Causal connectivity of evolved neural networks during behavior

Anil K. Seth

The Neurosciences Institute, 10640 John Jay Hopkins Drive, San Diego, CA 92121, USA

E-mail: seth@nsi.edu, www.nsi.edu/users/seth

Abstract.

To show how causal interactions in neural dynamics are modulated by behavior, it is valuable to analyze these interactions without perturbing or lesioning the neural mechanism. This paper proposes a method, based on a graph-theoretic extension of vector autoregressive modeling and 'Granger causality', for characterizing causal interactions generated within intact neural mechanisms. This method, called *causal connectivity analysis*, is illustrated via model neural networks optimized for controlling target fixation in a simulated head-eye system, in which the structure of the environment can be experimentally varied. Causal connectivity analysis of this model yields novel insights into neural mechanisms underlying sensorimotor coordination. In contrast to networks supporting comparatively simple behavior, networks supporting rich adaptive behavior show a higher density of causal interactions, as well as a stronger causal flow from sensory inputs to motor outputs. They also show different arrangements of 'causal sources' and 'causal sinks': nodes that differentially affect, or are affected by, the remainder of the network. Finally, analysis of causal connectivity can predict the functional consequences of network lesions. These results suggest that causal connectivity analysis may have useful applications in the analysis of neural dynamics.

PREPRINT: Network: Computation in Neural Systems (in press).

1. Introduction

Many neurobiological processes involve mutual interactions among behavior, environment and neural mechanisms [1, 2, 3, 4, 5]. Neural mechanisms generate behavior, and are at the same time modulated by the correlations imposed by behavior and environment. To advance our understanding of these interactions, this paper proposes a method for characterizing the *causal connectivity* of a neural system, i.e. the directed graph of dynamical interactions among elements of the neural system in which each edge reflects a causal influence between two nodes. The method is based on vector autoregressive modeling and 'Granger causality', adapted from time-series analysis [6, 7], together with techniques from graph theory [8]. Unlike alternative approaches for determining causality [9, 10, 11], the method does not require perturbation or lesioning of network elements, and hence is well suited to analyzing data sets acquired during behavior from intact (simulated or biological) neural systems.

The approach is illustrated by showing how behavior and environment modulate the causal connectivity of a simulated neural network engaged in sensorimotor coordination. Evolutionary algorithms are used to generate neural networks that support target fixation behavior in a simulated head/eye system, in which the complexity of the sensorimotor environment can be experimentally varied [4]. The causal connectivities of evolved networks are then compared with their structural and behavioral properties. Causal connectivity analysis suggests several novel predictions regarding neural mechanisms of sensorimotor coordination. For example, the analysis predicts that neural mechanisms supporting head-eye coordination in rich environments will show a higher density of causal interactions, and a stronger causal flow from sensory inputs to motor outputs, than mechanisms supporting comparatively simple coordination.

The concept of Granger causality is based on prediction: if a signal A causes a signal B, then past values of A should contain information that helps predict B, above and beyond the information contained in past values of B alone [12, 6]. In practice, Granger causality is usually determined by linear modeling of time series [7]. In the simplest case with two variables, if the variance of the prediction error for B is significantly reduced by including past observations of A in the regression model, then A can be said to cause B [6].

Granger causality analysis has been used previously to identify causal relations in neurobiological data. Bernasconi and Konig applied a spectral version of Granger causality [13] to local field potential data from cat visual cortex, identifying bidirectional causal influences between supragranular and infragranular layers during a go/no-go visual discrimination task [14]. Liang *et al* used a time-varying spectral technique to differentiate feedforward, feedback, and lateral dynamical influences in monkey ventral visual cortex during visual pattern discrimination [15]. Kaminski *et al* noted increasing anterior to posterior causal influences during the transition from waking to sleep by analysis of electroencephalographic (EEG) signals [16]. Hesse *et al* used an adaptive

estimation of Granger causality to identify causal interactions in human EEG data recorded during performance of a Stroop task; they found dense webs of posterior to anterior interactions that appeared ∼400ms following stimulus onset [17]. Recently, Brovelli *et al* identified causal influences extending from primary somatosensory cortex to motor cortex in the beta-frequency range (15-30 Hz) during lever pressing by awake monkeys [18].

The present research differs from these studies by applying Granger causality analysis to *simulated* neural systems that support adaptive behaviors, and by describing causal connectivity in terms of graph-theoretic properties. Concepts of *causal density, causal disequilibrium*, *causal flow*, and *causal sources* and *causal sinks* in networks are introduced in order to describe and compare causal connectivities. Causal density reflects the fraction of significant causal interactions present in a set of network dynamics, causal disequilibrium reflects the deviation of this pattern from reciprocity, causal flow reflects the balance between outgoing and incoming causal influences for a given node, and causal sinks and sources identify network nodes that respectively differentially affect, or are affected by, the remainder of the network. Applying these concepts to a simulation model allows causal connectivity to be related to structural connectivity, and permits a detailed analysis of the modulation of causal connectivity by behavior and environment. As well as yielding insights into neural mechanisms of sensorimotor coordination, the model illustrates the utility of causal connectivity analysis, providing a platform for application of the method within the neurobiology of behavior.

2. Causal connectivity analysis

Granger causality is usually tested in the context of linear autoregressive models that predict the evolution of a time series or of a set of time series [7]. Univariate autoregressive models describe a single time series in terms of linear combinations of past values (lags) of the time-series. Multivariate (vector) autoregressive (VAR) models include lags of multiple time-series. To illustrate Granger causality, consider two time series $X_1(t)$ and $X_2(t)$ of length T. Suppose that the temporal dynamics of $X_1(t)$ and $X_2(t)$ can be described by a bivariate autoregressive model:

$$\begin{aligned} X_1(t) &= \sum_{j=1}^{p} A_{11,j} X_1(t-j) + \sum_{j=1}^{p} A_{12,j} X_2(t-j) + E_1(t) \\ X_2(t) &= \sum_{j=1}^{p} A_{21,j} X_1(t-j) + \sum_{j=1}^{p} A_{22,j} X_2(t-j) + E_2(t) \end{aligned} \tag{1}$$

where p is the maximum number of lags included in the model (the model order, $p < T$), A contains the estimated coefficients of the model, and E_1, E_2 are residuals for each time series. If the variance of the prediction error E_1 (or E_2) is reduced by the inclusion of the X_2 (or X_1) terms in the first (or second) equation, then it is said that X_2 (or X_1) *Granger-causes* X_1 (or X_2). In other words, X_2 Granger-causes X_1

if all the coefficients in A_{12} are jointly significantly different from zero. This can be tested by performing an F-test of the null hypothesis that $A_{12} = 0$, given assumptions of covariance stationarity on X_1 and X_2 (see Appendix A). The magnitude of a given Granger causality interaction can be estimated by the logarithm of the corresponding F-statistic.

This concept can be readily extended to the multivariate case by estimating an N-variable VAR model. In this case, X_2 Granger-causes X_1 if knowing X_2 reduces X_1's prediction error when the activities of all other variables $X_3 \ldots X_N$ are also taken into account. Multivariate analyses can provide robustness to false positives in cases of common input. In a system in which X_1 and X_2 are both influenced by X_3 but are otherwise independent, a bivariate model of X_1 and X_2 may wrongly suggest the existence of a causal relationship between X_1 and X_2. A multivariate model including all three variables would not, since knowing $X_{1(2)}$ would not help predict $X_{2(1)}$ in the context of knowing X_3. For this reason this paper adopts a fully multivariate approach.

Significant Granger causality interactions between variables can be represented as edges in a graph, allowing the application of graph-theoretic techniques. Since Granger causality is in general not symmetric, these edges will be directed. Graphical representation can be used to summarize causal connectivity in several novel ways:

- **Causal density.** The *causal density* (cd) of a network's dynamics reflects the fraction of interactions among nodes that are causally significant. A set of independent nodes will have low cd, as will a network in which all nodes have identical dynamics. Causal density is defined as $cd = gc/(2N(N-1))$, where gc is the total number of significant causal links observed, and N is the network size. A related quantity, the *unit* causal density $cdu(i)$, is defined as the total number of significant causal links involving node i. For unweighted graphs (graphs in which all edges are equivalent), $cdu(i)$ is equivalent to the degree of node i, i.e. the total number of afferent (out-degree) and efferent (in-degree) connections.
- **Causal flow.** The *causal flow* (cf) of a node i in a Granger-causality graph is defined as the *difference* between its out-degree and in-degree.† The *causal flow profile* of the graph is the vector $F = [cf(1), cf(2), \ldots, cf(n)]$ for nodes $i = 1 \ldots n$. This profile identifies nodes that have distinctive causal effects on network dynamics: A node with highly positive cf exerts a strong causal influence over the network (a *causal source*); a node with negative cf may be said to be a *causal sink* of the network. Note that the present definition uses the difference between out-degree and in-degree, rather than the ratio (see [19]), in order to obtain a measure unaffected by how many balanced efferent/afferent connections a node may have.

