RT-NL METHOD SUPPLEMENTAL MATERIAL*

JAMES J. DICARLO AND JOHN H.R. MAUNSELL (2004)

*** THIS DOCUMENT OUTLINES WORK DONE BY DICARLO AND MAUNSELL RELATING TO THEIR RT-NL METHOD PUBLISHED IN THE JOURNAL OF NEUROPHYSIOLOGY (2004). PLEASE SEE THAT MANUSCRIPT FOR OVERVIEW AND MAIN RESULTS. THIS WORK IS NOT PEER-REVIEWED OR OTHERWISE SANCTIONED BY THAT PUBLISHER.**

THEORETICAL OVERVIEW

An underlying assumption of the RT-NL method (DiCarlo and Maunsell 2004) is that, in RT tasks, neuronal activity generated by sensory transducers is transmitted through a potentially branching path of neuronal connections that ultimately leads to motor neurons whose activity produces a behavioral response to the stimulus. We refer to the sensory transducers, motor neurons and the neuronal elements that link their activity in a feed-forward way as the neuronal 'processing chain' that mediates the behavior. Note that we do not assume that all RT tasks are carried out by a fixed set of neuronal connections. Clearly, the remarkable flexibility of behavior suggests that neuronal processing is flexible in that different processing chains are invoked for different tasks and contexts. However, we do assume that, when the RT task and context are held nominally constant, the processing pathway underlying that particular RT task is also constant in that a particular set of brain structures and neurons within those structures is responsible for the sensory-motor transformation from stimulus to behavioral response. We refer to this set of neurons as being 'on' the processing chain for the task under study. Specifically, neurons whose activity is modulated after sensory stimulation and before the behavioral response, and which contribute, however indirectly, to the initial activation of the motor neurons underlying the behavioral response are considered 'on' the processing chain. Neurons that do not contribute to the initial activation of the motor neurons, whether or not their activity is modulated by the sensory stimulus, are considered 'off' the processing chain.

Mean neuronal latency and neuronal-behavioral covariance

In RT tasks, neurons located at anatomically early stages of processing (e.g. "sensory" neurons) respond with mean latencies that are short relative to behavioral responses, while neurons late in processing (e.g. "motor" neurons) become active only shortly before the RT. Hypothetical examples of the trial-by-trial relationship between NL and RT are shown in *Figure S1*. The raster plot in *Figure S1*A shows responses that might be recorded from a sensory neuron to repeated presentation of a given stimulus, while the plot in *Figure S1*B shows activity that might be recorded from a motor neuron.

Figure S1 about here

Each point in the scatter plot in *Figure S1*C plots the latency for the change in activity for one of these neurons in one trial (NL) against the RT on that trial. Lines show the best least-squares linear fit to each set of data. These lines show that these neurons differ not only in the mean time of their responses (reflected by the vertical offsets of the lines), but also in the trial-by-trial covariance of those response times with RT (reflected by the slopes of the lines). Thus, *Figure S1* suggests that neurons near the beginning and the end of neuronal processing chains can be distinguished based on either their mean NL or the trial-by-trial covariation of their NL and RT, and raises the possibility that these measures could be used to assign neurons to intermediate positions on processing chains. To fix ideas and to develop these concepts, we start by considering a highly simplified model of a neuronal processing chain (*Figure S2A*) and make simple assumptions about how activity propagates along that chain to produce RT variability. The structure of this model is similar to those used by others (e.g. Akamatsu et al. 1997).

Mean neuronal latency and neuron-behavioral covariance in networks

In the model in *Figure S2A*, each node (circle) represents a neuron or a group of neurons. The filled circles numbered 1 to n are part of a processing chain. A sensory stimulus directly drives neuronal elements in the first node of this chain (level 1) with neuronal response latency T_1. After a transit time, T_2, excitation is transmitted to the next node in the chain (level 2), which in turn drives the next node (level 3) after another transit time, T_3. Eventually the final level in the processing chain (level n) is excited, which causes the behavioral response with no further delay. The delay between the onset of the sensory stimulus and the activation of the final level is the reaction time, RT. For simplicity, we ignore delays associated with muscle activation and the detection of that activation. Additional neuronal elements (open nodes) are indirectly activated by the sensory stimulus, but do not contribute to the motor response (i.e. these neurons are 'off' the processing chain).

