).
[9] J. A. Henderson and P. A. Robinson, Phys. Rev. Lett. **107**, 018102 (2011).
[10] O. Sporns, *Networks of the Brain* (MIT Press, Cambridge, MA, 2011).
[11] C. J. Honey, J.-P. Thivierge, and O. Sporns, NeuroImage **52**, 766 (2010).
[12] M. Rubinov, O. Sporns, J.-P. Thivierge, and M. Breakspear, Publ. Lib. Sci. Comp. Biol. **7**, e1002038 (2011).
[13] D. S. Bassett, A. Meyer-Lindenberg, S. Achard, T. Duke, and E. Bullmore, Proc. Nat. Acad. Sci. USA **103**, 19518 (2006).
[14] R. T. Gray and P. A. Robinson, J. Comput. Neurosci. **27**, 81 (2009).
[15] R. T. Gray, C. K. C. Fung, and P. A. Robinson, Neurocomputing **72**, 1565 (2009).
[16] M. G. Kitzbichler, M. L. Smith, S. R. Christensen, and E. Bullmore, Publ. Lib. Sci. Comp. Biol. **5**, e1000314 (2009).
[17] O. Sporns, G. Tononi, and R. Kötter, Publ. Lib. Sci. Comp. Biol. **1**, e42 (2005).
[18] J. M. Beggs and D. Plenz, J. Neurosci. **23**, 11167 (2003).
[19] C. J. Stam, and E. A. de Bruin, Hum. Brain Mapp. **22**, 97 (2004).
[20] K. Linkenkaer-Hansen, V. V. Nikouline, J. M. Palva, and R. J. Ilmoniemi, J. Neurosci. **21**, 1370 (2001).
[21] R. Albert and A. L. Barabasi, Rev. Mod. Phys. **74**, 47 (2002).
[22] M. Barthélemy, Phys. Rep. **499**, 1 (2011).
[23] R. F. Galán, PLoS ONE **3**, e2148 (2008).
[24] A. Aertsen and H. Preissl, in *Nonlinear Dynamics and Neuronal Networks*, edited by H. G. Schuster (VCH, New York, 1991), p. 281.
[25] S. A. Knock, A. R. McIntosh, O. Sporns, R. Kötter, P. Hagmann, and V. K. Jirsa, J. Neurosci. Methods **183**, 86 (2009).
[26] M. D. Grecius, K. Supekar, V. Menon, and R. F. Dougherty, Cereb. Cortex **19**, 72 (2009).
[27] C. J. Honey, O. Sporns, L. Cammoun, X. Gigandet, J. P. Thiran, R. Meuli, and P. Hagmann, Proc. Nat. Acad. Sci. USA **106**, 2035 (2009).
[28] P. Skudlarski, K. Jagannathan, V. D. Calhoun, M. Hampson, B. A. Skudlarska, and G. Pearlson, NeuroImage **43**, 554 (2008).
[29] M. Rubinov and O. Sporns, NeuroImage **52**, 1059 (2010).
[30] C. J. Stam, B. F. Jones, G. Nolte, M. Breakspear, and Ph. Scheltens, Cereb. Cortex **17**, 92 (2007).
[31] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964).
[32] W. Greiner and J. Reinhardt, *Quantum Electrodynamics*, 2nd ed. (Springer, Berlin, 1994).
[33] W. Greiner, *Relativistic Quantum Mechanics: Wave Equations* (Springer, Berlin, 1990).
[34] M. A. Buice and J. D. Cowan, Phys. Rev. E **75**, 051919 (2007).
[35] M. A. Buice and J. D. Cowan, Prog. Biophys. Molec. Biol. **99**, 53 (2009).
[36] M. A. Buice, J. D. Cowan, and C. C. Chow, Neural Comput. **22**, 377 (2010).
[37] G. Deco, V. K. Jirsa, P. A. Robinson, M. Breakspear, and K. Friston, Pub. Lib. Sci. Comp. Biol. **4**, e1000092 (2008).
[38] V. K. Jirsa and H. Haken, Phys. Rev. Lett. **77**, 960 (1996).
[39] P. L. Nunez, *Neocortical Dynamics and Human EEG Rhythms* (Oxford University Press, Oxford, 1995).
[40] P. A. Robinson, C. J. Rennie, and J. J. Wright, Phys. Rev. E **56**, 826 (1997).
[41] P. A. Robinson, J. Theor. Biol. **222**, 163 (2003).
[42] P. A. Robinson, Phys. Rev. E **72**, 011904 (2005).
[43] P. A. Robinson, C. J. Rennie, and D. L. Rowe, Phys. Rev. E **65**, 041924 (2002).
[44] M. Breakspear, J. A. Roberts, J. R. Terry, S. Rodrigues, N. Mahant, and P. A. Robinson, Cerebral Cortex **16**, 1296 (2006).
[45] M. L. Steyn-Ross, D. A. Steyn-Ross, J. W. Sleigh, M. T. Wilson, and L. C. Wilcocks, Phys. Rev. E **72**, 061910 (2005).
[46] P. L. Nunez, Math. Biosci. **21**, 279 (1974).
[47] H. R. Wilson and J. D. Cowan, Kybernetik **13**, 55 (1973).