† 'Causal flow' is distinct from a previous graph-theoretic definition of 'flow' in graph theory, which refers to the problem of assigning non-negative values to directed edges such that total inflow is equal to total outflow for all nodes except two [8].

- **Causal reciprocity.** The *causal reciprocity* (c_{recip}) of a Granger-causality graph is defined as the fraction of edges for which a directly reciprocal edge exists. This measure provides an estimate of the degree of functional reciprocity sustained by a network. It is analogous to the quantity f_{recip} calculated for anatomical networks [20].
- **Causal disequilibrium.** A more informative but more complex measure of functional reciprocity is given by *causal disequilibrium* (*cde*). This quantity reflects the *deviation* of directed dynamical interactions from reciprocity, measured over all network bipartitions. It is defined as

$$cde = (\frac{1}{Np})\sum_{i=1}^{Np}(|gc_{A_iB_i} - gc_{B_iA_i}|)/(N_{A_i}N_{B_i})$$

  where Np is the total number of bipartitions of the graph, $gc_{A_iB_i}$ is the number of links from nodes in partition A_i to nodes in its complement B_i (and *vice-versa* for $gc_{B_iA_i}$), and N_{A_i} (N_{B_i}) is the number of nodes in partition A_i (B_i). Calculation of *cde* allows identification of network partitions which maximize or minimize cross-partition causal interactions, which may correspond to informative functional decompositions. Causal disequilibrium is related to causal flow: networks with high causal disequilibrium will tend to have causal flow profiles with high variance (see Sections 5.5 and 5.6). It is also related to the property of node symmetry as calculated for structural graphs [19].

'Weighted' versions of *cd*, *cdu*, cf, and *cde* can be calculated by scaling the contribution of each connection by its magnitude, which is given by the logarithm of the corresponding F-statistic. The resulting values are labelled cd_w, cdu_w, cf_w, and cde_w, respectively. There is no weighted version of c_{recip}. MATLAB (Natick, MA) routines for calculating and graphically representing these measures are provided on the author's website `www.nsi.edu/users/seth`.

3. Simulation model of target fixation

The modulation of causal connectivity by behavior and environment is illustrated here by analysis of a simulation model of target fixation. This model has been previously described and analyzed using an information-theoretic approach [4, 5], the results of which will be compared to the present analysis in Section 5. Full details of the model are given in [4]; here only a minimal set of features are described.

The model consists of a simulated head/eye system in which a neural network controls the movements of a head (H) and an eye (E) in an x, y plane (figure 1). Good performance in the model corresponds to fixation of a target (T) with H and E aligned. Networks consist of $N = 32$ neurons and $K = 256$ weighted connections (inhibitory and excitatory), with a per-neuron in-degree (*indeg*) of 8. Network dynamics are implemented as a continuous system in which neuron output is a sigmoidal function

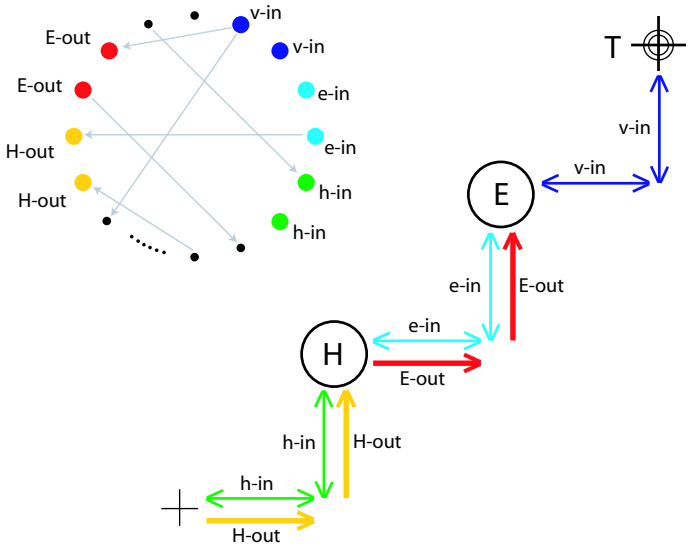


Figure 1. Simulation model of target fixation. A neural network (top left, $N = 32$, $K = 256$) controls 'head' (H) and 'eye' (E) movements to fixate a target (T). Only 5 out of 22 interneurons (INs, black) are shown, and only a small number of possible connections are illustrated. Visual inputs (v-inputs, blue) respond to displacement of gaze direction from the target. 'Proprioceptive' inputs reflect the offset between H and E (e-inputs, cyan), and the displacement of H from a central axis (h-inputs, green). Motor outputs control H (H-outputs, yellow), as well as movements of E relative to H (E-outputs, red).

of the sum of its inputs. In a given network, six neurons are sensory inputs and four neurons are motor outputs (see figure 1 for details). The remaining 22 neurons are 'interneurons' (INs). Note that input neurons are also modulated by the remainder of the network ($indeg = 8$ for *all* neurons).

Evolutionary algorithms were used to generate networks able to support target fixation in a 'simple' context ϕ_S, in which the target is stationary. A second set of networks were evolved in a 'complex' context ϕ_C, in which the target drifts and jumps, and in which movements of H and E are constrained by more complex parameters including, for example, time-lags and differing momenta. These parameters constitute the 'phenotype' of the model. During evolution, connection and weight distributions of networks were allowed to mutate but N, K, and $indeg$ were maintained.

Ten networks were evolved in each context. Those evolved in ϕ_S are referred to as S-networks, and those evolved in ϕ_C as C-networks. Each evolved network was analyzed in both contexts, as well as in a random noise context (ϕ_R). Ten randomly generated networks were also analyzed in each context (R-networks: $N = 32$, $K = 256$, $indeg = 8$). The present analysis therefore consists of nine different combinations of network type and context, with ten sets of neural dynamics for each combination. These combinations are labelled by the shorthand: CC for a C-network in ϕ_C, CS for a C-network in ϕ_S, SC for an S-network in ϕ_C, and so on.

Behavioral results from the model, described in [4], show that C-networks are

able to fixate targets in both contexts, whereas S-networks show poor fixation in ϕ_C. Moreover, C-networks are able to compensate for unexpected perturbations to the head position, whereas S-networks are not. These results suggest that network optimization in a rich sensorimotor environment can facilitate the emergence of robust and adaptive behavior. In the following analysis, these behavioral observations are related to a causal connectivity analysis of the corresponding network dynamics.

4. Causal connectivity analysis

4.1. VAR model estimation

Separate (multivariate) VAR models were estimated for each of the 90 sets of neural dynamics. To keep the number of parameters to be estimated within a reasonable range, from each set a subset was excised corresponding to the 6 input and 4 output neurons. Each time-series was pre-treated by first-order differencing (see Appendix A) and by removing the first 25 values. The resulting matrices (10x575) are referred to as 'activity profiles'. It should be emphasized that these profiles reflect rates of change of neural activity rather than absolute values.

To select the appropriate model order (i.e. the number of lagged observations to include), the Bayesian Information Criterion (BIC) was used [21]. The model order p_{min} resulting in the lowest BIC represents the best compromise between accuracy of fit and parameter parsimony. For an observation of length T and model order p, $\text{BIC}_p = log(\det(\hat{\Omega})) + log(T)n_{est}/T$, where $\hat{\Omega}$ is an estimate of the covariance matrix of the residuals of the corresponding VAR (see equation 1), and n_{est} is the number of freely estimated parameters (n^2 x p for an n-dimensional VAR). For each activity profile, BIC_p for $p \in [2, 10]$ was calculated by estimating a 10-dimensional VAR model for each model order, using the method of ordinary least squares. The resulting values of p_{min} differed within and between conditions, ranging from 2 to 4. To maintain consistency across conditions, $p = 4$ was chosen for all subsequent analyses. To confirm that the temporal relationships among variables were captured by each $p = 4$ model, it was verified that the residuals were serially uncorrelated ($p < 0.01$ in all cases, Ljung-Box 'Q' statistic [22]). It was also verified that each model captured most of the variance in the data (R^2_{adj} in all cases lay in the range [0.5,0.8] [23]).