This simple model is not intended to explain how the brain works--it is simply a minimal model for exploring some of the ways that timing can vary along neuronal processing pathways. It does not describe all the pathways that contribute to a behavioral response nor does it include the effects of feedback. It captures the mean and covariance of the latency of only the initial neuronal and behavioral responses, so slow pathways that influence responses only after they start will not appear on the processing chain. This makes the model immune to complexities related to parallel processing chains of different speeds, but it means that it does not provide a complete description of the pathways that contribute to the entire execution of a behavioral response. In sum, the model in *Figure S2* should simply be taken a starting point to fix ideas and to begin quantitative discussion about how neuronal latencies might propagate through a neuronal network in the context of a reaction time task.

Figure S2 about here

To introduce RT variability across trials, we take all the transit times (T_i) to be random variables (with a mean greater than 0), so that the time to propagate activity between each nodal connection in *Figure S2A* varies from trial to trial. For the moment, we assume these random variables are independent and have the same distribution, but we make no assumption about the form of this distribution (except that it has a well defined variance). The mean NL will increase monotonically at each successive level, and the mean RT will be the mean NL for the final level. The mean NL for each neuron in *Figure S1C* corresponds to the mean of the data points projected on the y-axis. We describe a neuron's normalized mean NL with a value termed λ, which is found by dividing its mean NL by the mean RT. In the model, λ progresses linearly from ~0 to 1, as shown in *Figure S2B*.

This progression from 0 to 1 suggests that normalized mean NL (λ) may be useful for determining a neuron's position along the processing chain. Indeed, measures of mean NL have been used in many studies in an effort to order neurons or brain areas along processing pathways (e.g. Bullier and Henry 1979; Gawne et al. 1996; Maunsell and Gibson 1992; Nowak and Bullier 1997; Robinson and Rugg 1988; Schroeder et al. 1998; Zeki 2001). However, mean NL is not a panacea for revealing processing chains. One fundamental problem is that mean NL cannot distinguish between neurons that are on the processing chain and those that respond to the stimulus but are not on the processing chain. This is illustrated by considering the activation of nodes that are off the processing chain (open nodes) in *Figure S2A*. As neuronal activity propagates along these off-chain pathways, the mean NL will increase just as it does along the on-chain pathway. In this simple model, values of λ are the same for all nodes at a particular level, whether or not they are on the processing chain, and the λ values of all nodes at a given level in *Figure S2A* superimpose in *Figure S2B*. Thus, while measures of mean NL (e.g. λ) can provide a constraint on the relative ordering of neuronal elements, such measures cannot distinguish neurons that are on or off the processing chain.

To overcome this weakness, we show here that a statistical measure of the trial-by-trial association of RT and NL provides the possibility of distinguishing between neurons that are on and off the processing chain. Intuitively, the aim is to measure the amount of association that a neuron's response time (NL) has with RT across trials-the more association, the more likely the neuron is involved in producing that RT (i.e. being on the processing chain). We measure the association of RT and NL as the covariance of those variables (i.e. $\mathrm{cov}(RT,NL)$). Because we are not interested in the absolute covariance per se, but the fraction of the RT variance that is associated with the NL, we define a normalized measure of association (β), the covariance of NL and RT divided by the RT variance ($\beta = \mathrm{cov}(RT,NL)/\sigma_{RT}^2$). Graphical intuition about this measure can be gained by realizing that this definition of normalized covariance (β) is also the definition of the slope of the best fitting line resulting from the linear regression of unbiased trial-by-trial estimates of NL on RT (i.e. the linear regression for hypothetical data plotted like those in *Figure S1*C). However, it is important to emphasize that our analyses do not estimate the normalized covariance (β) by performing this linear regression, nor do they rely on a linear relationship between RT and NL. Moreover, although regression is often

used to provide a prediction of one variable given the value of another variable, this makes little physical sense for *Figure S1C* because the value of RT on a particular trial cannot cause the value of an earlier NL (see Discussion). We simply use this definition of β because it captures the association between NL and RT, and thus can inform about neuronal processing chains (described next).

A mathematical derivation of β_i for the nodes in *Figure S2A* is provided in the Appendix, but its behavior is easily understood. The covariance between RT and NL for the first level is ~0 because little latency variance has been introduced at this level. Because RT is defined as the NL of the last level in the processing chain (level n), NL for the final level covaries exactly with RT across trials. The covariance of RT and NL for the final level therefore equals the variance of RT ($\mathrm{cov}(RT, NL_n) = \sigma^2_{RT}$), and the normalized covariance, β_n, is 1. The RT-NL covariance at intervening levels is the amount of the total RT variance that has been introduced up to that level, so that the normalized covariance (β) at each level increases linearly (in the case of the simple model) from ~0 to 1 along the processing chain (*Figure S2C*, solid line).