[48] J. J. Wright and D. T. J. Liley, Behav. Brain Sci. **19**, 285 (1996).
[49] R. L. Beurle, Philos. Trans. R. Soc. London B **240**, 55 (1956).
[50] P. A. Robinson and J. W. Kim, in *Advances in Cognitive Neurodynamics III* (Springer, Berlin, 2012).

[51] P. A. Robinson and J. W. Kim, J. Neurosci. Meth. (submitted).
[52] P. A. Robinson, C. J. Rennie, D. L. Rowe, and S. C. O'Connor, Hum. Brain Mapp. **23**, 53 (2004).
[53] P. A. Robinson, C. J. Rennie, D. L. Rowe, S. C. O'Connor, and E. Gordon, Philos. Trans. R. Soc. London B **360**, 1043 (2005).
[54] P. A. Robinson, C. J. Rennie, J. J. Wright, H. Bahramali, E. Gordon, and D. L. Rowe, Phys. Rev. E **63**, 021903 (2001).
[55] M. L. Steyn-Ross, D. A. Steyn-Ross, J. W. Sleigh, and D. T. J. Liley, Phys. Rev. E **60**, 7299 (1999).
[56] B. Ermentrout, Rep. Prog. Phys. **61**, 353 (1998).
[57] S. Coombes, Bio. Cybern. **93**, 91 (2005).
[58] S. Coombes, N. A. Venkov, L. Shiau, I. Bojak, D. T. J. Liley, and C. R. Laing, Phys. Rev. E **76**, 051901 (2007).
[59] D. J. Pinto and G. B. Ermentrout, SIAM J. Appl. Math. **62**, 226 (2001).
[60] F. M. Atay and A. Hutt, SIAM J. Appl. Math. **65**, 644 (2005).
[61] I. Bojak and D. T. J. Liley, PLoS Comp. Biol. **6**, e1000653 (2010).
[62] S. Amari, Biol. Cybern. **27**, 77 (1977).
[63] P. C. Bressloff, Phys. Rev. Lett. **89**, 088101 (2002).
[64] P. C. Bressloff and J. D. Cowan, Physica D **173**, 226 (2002).
[65] G. B. Ermentrout and J. D. Cowan, Biol. Cybern. **34**, 137 (1979).
[66] S. Coombes, G. J. Lord, and M. R. Owen, Physica D **178**, 219 (2003).
[67] C. R. Laing and W. C. Troy, SIAM J. Appl. Dyn. Syst. **2**, 487 (2003).
[68] D. J. Pinto, R. K. Jackson, and C. E. Wayne, SIAM J. Appl. Dyn. Syst. **4**, 954 (2005).
[69] F. H. Lopes da Silva, A. Hoeks, H. Smits, and L. H. Zetterberg, Kybernetik **15**, 27 (1974).
[70] P. A. Robinson, Phys. Rev. E **73**, 041904 (2006).
[71] P. A. Robinson, Biol. Cybern. **97**, 317 (2007).
[72] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark (eds.), *NIST Handbook of Mathematical Functions* (Cambridge University Press, Cambridge, 2010).
[73] F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, New York, 1965).
[74] P. M. Drysdale, J. P. Huber, P. A. Robinson, and K. M. Aquino, J. Theor. Biol. **265**, 524 (2010).
[75] K. M. Aquino, M. M. Schira, P. A. Robinson, P. M. Drysdale, and M. Breakspear, PLoS Comp. Biol. (submitted).
[76] K. J. Friston, A. Mechelli, R. Turner, and C. J. Price, NeuroImage **12**, 466 (2000).
[77] W. J. Freeman, *Mass Action in the Nervous System* (Academic Press, New York, 1975).
[78] V. Braitenberg and A. Schüz, *Cortex: Statistics and Geometry of Neuronal Connectivity* (Springer, Berlin, 1991).
[79] C. Koch, *Biophysics of Computation* (Oxford University Press, Oxford, 1999).
[80] P. A. Robinson, P. N. Loxley, S. C. O'Connor, and C. J. Rennie, Phys. Rev. E **63**, 041909 (2001).
[81] P. L. Nunez and R. Srinivasan, *Electric Fields of the Brain: The Neurophysics of EEG*, 2nd ed. (Oxford University Press, Oxford, 2006).
[82] J. Schummers, H. Yu, and M. Sur, Science **320**, 1638 (2008).
[83] S. B. Laughlin, Curr. Opin. Neurobiol. **11**, 475 (2001).
[84] P. Jezzard, P. M. Matthews, and S. M. Smith (eds.), *Functional MRI: An Introduction to Methods* (Oxford University Press, Oxford, 2001).
[85] K. E. Stephan, N. Weiskopf, P. M. Drysdale, P. A. Robinson, and K. J. Friston, NeuroImage **38**, 387 (2007).
[86] P. A. Robinson, P. M. Drysdale, H. Van der Merwe, E. Kyriakou, M. K. Rigozzi, B. Germanoska, and C. J. Rennie, NeuroImage **31**, 585 (2006).
[87] P. A. Robinson, P.-C. Chen, and L. Yang, Biol. Cybern. **98**, 1 (2008).
[88] P. C. Bressloff, SIAM J. Appl. Math. **70**, 1488 (2009).