4.2. Causal connectivity

Given a $p = 4$ VAR model for each activity profile, significant Granger causality interactions between input neurons and output neurons were calculated using an F-test corrected for multiple comparisons ($p < 0.01$). Causal interactions between input neurons were not considered, nor were interactions between output neurons. Figure 2 shows representative casual connectivity graphs from each of the nine conditions. Each arrow indicates a significant Granger causality interaction and the width of the arrow is scaled by the magnitude of the interaction, as determined by the logarithm of the

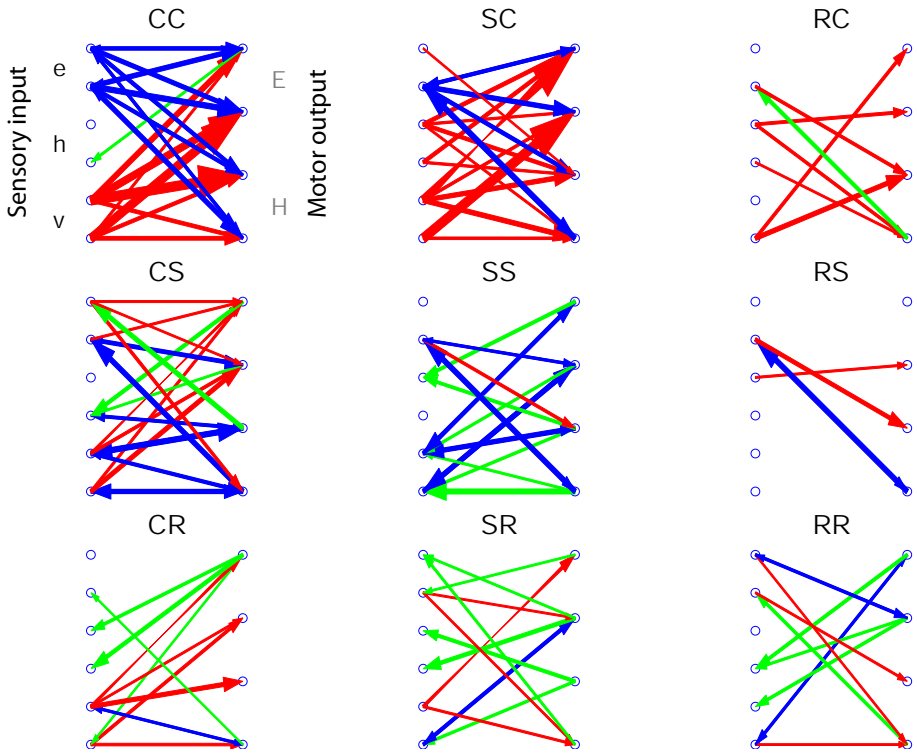


Figure 2. Causal connectivity graphs for nine representative activity profiles. CC: C-network in context ϕ_C, CS: C-network in ϕ_S, CR: C-network in ϕ_R, SC: S-network in ϕ_C, SS: S-network in ϕ_S, SR: S-network in ϕ_R, RC: R-network in ϕ_C, RS: R-network in ϕ_S, and RR: R-network in ϕ_R. Each panel shows six input neurons (black lettering, v: v-inputs, e: e-inputs, h: h-inputs) and four output neurons (grey lettering, H: H-outputs, E: E-outputs). The labelling of the top-left panel applies to all panels. Red arrows show input to output causal connections, green arrows show output to input causal connections, and blue arrows show reciprocal causal connections. Causal links are included if significant at the $p < 0.01$ level (Bonferroni-corrected F-test). The width of each arrow and size of the arrowhead reflect the magnitude of the causal influence, as determined by the logarithm of the corresponding F-statistic.

corresponding F-statistic. It bears repeating that a significant causal interaction from A to B shows a statistical relation, i.e. that the activity of B can be better predicted by including past observations of A in a multivariate linear model of the dynamics of all (10) neurons in the network.

4.2.1. C-networks. Figure 2 reveals several features of dynamical organization. For example, causal graphs for C-networks tested in ϕ_C (CC) have strong causal connectivity from v-inputs to motor outputs, suggesting that their dynamics are driven largely by visual signals. There is also strong reciprocal causal connectivity between the motor outputs and (proprioceptive) e-inputs, suggesting that the corresponding dynamics are modulated by the displacement between head position and eye position, and little if any among motor outputs and (proprioceptive) h-inputs, suggesting that these inputs are not functionally significant in this case. This profile is altered when the same

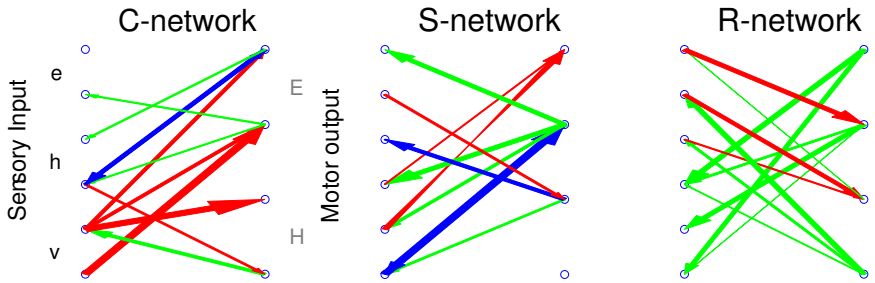


Figure 3. Connectivity profiles for a representative C-network, S-network, and R-network. Each panel shows six input neurons and four output neurons (labelled as in figure 2). Red arrows show input to output connections, green arrows show output to input connections, and blue arrows show reciprocal connections. Line widths and arrowhead sizes are scaled by the corresponding connection strengths.

network is tested in contexts ϕ_S and ϕ_R (CS and CR respectively): Causal projections from visual inputs are weaker, and the fraction of unidirectional output to input connections increases. These observations reflect a network that is organized to be driven preferentially by visual input in rich sensorimotor contexts, with additional modulation from proprioceptive sensors, particularly e-inputs.

One may compare causal graphs with the underlying network structure, as well as with dynamical covariances among network elements. Figure 3 shows the (direct) structural connectivity among input neurons and output neurons for the networks whose causal connectivity is depicted in figure 2. Not surprisingly, there is some overlap between structural and causal representations. For example, for the C-network, the strong causal connectivity from v-inputs to motor outputs is reflected in the structure. But there are also differences: Structurally, the C-network shows only one reciprocal connection, whereas the corresponding causal graph (CC) shows strong reciprocal connectivity between motor outputs and e-inputs. More generally, causal connectivity graphs show the influence of interactions among network structure, phenotype, and environment (hence the differences among panels CC, CS, and CR), whereas representations of network structure are necessarily invariant to these contextual influences. Note that structural graphs omit indirect structural connections via INs. One reason for this is that in all networks, all neurons are structurally connected to all other neurons via many differently-weighted paths of diverse lengths. Evaluating only direct connections between the observed neurons provides a parsimonious representation that is consistent with the causality analysis (and the following covariance analysis): Each considers the same restricted set of variables.

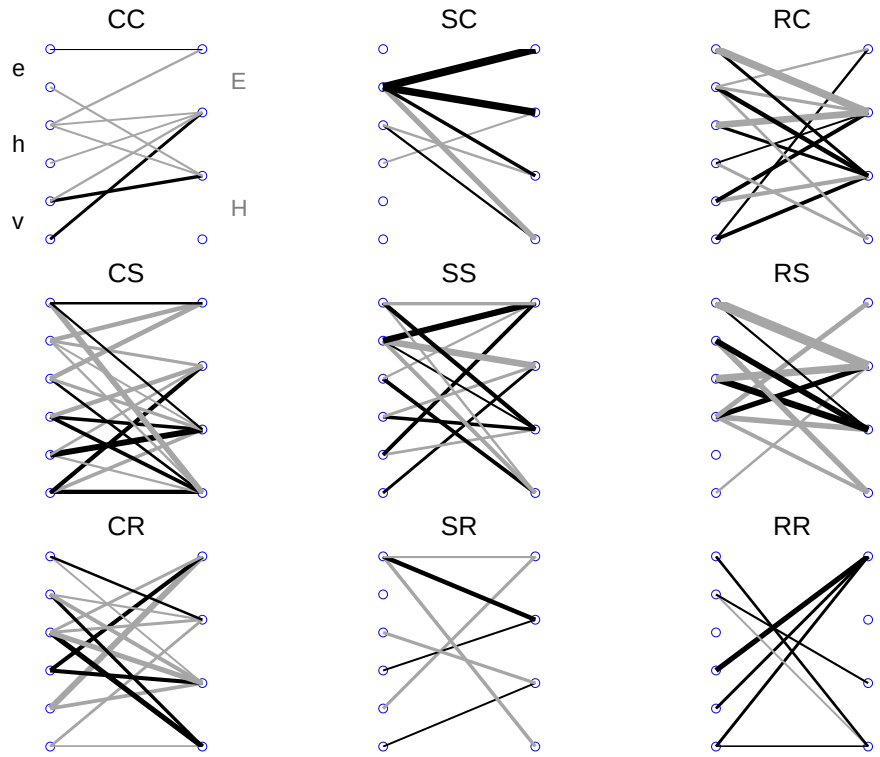


Figure 4. Covariance patterns by condition (labelled as in figure 2). Two neurons are connected if their correlation is significant at the $p < 0.01$ level (Bonferroni-corrected t-test). Only covariances between input and output neurons are shown. Each panel shows six input neurons and four output neurons (labelled as in figure 2). The width of each line reflects the magnitude of the covariance. Light grey lines indicate positive covariance, dark grey lines indicate negative covariance.