How can the normalized covariance of NL and RT (β) distinguish neurons that are on and off the processing chain? Unlike λ, β can do more than simply increase with increasing level for neurons that are on the processing chain. For neuronal pathways that diverge from the processing chain, β will no longer increase from level to level, but will remain at the value of the level where divergent chain separated from the processing chain (*Figure S2*C). This is because the NL of neuronal elements that have diverged from the processing chain will contain variability that accumulated before the point of divergence, which will covary with RT variability, but (under the current set of assumptions) variability that accumulates along the divergent chain will not covary with the RT. *Figure S2D* shows simulations of the relationships between trial-by-trial NL and RT that might be observed for nodes that are early (a), intermediate (b) and late (c) on the processing chain, and additional nodes (d, e and f) that are off the processing chain (see legend). In sum, *Figure S2* shows that, at least under these simple assumptions, a measure of the covariance of RT and NL (β) can provide insight into neuronal function not provided by measures of mean NL alone (e.g. λ).

To determine if the RT-NL covariance measure (β) has utility beyond the highly simplified set of assumptions portrayed in Figure S2, we considered the consequences of relaxing those assumptions. Specifically, we considered the effects of: 1) lack of independence in the NL variability within and across processing levels, 2) the effects of neuronal convergence, and 3) different amounts of NL variability occurring at different levels. The main effects are summarized in *Figure S3*, and further details are provided in the Appendix. *Figure S3*A plots λ and β for the network in *Figure 2*B, keeping the assumption that all sources of NL variability are independent. Whereas *Figure S2B* and *C* plotted these values against level within the network, here they are plotted against each other. Because λ rises linearly with level (see *Figure S2B* and Appendix), it can be used as an observable surrogate for processing level, and *Figure S3*A thus has the same appearance as *Figure S2C*.

*Figure S3*B shows the expected effect of including some global covariation of neuronal latency. All of the β values for neurons along paths off the processing chain are increased, bringing them closer to the β values for neurons on the processing chain. This occurs because the global covariation causes some of the variance introduced on the divergent pathways to covary with RT. In the limit, all of the NL and RT variability is due to a single, global source, and β cannot distinguish on and off chain neurons.

*Figure S3*C shows that the effect of neuronal convergence. If each node in the network consists of many neurons that project in a highly convergent way to neurons in the next node, neurons in the later level can "average out" the variability that exists across the many neurons in the earlier level (Marsalek et al. 1997). The effect of assuming such convergence is to compress all of the β values toward 0, especially for neurons at early processing levels (*Figure S3C*). In the limit, all of the β values stay near 0 along the neuronal chains (rising toward 1 only at the final level), and β cannot distinguish on and off chain neurons.

Figure S3D shows the expected effect of the introduction of a very large amount of NL variability at a single processing level (relative to all other levels). The effect is a larger increase in β at that processing level (relative to other levels), and a compression of the β values for all other levels toward either 0 or 1, again making it difficult to distinguish on and off chain neurons.

In sum, *Figure S2* shows that a measure of the covariance of RT and NL (e.g. β) can provide information about a neuron's role in a neuronal processing chain that is not provided be measures of mean NL (e.g. λ). However, *Figure S3* cautions that the specific value of β for neurons in even simple neuronal networks strongly depends on the way in which neuronal latency variance accumulates in the network. Nevertheless, *Figure S3* reveals two important results that generalize across a wide range assumptions: β increases monotonically along any processing chain, and, among neurons at a particular processing level (i.e. the same λ), neurons on the processing chain will have the largest β value. Thus, measurement of β values among many neurons at a particular processing level (i.e. a particular value of λ or a particular brain area) could reveal the neurons in that area that are most likely involved in the processing chain (i.e. those with the highest values of β). *Figure S3* also reveals that recordings from many individual neurons over the entire range of neuronal processing levels could ultimately reveal the entire processing chain (i.e. the set of neurons that lie along the upper edge of the measured distribution of λ and β values). Finally, *Figure S3* shows how this method may provide insight into the factors that dominate NL and RT variability (i.e. which of the panels in *Figure S3* most resembles the recorded distribution of λ and β values in the brain?).