Figure 4 shows covariances among input neurons and output neurons for the same activity profiles as in figure 2. Two neurons are connected if their correlation is significant at the $p < 0.01$ level (t-test corrected for multiple comparisons) and the width of each link indicates the strength of covariance (thicker lines correspond to higher covariance values). Again, these graphs provide an impoverished representation of network dynamics as compared to causal connectivity. Although covariance graphs vary with context (compare CC with CS and CR), since covariances are by definition symmetric they do not give any indication of directionality. Also, covariance patterns clearly differ from the corresponding causal connectivities. For example, covariances in condition CC are sparse, give less emphasis to v-inputs, and do not distinguish between proprioceptive e-inputs and h-inputs. By contrast, the covariance pattern in condition CS is dense and contains little obvious organization.

4.2.2. S-networks and R-networks. Networks adapted to different contexts show different patterns of causal connectivity. In contrast to the C-network, S-network motor outputs are driven strongly by h-inputs as well as by v-inputs in ϕ_C (figure 2). This suggests that the S-network is less specifically sensitive to visual signals than the C-

network in this context. Also, the C-network has more reciprocal causal connections between e-inputs and motor outputs than the S-network (or R-network), suggesting that proprioceptive visual signals play a greater functional role in the C-network. Despite these differences, both networks are affected by changes in context in a similar way, showing a decrease in input→output causal connectivity as the testing environment is switched from rich (ϕ_C) to simple (ϕ_S) to random (ϕ_R) (figure 2, middle panels). Finally, R-networks show sparse and disorganized causal connectivity irrespective of the environment of testing (figure 2, right panels).

These patterns of causal connectivity can be compared with the corresponding structural and covariance patterns (figures 3 and 4). S-network causal projections from motor outputs to e-inputs are not reflected in the S-network structure. The strong causal projections from v-inputs are also largely absent in the corresponding structure graph. Covariance patterns for the S-network are highly variable among conditions and - as for the C-network - differ from the corresponding causal patterns. For example, covariances in condition SC lack connections from v-inputs. R-networks, as expected, show little organization in either structure or covariance in all conditions.

Taken together, comparisons among causal, structural, and covariance patterns suggest that causal connectivity provides a comparatively rich representation of network dynamics. Unlike covariance patterns, causal connectivity graphs are directed, and unlike structural graphs, they are sensitive to context. Moreover, causal patterns appear to provide an intuitive interpretation of the relation between network dynamics and behavior, for example by highlighting the importance of visual signals in rich sensorimotor environments.

4.3. Group analysis of causal connectivity

To test the consistency of the above observations, 'composite causal connectivity' graphs were derived for each set of 10 activity profiles from each condition. Each panel in figure 5 shows the causal interactions that are reliably present in each set of profiles. Black indicates a strong presence of a given connection across all activity profiles for the condition, white indicates that a connection is never present. For example, the top-left panel shows that causal connections from v-inputs to motor outputs are strong across all activity profiles for condition CC, whereas connections from h-inputs to motor outputs are less prevalent.

These composite patterns warrant several observations consistent with figure 2. The causal connectivity of a network depends on the environment in which it operates. C-networks in ϕ_C tend to have rich causal interactions from v-inputs and e-inputs to motor outputs, as well as a high proportion of causal connections from motor outputs back to e-inputs. By contrast, connections from h-inputs to motor outputs are rare. This pattern changes with progressive simplification of the environment (CS and CR). In condition CS, H-outputs are less responsive to v-inputs, and h-inputs are less disconnected from motor outputs. In condition CR, there remains only a general tendency for input neurons

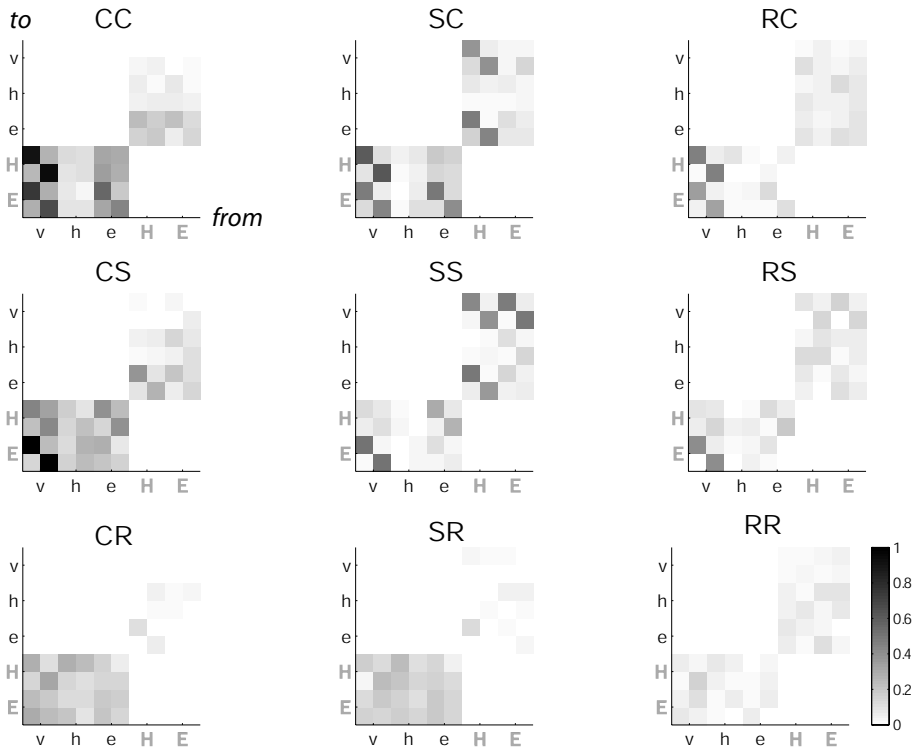


Figure 5. Composite causal connectivity by condition (conditions are labelled as in figure 2). Each panel shows the causal interactions that are reliably present in each set of 10 activity profiles from each condition. The composite causal connectivity between a pair of neurons is the sum of the magnitudes of causally significant interactions over all activity profiles, divided by the largest value of this sum across all connections and all conditions. Horizontal axes show source neurons and the vertical axes show target neurons. The labelling of the top-left panel applies to all panels. Black indicates a strong presence of a given connection across all profiles, white indicates that an interaction is never present. Six input neurons are shown (black lettering, v: v-inputs, e: e-inputs, h: h-inputs), and four output neurons (grey lettering, H: H-outputs, E: E-outputs).

to drive output neurons. The same general point holds for S-networks and R-networks: Environmental simplification leads to a reduction in the range of causal interactions reliably sustained by a network.

The environment to which a network has adapted influences the richness of causal connectivity. For a given testing context, C-networks tend to have a higher density of input→output causal interactions than either S-networks or R-networks. S-networks, however, show strong causal connections from H-outputs to v-inputs (except in ϕ_R), implying that visual signals are predictable from preceding head movements. This suggests that for S-networks (unlike C-networks) the generation of behavior does not involve complex coordination of head and eye movements. As expected, R-networks in ϕ_R show no reliably present causal connections. However, in conditions RC and RS there *are* strong causal connections between v-inputs and motor outputs, which demonstrates that the sensorimotor correlations imposed by an environmental and phenotypic context

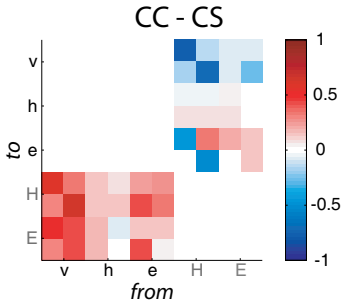


Figure 6. Effect of context on causal connectivity. Shown are composite causal connectivity differences between conditions CC and CS, scaled by magnitude and labelled as in figure 5. Warm colors (cool colors) show causal interactions that are reliably stronger (weaker) when a network is switched from context ϕ_S to ϕ_C. The color scale is normalized to the largest difference in composite causal connectivity.

can be sufficient to generate reliable causal interactions even in randomly connected networks.

The effect of context on network dynamics is illustrated further in figure 6, which shows differences between composite causal connectivity patterns for conditions CC and CS. Warm colors and cool colors respectively indicate the causal interactions that are reliably strengthened or weakened, when a C-network is switched from ϕ_S to ϕ_C. Consistent with figure 5, switching the network to a richer sensorimotor context leads to greater causal connectivity from sensory inputs to motor outputs. Connections from motor output to e-inputs are also generally stronger in the rich environments, with the exception of two connections from H-outputs to e-inputs that may reflect a greater dissociation of head and eye movements in ϕ_C. Finally, connections from motor outputs to v-inputs are consistently weaker in the richer context, supporting the suggestion that visual signals are less predictable from prior head movements in rich environments.

4.4. Causal flow

Causal flow profiles are consistent with the foregoing analyses. Figure 7 shows that C-networks in ϕ_C have a prevalence of causal connections extending from v-inputs (these neurons are causal *sources*). Motor output neurons have predominantly incoming causal connections (causal *sinks*). Environmental simplification (CS and CR) lessens the prominence of this pattern. S-networks and R-networks in ϕ_C show the same general pattern of sources and sinks, but v-inputs are less prominent in these cases. In contrast, flow profiles for graphs based on covariances or structural connectivity, are generally flat (not shown).

4.5. Causal density

Figure 8(a) shows that weighted causal density cd_w falls significantly as the environment is simplified, and that the highest overall values are for S- and C-networks in ϕ_C.

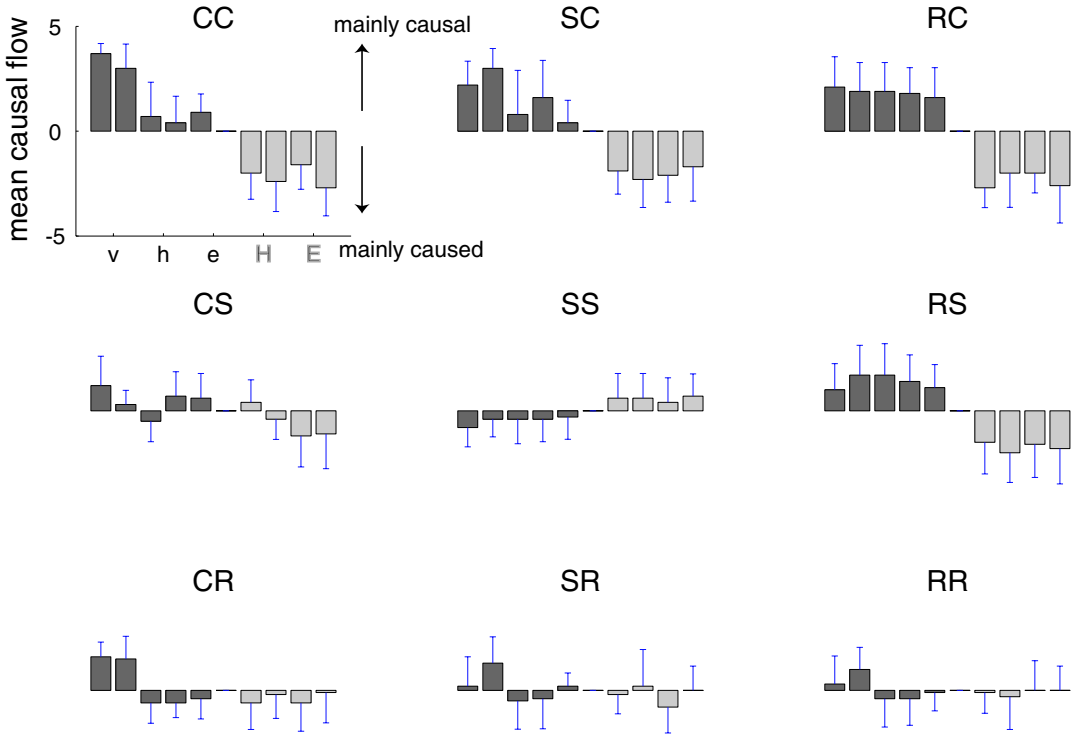
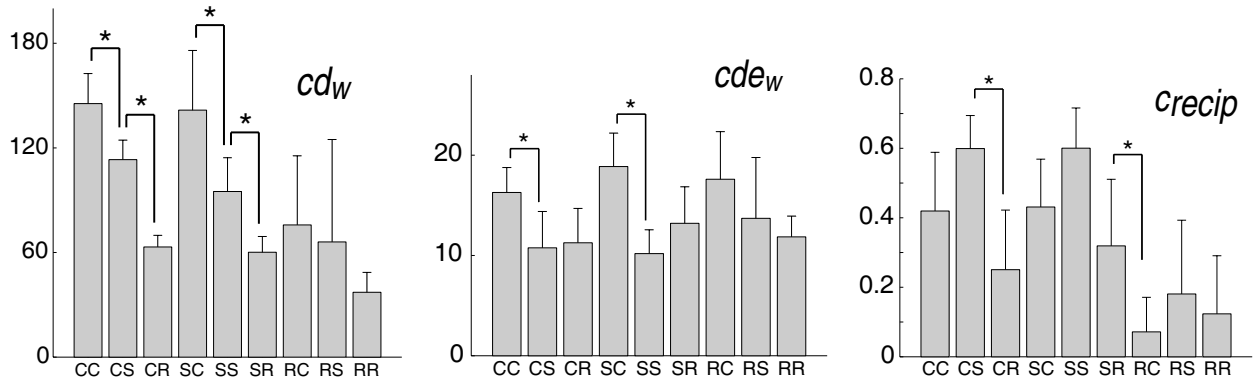


Figure 7. Causal flow profiles for each condition (conditions are labelled as in figure 2). Each panel shows the mean and standard deviation causal flow $cf(i)$ for each neuron (the difference between the out-degree and the in-degree), averaged across 10 activity profiles for each condition. Dark grey bars correspond to inputs (black lettering: v, h, e), light grey to outputs (grey lettering: H, E). The labelling of the top-left panel applies to all panels.

Random R-networks show low cd_w regardless of the testing context. A qualitatively similar pattern was observed for unweighted causal density cd. Therefore, for networks adapted to a structured environment (S- and C-networks), behavior within a rich environment evokes high causal density in the corresponding network dynamics.

4.6. Causal disequilibrium

Figure 8(b) shows how weighted causal disequilibrium cde_w varies across conditions. The results are broadly similar to cd_w: Rich environments evoke high cde_w as compared to simple or random environments, for both C-networks and S-networks (CC and SC respectively). Interestingly, unlike cd_w, cde_w does not distinguish between ϕ_S and ϕ_R for either C-networks or S-networks. This suggests that while imposing simple sensorimotor correlations can increase the density of causal interactions (as compared to random activation, see figure 8a), this increase is not accompanied by a corresponding increase in the deviation from overall reciprocity of these interactions. For such an increase reliably to take place, behavior in a rich sensorimotor environment (ϕ_C) is necessary.



(a) cd_w (b) cde_w (c) c_{recip}

Figure 8. (a). Mean and standard deviation weighted causal density cd_w for each set of 10 activity profiles from each condition (conditions are labelled as in figure 2). (b). Weighted causal disequilibrium cde_w. (c) Causal reciprocity c_{recip}. Asterisks shows significant differences by two-tailed t-test ($p < 0.01$).

4.7. Causal reciprocity

Figure 8(c) shows that causal reciprocity c_{recip} follows a pattern that differs from both cd_w and cde_w. Highest values are found for non-random networks tested in ϕ_S, and lowest values are found for random networks in any context. This suggests that c_{recip} responds to uniformity and regularity in network dynamics.