Figure S3 about here

The analysis described above shows that RT-NL measures (λ and β) may help distinguish on and off chain neurons in the brain. However, even in the simple network considered so far, extreme values of any of the factors considered in *Figure S3* might severely limit the utility of these measures. Other factors that we have not yet considered, such as correlated firing between neurons within one node, feedback, and pathways that change on each behavioral trial might also limit their utility.

APPENDIX

We characterized neuronal responses using two parameters that reflect the mean latency of the neuronal response (NL) and the trial-by-trial covariance between the NL and behavioral response (RT). The normalized mean latency, λ, is the mean NL divided by the mean RT: $\lambda = \mu_{NL} / \mu_{RT}$. The normalized covariance, β, is the covariance of the neuronal and behavioral latencies divided by the variance of the RT: $\beta = Cov(NL, RT) / \sigma^2_{RT}$. To analyze how λ and β might vary in neuronal networks, we extend the model introduced in *Figure S1* to that shown in *Figure S4* to allow for sources of latency variability that may be common to more than one node and for reduction in variance that might occur from converging inputs.

Each column in *Figure S4* represents a level of processing, and the index i specifies the column in which a node lies. The top row is the processing chain that mediates the behavior. Nodes in lower rows respond to the stimulus, but do not contribute to the response. The index j specifies the column in the processing chain at which a node's branch diverged. For nodes on the processing chain, $j = i$.

Figure S4 about here

On each behavioral trial, the latency of the response of each node is the sum of the latency of the preceding node and a delay introduced between those two nodes. The delay introduced between each pair of connected nodes is the sum of three sources, each a random variable. One source of variability ($T_{i,j}$) is independent of all other sources. The number of these sources is therefore equal to the number of pairs of connected nodes. Such variability could be caused by random effects that are *local* to each receiving node, so we refer to each $T_{i,j}$ as local random delay. A second source of variability is assumed common to all node connections, and is thus referred to as a *global* random delay (G). There is only one source of global random delay (i.e. one random variable, G). On each behavioral trial, a single random amount of global delay is added to all the node connections. The third source of latency variability between connected nodes is meant to allow for the possibility of delay mechanisms that are common to all nodes within a level of processing, but independent at each processing level. The number of these sources is therefore the number of processing levels. Because one interpretation of this variability

is a source that is common within each brain *area* (i.e. processing level), but is independent of other areas, we refer to this as areal random delay (A_i).

The analysis that follows makes no assumptions about the distributions of the random variables $T_{i,j}$, A_i and G except that each has a well defined variance (which is true for neurophysiologically plausible distributions). Although a physical interpretation of these sources of delay requires that they always assume positive values, the following analysis applies whether or not that restriction is imposed. There could be sources of variability that are distributed in many ways beyond those considered here, but these three let us describe a range of possible effects from the combination of different sources. While this model can be used to explore some effects of convergence and divergence, it is not intended to capture the complexity of the interacting pathways that are likely to be engaged during the execution of complex behaviors. It does not address details of how neurons at a given level interact, or how feedback might affect processing. The structure of the processing chain is task dependent: different behaviors will depend on different sets of nodes.

Equations:

The latency of any node is a random variable ($N_{i,j}$) that is the sum of all random delays up to that node (see *Figure S4*):

$$(1) \qquad N_{i,j} = \sum_{k=1}^{j} T_{k,k} + \sum_{k=j+1}^{i} T_{k,j} + \sum_{k=1}^{i} A_k + (i \cdot G)$$

Note that i is the processing level and r is the processing level where one node directly produces the response (see *Figure S1* and *S4*), and thus $i \in [1,r]$. Also note that, by construction, $i \geq j$ for all nodes. When $i > j$, the node is off the processing chain, and the first term represents the accumulation of local random delays (T) along the processing chain until the branch point, and the second term represents the accumulation of local random delays after the branch point. The remaining two terms show that delays common to an entire processing level (A) and the global delay (G) accumulate in exactly the same way as the processing level (i) increases, regardless of whether or not the node is on the processing chain.