4.8. Predicting the influence of network lesions

A potential application of causal connectivity analysis is prediction of the effects of network lesions on behavior. To test this possibility, a C-network (shown in figure 3) was retested in ϕ_C after removing all connections to and from each of the six input neurons in turn. Each lesioned network was retested 50 times.

Figure 9(a) shows the mean residual fitness for each lesioned network, as a proportion of the mean fitness value obtained by the non-lesioned network. Lesions to v-inputs had severe effects on fitness, lesions to e-inputs had moderate effects, and lesions to h-inputs had very mild effects. These results are consistent with the causal connectivity of the network (figure 2, top-left panel): v-inputs have strong causal connections to the motor outputs, e-inputs have reciprocal causal connections with the outputs, and h-inputs are mostly disconnected from the outputs. By contrast, the corresponding structural representation is only partly consistent with the lesion data. As figure 3 shows, there are strong projections from v-inputs to output neurons; however, h-inputs also have widespread (albeit weaker) structural connectivity with output neurons, suggesting that lesions to h-inputs would have substantial effects on fitness, which is not the case. Finally, e-inputs have sparse and weak structural connectivity with outputs, which is at odds with the significant effects on fitness of lesions to these inputs. There is no obvious similarity between the lesion data and the corresponding covariance graph

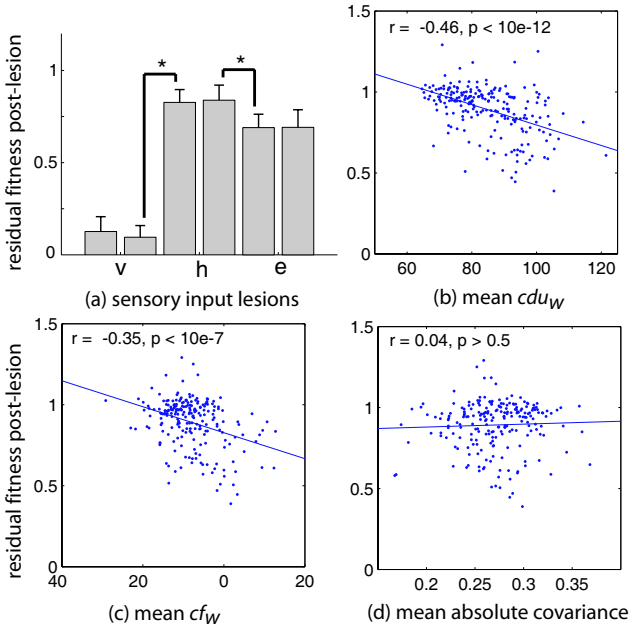


Figure 9. (a) Mean post-lesion fitness of a representative C-network in ϕ_C following sensory input lesions, shown as a proportion of the fitness of a non-lesioned network (v: v-inputs, e: e-inputs, h: h-inputs). Asterisks shows significant differences by two-tailed t-test ($p < 0.01$). (b-d) Residual fitness following lesions to INs of C-networks, as a proportion of the fitness of the intact network, plotted against (b) mean cdu_w of the IN, (c) mean cf_w of the IN, and (d) mean absolute covariance of the IN with the remainder of the network. Intact networks and lesioned networks were each tested 12 times in ϕ_C. All 22 INs were lesioned in turn for all 10 C-networks. Each panel shows Pearson's correlation coefficient (r) as well as the corresponding p-value.

(figure 4, top-left panel).

This analysis was extended by calculating the causal connectivity among all 32 neurons in a C-network during behavior in ϕ_C.§ Each of the 22 'intermediate' neurons (INs) were then lesioned, and the fitness consequences assessed. Each lesioned network was tested 50 times, allowing identification of the four INs that had the most severe fitness effects when lesioned, and the four that had the mildest. Figure 10 shows, for each of these neurons, the causal connections involving the IN and the sensory and motor neurons. For severe lesions (top row), the corresponding INs mediate input-output causal pathways; that is, they are causally affected by sensory input, and they causally affect motor output. For mild lesions, the corresponding INs are causally isolated (they are caused, but they do not cause). Neither structural representations nor covariance representations showed equivalent correspondences. INs associated with mild fitness effects are not structurally isolated: all four project to input neurons, and #7 and #22 project also to output neurons. One IN associated with severe fitness effects (#10) had no significant covariances with either input or output neurons, and, conversely, one IN

§ This required estimating a 32-dimensional $p = 4$ VAR. The mean R^2_{adj} was 0.9, and the residuals were uncorrelated ($p < 0.01$, Ljung-Box 'Q' statistic).

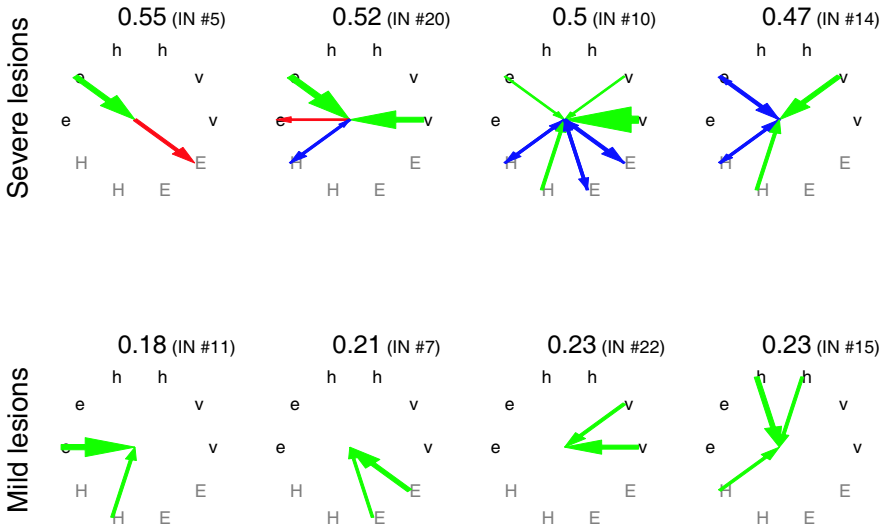


Figure 10. Causal connectivity of selected 'intermediate' neurons (INs) with sensory and motor neurons. Connections among sensory inputs (black lettering) and motor outputs (grey lettering) are omitted for clarity. Red (green) arrows indicate causal connections from (to) the IN, blue arrows indicate reciprocal causal connections. Arrow width and arrowhead size are scaled by the magnitude of causal influence (the logarithm of the corresponding F-statistic). Each graph is labelled by the index of the corresponding IN (for example, IN #5) as well as by the fractional fitness decrement caused by lesioning the IN (for example, 0.55).

associated with mild fitness effects (#15) covaried strongly with both input and output neurons.

This analysis was generalized across all INs in all ten C-networks (see figure 9b-d), showing that post-lesion fitness correlates with both cdu_w and cf_w of the lesioned IN. By contrast, there is no correlation between post-lesion fitness and mean absolute covariance of the lesioned IN with the remainder of the network. These results confirm that the causal connectivity of a neuron is a useful predictor of the dynamical and behavioral consequences of network damage.

5. Discussion

This paper has described a method for characterizing directed dynamical interactions within intact neural systems during behavior. Causal connectivity analysis is based on Granger causality [6], which formalizes the ability of a linear autoregressive model to elucidate significant causal relations within a set of variables. Given such a model, graph-theoretic concepts can be applied to characterize the resulting patterns of causal connectivity. Specifically, one may measure the causal density, causal reciprocity, causal disequilibrium and causal flow profile of a network, and identify causal sinks and causal sources.

Throughout this paper the term 'causality' has been used following Wiener [12] and Granger [6]. It is important to stress that this concept of causality is statistical rather

than physical. If a system is only partially observed, there are several situations in which Granger causality may not correspond to physical causal chains. For example, A may enhance the predictability of B if both A and B are driven by a common (unobserved) input C, or if A physically causes an (unobserved) intermediate process D which itself physically causes B. ‖

This sensitivity to unobserved variables is true in general for dynamical analyses, and can in fact be useful: many network systems cannot be observed *in toto*, especially when the network does not form an isolated system; biological networks, for example, are embodied in phenotypes and embedded in environments [1, 2]. Causal connectivity analysis, in virtue of its sensitivity to indirect interactions, remains able to describe causal influences among observable elements of a system, even if the activity of these elements is physically caused in part by factors outside the scope of observation.

An alternative approach to determining causal interactions in a network is to measure the effects of selective perturbations. For example, Tononi and Sporns stimulate selected subsets of a network with Gaussian noise and interpret the resulting mutual information between the subset and the rest of the network as a measure of their causal connectivity [10]. Similarly, Keinan *et al* assess the functional contribution of individual network elements by measuring network performance following lesions to subsets of elements [11]. While these approaches are in principle robust to artifacts induced by common input, in practice their use is restricted to situations in which networks can be repeatedly and reversibly perturbed, which for many biological systems is currently not possible. Nor are these approaches suitable for analyzing the influence of behavior and environment on network dynamics, since such analyses require data derived from *intact* neural networks in different contexts. Another approach which does not require network perturbation is 'structural equation modeling' (SEM) [24, 25]. However, whereas causal connectivity analysis infers causal interactions directly from data, SEM tests *a priori* hypotheses about causal relationships. It may be useful to compare these approaches for situations in which structural connectivity is known.

There are many opportunities to extend the graph-theoretic components of causal connectivity analysis. Kötter and Stephan [19] have recently proposed a series of 'network participation indices' for analyzing neuroanatomical connectivity: Some of these may have useful interpretations for causal graphs. It may be useful to look for 'small-world' properties in causal connectivity graphs. Small-world networks combine high local clustering with short characteristic path lengths [26], and are usually identified by analysis of network topology. The question remains open: What network topologies would support small-world causal connectivity, and what functional properties would such causal connectivity provide?

‖ As was remarked in the Introduction, if common inputs or intermediate variables (e.g. C and D in the above example) *are* part of the observed system, a multivariate Granger causality analysis *will* reveal the causal interactions mediated by these variables, instead of the indirect causal interactions that they support.

5.1. Linearity and stationarity

Causal connectivity analysis involves linear modeling of stationary time-series. The present application investigated the validity of assumptions of stationarity and linearity and found them satisfactory. In general, linear models are simple to estimate, and for signals generated by a Gaussian process are superior to any other estimator [14]. Fortunately, large-scale neurodynamics such as interactions among distinct brain areas appear to be well described by linear models [20, 25, 27]. In contrast, interactions among individual spiking neurons are generally nonlinear [28] and future work may usefully address the application of suitable nonlinear modeling techniques [29, 30] to assessing causal interactions in such systems.

Non-stationarities in time-series data can be addressed in several ways, for example by differencing (see Appendix A), or by the use of adaptive VAR models that have time-varying coefficients (in these models, a time-series is broken into short segments which may approximate stationarity). An advantage of adaptive models is that they can identify changes over time in causal influences [17], however they generally require a greater number of parameters to be estimated with less data [14], and as such may provide a weaker basis for statistical inference of causal interactions.

5.2. Environmental and behavioral modulation of network dynamics

This paper illustrated causal connectivity analysis using a simulation model of target fixation [4]. The method provided several insights into network dynamics and their modulation by behavior. In the model, networks evolved and tested in rich environments (C-networks in ϕ_C, figure 2) were driven largely by visual signals (v-inputs), with proprioceptive modulation (e-inputs). In these networks, as expected, the visual inputs are causal sources and the motor outputs are causal sinks. Networks adapted to comparatively simple environments (S-networks) were less specifically sensitive to visual signals, and less responsive to modulation by e-inputs (figure 2). Unlike C-networks, S-networks showed strong causal projections from head motor outputs (H-outputs) to v-inputs, suggesting that for these networks the generation of behavior did not involve a complex coordination of head and eye movements. In general, these features were not recapitulated in either structural or covariance representations of the same networks (figures 3 and 4).

For all networks, including randomly connected networks, behavior in a comparatively rich sensorimotor environment tended to evoke input-to-output causal interactions in network dynamics. This trend was noted in representative causal graphs (figure 2), summaries of all networks (figures 5 and 6), causal flow profiles (figure 7), and mean causal disequilibrium (figure 8).

A useful summary of global causal connectivity is given by *causal density*, which reflects the fraction of interactions among network elements that are causally significant. In general, dense causal interactivity in a network signifies that nodes are both globally coordinated in their activity (in order to be useful for predicting each other's activity)

and at the same time dynamically distinct (so that nodes contribute in different ways to these predictions). High causal density may therefore reflect an intuitive property of complex systems in general, i.e. that they are intermediate between completely ordered systems and completely random systems [31, 5]. In our model, adaptation to a structured sensorimotor environment evoked dense causal interactions when compared to random networks, and behavior within such a context evoked dense causal interactions when compared to behavior in a simple environment (figure 8). These observations are consistent with the hypothesis that adaptation to complex environments yields complex systems [3, 4].

Causal connectivity analysis was able to predict the functional consequences of network lesions (figures 9 and 10). Lesions to input neurons with strong causal projections to outputs yielded correspondingly large fitness deficits. Lesions to interneurons (INs) that mediated input-to-output causal pathways yielded large fitness decreases, whereas INs associated with mild post-lesion fitness decreases tended to be causally isolated. Finally, post-lesion fitness correlated with both cdu_w and cf_w of the lesioned IN, but not with the covariance of the IN with the remainder of the network. It is an exciting possibility that similar analyses will allow assessment of the context-specific functional robustness of biological as well as simulated neural networks.

Although the present model is not intended as a realistic model of the neurobiology of eye movements (see [32, 33]), the results nevertheless suggest several hypotheses. They predict that head-eye coordination in rich sensorimotor environments will be accompanied by increased causal density and increased input→output causal flow in the underlying neural dynamics. They suggest that signals reflecting head-eye displacement will affect performance more than signals reflecting head position in a global reference frame. Lastly, they predict that causal connectivity will be differentially affected by adaptation and by behavior. For example, adaptation to a structured environment (whether simple or rich) will yield neural mechanisms capable of displaying high causal density, but behavior within a rich environment will be necessary for this density to be expressed.

A previous study [4] explored how behavior and environment modulated the 'neural complexity' of the same simulation model described in this paper. Neural complexity is an information-theoretic measure of global network dynamics which reflects the extent to which a system balances dynamical integration and dynamical segregation [31]. It was found that neurally complex dynamics accompanied adaptive behavior in rich sensorimotor environments. Other related dynamical measures, such as entropy (the level of statistical independence among network elements) and integration (the converse), did not show such correlations with adaptive behavior (for details of these measures see [20, 5]). The present results are consistent with these findings inasmuch as causal density may also reflect system complexity. Nevertheless, causal density and neural complexity are distinct concepts and measure system dynamics in very different ways.

Furthermore, a fundamental limitation of the 'neural complexity' study was that

it was not able to provide insight into the causal transactions in networks that supported adaptive behavior: neural complexity, being a global measure, yields only a scalar quantity for each analyzed network. Also, neural complexity is based on the symmetric quantity of mutual information and as such cannot take account of or describe directional interactions in network dynamics. By contrast, the present analysis not only characterizes directed dynamical interactions, but is able to link global descriptions of these interactions (causal density, causal disequilibrium) to the specific causal transactions in networks that underlie their adaptive interactions with behavior and environment.

The present approach may generalize to many cases of neurobiological significance. Environment and behavior can modulate neural dynamics over a number of different timescales. For example, during behavior, active vision involves neural mechanisms which control gaze direction; gaze direction modulates visual input, and this in turn drives subsequent neural dynamics [34]. During the lifetime of an individual, differences in perceptual history can affect the organization of the nervous system [1]; for instance, more neurons in monkey inferotemporal cortex respond to familiar than to unfamiliar stimuli [35]. Although relevant empirical data are hard to collect, it has often been argued that exposure to rich environmental niches over the course of evolution promotes the evolution of complex neural mechanisms [36, 3]. While previous computational and neurorobotic models have addressed each of these phenomena [37, 38, 39, 40, 41], they have mostly lacked a quantitative analysis of the influences of behavior and environment on the directed dynamics of the corresponding neural networks. As suggested above, causal connectivity analysis is well suited for this task by its ability to elucidate causal interactions from data generated by intact neural networks during behavior.

This paper has examined the causal interactions generated by networks acting as sensorimotor controllers. It is likely that other networks which generate dynamics in the service of other functions will show different and distinctive causal connectivity patterns. The internet [42], protein interaction networks [43], and co-authorship networks of scientists [44] all present complex topologies that support diverse dynamical interactions. Extending causal connectivity analysis to these cases should advance our understanding of network dynamics in many different ways. For example, causal sources and causal sinks might suggest potential targets for pharmacological intervention in intracellular metabolic pathways or in protein interaction networks, and causal density patterns may help distinguish between normal and pathological states of these systems.