The mean latency of each node is the expected value of the latency of the node and is thus the sum of the mean latencies up to that node:

$$(2) \qquad \mu_{N\,i,j} = E\left[N_{i,j}\right]$$
$$= \sum_{k=1}^{j} \mu_{T_{k,k}} + \sum_{k=j+1}^{i} \mu_{T_{k,j}} + \sum_{k=1}^{i} \mu_{A_k} + (i \cdot \mu_G)$$
$$\textit{where} \quad \mu_X = E\left[X\right]$$

Because each of the random delays $T_{i,j}, A_i$, and G are independent, the variance of the latency at each node ($\sigma^2_{N_{i,j}}$) is simply the sum of the variances of the random delays up to that node:

$$(3) \qquad \sigma^2_{N_{i,j}} = \sum_{k=1}^{j} \sigma^2_{T_{k,k}} + \sum_{k=j+1}^{i} \sigma^2_{T_{k,j}} + \sum_{k=1}^{i} \sigma^2_{A_k} + (i^2 \cdot \sigma^2_G)$$
$$where\ \sigma^2_X = Var(X)$$

The covariance between the latency of a node ($N_{i,j}$) and RT, defined as the latency of the final node on the processing chain ($N_{r,r}$), is straightforward to compute because the defined random delays are independent and thus have zero covariance:

$$(4) \quad Cov(N_{i,j}, N_{r,r}) =$$
$$= Cov\left(\left[\sum_{k=1}^{j} T_{k,k} + \sum_{k=j+1}^{i} T_{k,j} + \sum_{k=1}^{i} A_k + (i \cdot G)\right], \left[\sum_{k=1}^{r} T_{k,k} + \sum_{k=1}^{r} A_k + (r \cdot G)\right]\right)$$
$$= Cov\left(\left[\sum_{k=1}^{j} T_{k,k}\right], \left[\sum_{k=1}^{r} T_{k,k}\right]\right) + Cov\left(\left[\sum_{k=1}^{i} A_k\right], \left[\sum_{k=1}^{r} A_k\right]\right) + Cov\left([i \cdot G], [r \cdot G]\right)$$
$$= \sum_{k=1}^{j} Cov(T_{k,k}, T_{k,k}) + \sum_{k=1}^{i} Cov(A_k, A_k) + Cov\left([i \cdot G], [r \cdot G]\right)$$
$$= \sum_{k=1}^{j} \sigma^2_{T_{k,k}} + \sum_{k=1}^{i} \sigma^2_{A_k} + (i \cdot r \cdot \sigma^2_G)$$

The values of λ and β for any node (i,j) are defined as:

$$(5) \quad \lambda_{i,j} \equiv \mu_{N_{i,j}} / \mu_{N_{r,r}}$$
$$(6) \quad \beta_{i,j} \equiv Cov\left(N_{i,j}, N_{r,r}\right) / \sigma^2_{N_{r,r}}$$

Equations (2-6) are the general cases for any combination of the three types of random delays shown in *Figure S4*. We now examine specific cases under additional assumptions.

If we assume that the local random delays (T) and the delays common to a processing level (A) have the same mean and variance ($\mu_{T_{i,j}} = \mu_T; \mu_{A_i} = \mu_A; \sigma^2_{T_{i,j}} = \sigma^2_T; \sigma^2_{A_i} = \sigma^2_A; \forall i,j$), then the values of λ and β for any node (i,j) are:

$$(7) \quad \lambda_{i,j} = i \cdot \left(\mu_T + \mu_A + \mu_G\right) / r \cdot \left(\mu_T + \mu_A + \mu_G\right)$$
$$= i / r$$
$$(8) \quad \beta_{i,j} = \left(j \cdot \sigma^2_T + i \cdot \sigma^2_A + i \cdot r \cdot \sigma^2_G\right) \Big/ \left(r \cdot \sigma^2_T + r \cdot \sigma^2_A + r^2 \cdot \sigma^2_G\right)$$
$$= (j/r) + \left((i-j)/r\right) \cdot \left[\left(\sigma^2_A + r \cdot \sigma^2_G\right) / \left(\sigma^2_T + \sigma^2_A + r \cdot \sigma^2_G\right)\right]$$

Equations (7) and (8) show that, for nodes on the processing chain ($i = j$), $\lambda = \beta = (i/r)$. That is, λ and β values on the processing chain increase linearly with processing level, up to a value of 1 at the final processing level.

For nodes off the processing chain, λ is the same as if the neuron were on the processing chain, but β generally differs. In particular, the first term of Equation (8) shows that β increases to the level at which the chain leading to the node branched (j). The second term of Equation (8) shows how β will increase for nodes after the branch point. For example, if only local sources of random delay are assumed ($\sigma_A^2 = \sigma_G^2 = 0$), then this term is 0 and β does not increase beyond the value achieved at the branch point ($\beta_{i,j} = (j/r)$). *Figure S2C, D* and *Figure S3A* shows these effects by plotting λ and β for a 10 level network in which $\sigma_A^2 = \sigma_G^2 = 0$. Although 10 levels were used for these plots, their form is unaffected by the number of levels.

Equation (8) also shows that if either areal variability (σ_A^2) or global variability (σ_G^2) grow large (relative to σ_T^2), then $\beta_{i,j} \to i/r$. That is, in the limit of very large global or areal variability (or both), nodes off the processing chain have the same β values as nodes on the processing chain at the same processing level (i.e. β loses its utility to distinguish among these). *Figure S3B* shows the effect of global delay ($\sigma_G^2 = 0.1 \cdot \sigma_T^2$; $\sigma_A^2 = 0$).

Inspection of Equation (4) is useful for intuition about the utility of β for distinguishing nodes that are on and off the processing chain. The reason that β has utility is that, as we travel down any chain, the covariance of NL with RT accumulates differently for nodes on the processing chain relative to those that have branched off the processing chain. The last line of Equation (4) shows that the only difference in this accumulation of covariance with RT is the number of local variances that are accumulated along the chain.

Convergence and Non-uniform Sources of Latency Variance:

The final factors we will consider are the effect of convergence and sources of latency variance that are not uniformly distributed over all processing levels. Inputs from many neurons typically converge as signals are relayed from one level to the next in the brain. By summing the inputs from different sources, neurons may reduce the variance of their responses. The model presented here does not have any explicit convergence, but each node can be thought of containing many neurons, each of which receives convergent input from many neurons in the preceding node.

The exact effect of convergence on latency variance depends on assumptions that determine how inputs sum and give rise to spikes. Because there is little to constrain such assumptions, we model convergence as simply as possible. The results of Marsálek et al. (Marsalek et al. 1997) show, across a wide range of variance and in models of

different forms and complexity, that convergence acts to reduce latency variance by a fixed proportion. We therefore represent convergence by reducing latency variance by a fixed proportion at each level.

Because variance that is common to all pre-synaptic elements cannot be reduced by convergence, we implement convergence by allowing it to reduce only the variance from $T_{i,j}$ by a fixed proportion at each node. This is equivalent to taking each node to be comprised of many neurons that each receives an independent source of latency variance $T_{i,j}$. Convergence does not reduce variance from A_i and G, which affect all neurons within a level. We incorporate the effects of convergence in Equations (3) and (4) by applying a factor, s, to $\sigma^2_{T_i}$, where $0 < s < 1$:

$$(9) \qquad \sigma^2_{N_{i,j}} = \sum_{k=1}^{j} \sigma^2_{T_{k,k}} \cdot s^{i-k} + \sum_{k=j+1}^{i} \sigma^2_{T_{k,j}} \cdot s^{i-k} + \sum_{k=1}^{i} \sigma^2_{A_k} + (i^2 \cdot \sigma^2_G)$$

$$(10) \quad Cov(N_{i,j}, N_{r,r}) = \sum_{k=1}^{j} \sigma^2_{T_{k,k}} \cdot s^{r-k} + \sum_{k=1}^{i} \sigma^2_{A_k} + (i \cdot r \cdot \sigma^2_G)$$

Intuitively, the contribution from a particular local source $T_{i,j}$ to the final latency variability at the output (RT) decreases exponentially at successive levels in the network. Together, Equations (9), (10), (5) and (6) can be used together to compute λ and β values for any node in the network when convergence occurs at every level and for any combination of mean and variance values of the random latency sources.

Figure S3C shows the effect of convergence on β. The value of s was set to 0.25, corresponding to 75% of the variance arising from all T being eliminated at each node. Thus, the contribution from a particular local source $T_{i,j}$ decreases exponentially at successive levels in the network. Global and areal variances were all zero ($\sigma^2_{A_i} = \sigma^2_G = 0$). Convergence slows the increase in β along the processing chain, leaving most of the accumulation of covariance with RT to the final few levels. This makes it more difficult to distinguish nodes on the processing chain from those that are not.