Acknowledgments

This research was made possible by the Neurosciences Research Foundation, which supports the work of The Neurosciences Institute. For providing many useful comments I thank Raffaella Giacomini, Gerald Edelman, Jason Fleischer, Jeff Krichmar, Joseph Gally, Eugene Izhikevich, and two anonymous reviewers.

Appendix A. Covariance stationarity

VAR modeling and Granger causality analysis assume that time-series are covariance-stationary, i.e. that the mean and variance of the process are constant over time [7]. In the context of an autoregressive model $x(t) = a_1 x(t-1) + \ldots + a_p x(t-p)$ covariance stationarity is assessed by testing for 'unit roots' as solutions of the equation:

$$1 - a_1 z - a_2 z^2 - \ldots - a_p z^p = 0 \tag{A.1}$$

The existence of one or more unit roots, indicated by one or more solutions to this equation lying on the unit circle, implies that the assumption of covariance-stationarity is violated. If the series has a unit root, covariance-stationarity can be induced by differencing, i.e. by transforming the time-series as follows:

$$\Delta x(t) = x(t) - x(t-1) \tag{A.2}$$

All time-series analysed in this paper were pre-treated by first-order differencing and subsequently tested for unit roots: none were found (Dickey-Fuller test, $p < 0.01$).

References

[1] G.M. Edelman. *Neural Darwinism.* Basic Books, New York, 1987.

[2] A. Clark. *Being there: Putting brain, body, and world together again.* MIT Press, Cambridge, MA, 1997.

[3] A.K. Seth. *On the relations between behaviour, mechanism, and environment: Explorations in artificial evolution.* PhD thesis, University of Sussex, 2000.

[4] A.K. Seth and G.M. Edelman. Environment and behavior influence the complexity of evolved neural networks. *Adaptive Behavior*, 12:5–21, 2004.

[5] A.K. Seth and G.M. Edelman. Theoretical neuroanatomy: Analyzing the structure, dynamics and function of neuronal networks. In E. Ben Naim, H. Fraunfelder, and Z. Toroczkai, editors, *Complex networks*, Lecture Notes in Physics, pages 487–518. Springer-Verlag, Berlin, 2004.

[6] C.W.J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37:424–438, 1969.

[7] J.D. Hamilton. *Time series analysis.* Princeton University Press, Princeton, NJ, 1994.

[8] B. Bollobás. *Random graphs.* Academic Press, London, 1985.

[9] J. Pearl. *Causality: Models, reasoning, and inference.* Cambridge University Press, Cambridge, UK, 1999.

[10] G. Tononi and O. Sporns. Measuring information integration. *BMC Neuroscience*, 4:31, 2003.

[11] A. Keinan, B. Sandbank, C.C. Hilgetag, I. Meilijson, and E. Ruppin. Fair attribution of functional contribution in artificial and biological networks. *Neural Computation*, 16:1887–1915, 2004.

[12] N. Wiener. The theory of prediction. In E. Beckenbach, editor, *Modern mathematics for engineers.* McGraw-Hill, New York, 1956.

[13] J. Geweke. Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, 77:304–13, 1982.

[14] C. Bernasconi and P. Konig. On the directionality of cortical interactions studied by structural analysis of electrophysiological recordings. *Biological Cybernetics*, 81:199–210, 1999.

[15] H. Liang, M. Ding, R. Nakamura, and S.L. Bressler. Causal influences in primate cerebral cortex during visual pattern discrimination. *NeuroReport*, 11(13):2875–2880, 2000.

[16] M. Kaminski, M. Ding, W.A. Truccolo, and S.L. Bressler. Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biological Cybernetics*, 85:145–157, 2001.

[17] W. Hesse, E. Möller, M. Arnold, and B. Schack. The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies. *Journal of Neuroscience Methods*, 124:27–44, 2003.

[18] A. Brovelli, M. Ding, A. Ledberg, Y. Chen, R. Nakamura, and S. Bressler. Beta oscillations in a large-scale sensorimotor cortical network: Directional influences revealed by Granger causality. *Proceedings of the National Academy of Sciences, USA*, 101(26):9849–9854, 2004.

[19] R. Kötter and K. Stephan. Network participation indices: Characterizing component roles for information processing in networks. *Neural Networks*, 16:1261–1275, 2003.

[20] O. Sporns, G. Tononi, and G.M. Edelman. Theoretical neuroanatomy: Relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cerebral Cortex*, 10:127–141, 2000.

[21] G. Schwartz. Estimating the dimension of a model. *The Annals of Statistics*, 5(2):461–464, 1978.

[22] G.E.P. Box, G.M. Jenkins, and G.C. Reinsel. *Time series analysis: Forecasting and control.* Prentice Hall, Englewood Clifs, NJ, 1994. Third edition.

[23] N.R. Draper and H. Smith. *Applied regression analysis: 3rd edition.* John Wiley and Sons, New York, 1998.

[24] T. Haavelmo. The statistical implications of a system of simultaneous equations. *Econometrica*, 11:1–12, 1943.

[25] A.R. McIntosh and F. Gonzalez-Lima. Structural equation modeling and its application to network analysis in functional brain imaging. *Human Brain Mapping*, 2:2–22, 1994.

[26] D.J. Watts and S.H. Strogatz. Collective dynamics of 'small world' networks. *Nature*, 393:440–442, 1998.

[27] A.R. McIntosh, C.L. Grady, L.G. Ungerleider, J.V. Haxby, S.I. Rapoport, and B. Horwitz. Network analysis of cortical visual pathways mapped with PET. *Journal of Neuroscience*, 14:655–666, 1994.

[28] E. Izhikevich. Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, 14:1569–1572, 2003.

[29] Y. Chen, G. Rangarajan, J. Feng, and M. Ding. Analyzing multiple nonlinear time series with extended Granger causality. *Physics Letters A*, 324:26–35, 2004.

[30] W.A. Friewald, P. Valdes, J. Bosch, R. Biscay, J.C. Jimenez, L.M. Rodriguez, V. Rodriguez, A.K. Kreiter, and W. Singer. The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies. *Journal of Neuroscience Methods*, 94:105–119, 1999.

[31] G. Tononi, O. Sporns, and G.M. Edelman. A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Science (USA)*, 91:5033–5037, 1994.

[32] T. Shibata, S. Vijayakumar, J. Conradt, and S. Schaal. Biomimetic oculomotor control. *Adaptive Behavior*, 9(4):189–209, 2001.

[33] E.G. Freedman and D.L. Sparks. Coordination of the eyes and head: Movement kinematics. *Exp. Brain Res.*, 131:22–32, 2000.

[34] D.H. Ballard. Animate vision. *Artificial Intelligence*, 48:57–86, 1991.

[35] E. Kobatake, G. Wang, and K. Tanaka. Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. *Journal of Neurophysiology*, 80(1):324–30, 1998.

[36] P. Godfrey-Smith. *Complexity and the function of mind in nature.* Cambridge University Press, Cambridge, 1996.

[37] D. Floreano, T. Kato, D. Marocco, and E. Sauser. Coevolution of active vision and feature selection. *Biological Cybernetics*, 90:218–228, 2004.

[38] A.K. Seth, J.L. McKinstry, G.M. Edelman, and J.L. Krichmar. Visual binding through reentrant connectivity and dynamic synchronization in a brain-based device. *Cerebral Cortex*, 14:1185–99, 2004.

[39] J.L. Krichmar and G.M. Edelman. Machine psychology: Autonomous behavior, perceptual categorization and conditioning in a brain-based device. *Cerebral Cortex*, 12(8):818–30, 2002.

[40] J.A. Fletcher, M. Zwick, and M.A. Bedau. Effect of environmental texture on evolutionary adaptation. In C. Adami, R.K Belew, H. Kitano, and C.E. Taylor, editors, *Proceedings of the Sixth International Conference on Artificial Life*, pages 189–199, Cambridge. MA, 1998. MIT Press.

[41] A.K. Seth. The evolution of complexity and the value of variability. In C. Adami, R.K Belew, H. Kitano, and C.E. Taylor, editors, *Artificial Life VI: Proceedings of the Sixth International Conference on the Simulation and Synthesis of Living Systems*, pages 209–221, Cambridge, MA, 1998. MIT Press.

[42] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. *Computer Communications Review*, 29:251–262, 1999.

[43] S.H. Yook, Z.N. Oltvai, and A.-L. Barabàsi. Functional and topological characterization of protein interaction networks. *Proteomics*, 4(4):928–42, 2004.

[44] M.E.J. Newman. Who is the best connected scientist? A study of scientific coauthorship networks. *Physical Review E*, 64:016131, 2001. cond-mat/0011144.