It is also useful to consider the effects of sources of latency variance that are not uniformly distributed over all processing levels. Although the equations above consider the sources of A_i to have the same variance, σ^2_A, this variance might differ between levels. *Figure S3D* shows that, in some such conditions, β can undergo a large, step-like increment at a particular processing level. A high convergence factor ($s = 0.75$) keeps β from rising before the step and, to a lesser extent, after the step. The global correlation, σ^2_G, and most values of $\sigma^2_{A_i}$ are set low to zero, also contributing to relatively little change in β from level to level. An exception is σ^2_A for the intermediate level, where the step occurs. For this level σ^2_A is high ($25 \cdot \sigma^2_T$). The large variance at this level is the source for most of the increase in β and most of the variance in RT.

REFERENCES

Akamatsu M, Terao A, Hasbroucq T, Mouret I, and Seal J. Why does the single neuron activity change from trial to trial during sensory-motor task? *Methods of Information in Medicine* 36: 322-325, 1997.

Bullier J and Henry GH. Neural path taken by afferent streams in striate cortex of the cat. *J Neurophysiol* 42: 1264-1270, 1979.

DiCarlo JJ and Maunsell JH. Using neuronal latency to determine sensory-motor processing pathways in reaction time tasks. *J Neurophysiol*, 2004.

Gawne TJ, Kjaer TW, Hertz JA, and Richmond BJ. Adjacent visual cortical complex cells share about 20% of their stimulus-related information. In: *Cereb Cortex*, 1996, p. 482-489.

Marsalek P, Koch C, and Maunsell J. On the relationship between synaptic input and spike output jitter in individual neurons. *Proc Natl Acad Sci U S A* 94: 735-740, 1997.

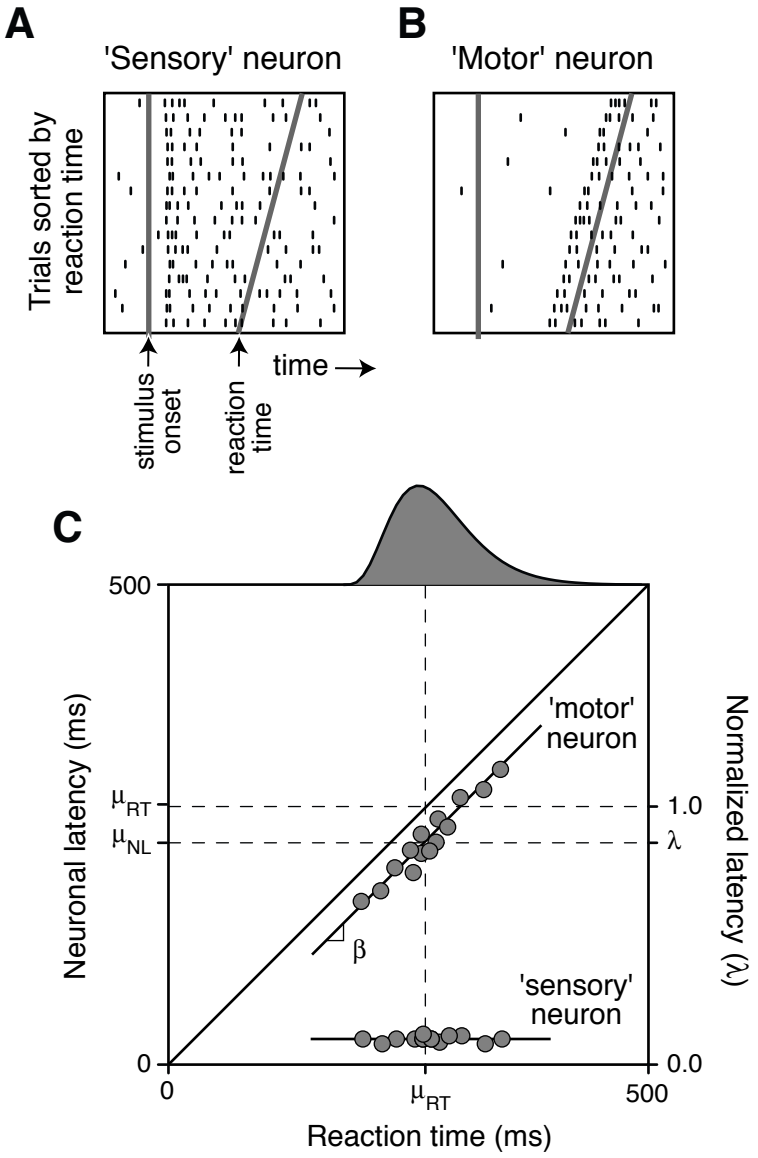
Maunsell JHR and Gibson JR. Visual response latencies in striate cortex of the macaque monkey. *J Neurophysiol* 68: 1332-1344, 1992.

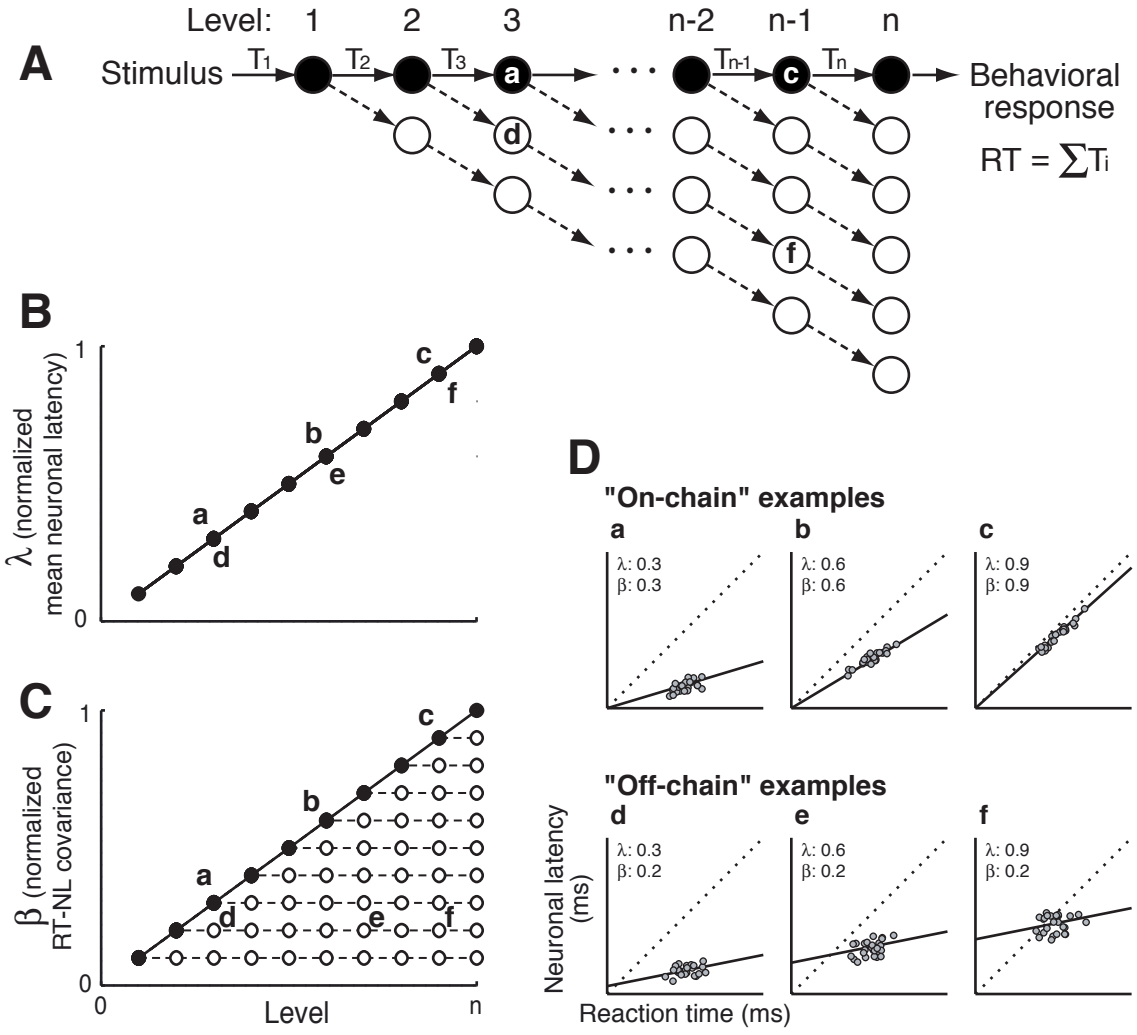
Nowak LG and Bullier J. The timing of information transfer in the visual system. In: *Cerebral Cortex: Extrastriate Cortex in Primate*, edited by Rockland K, Kaas J and Peters A: Plenum Publishing Corporation, 1997, p. 870.

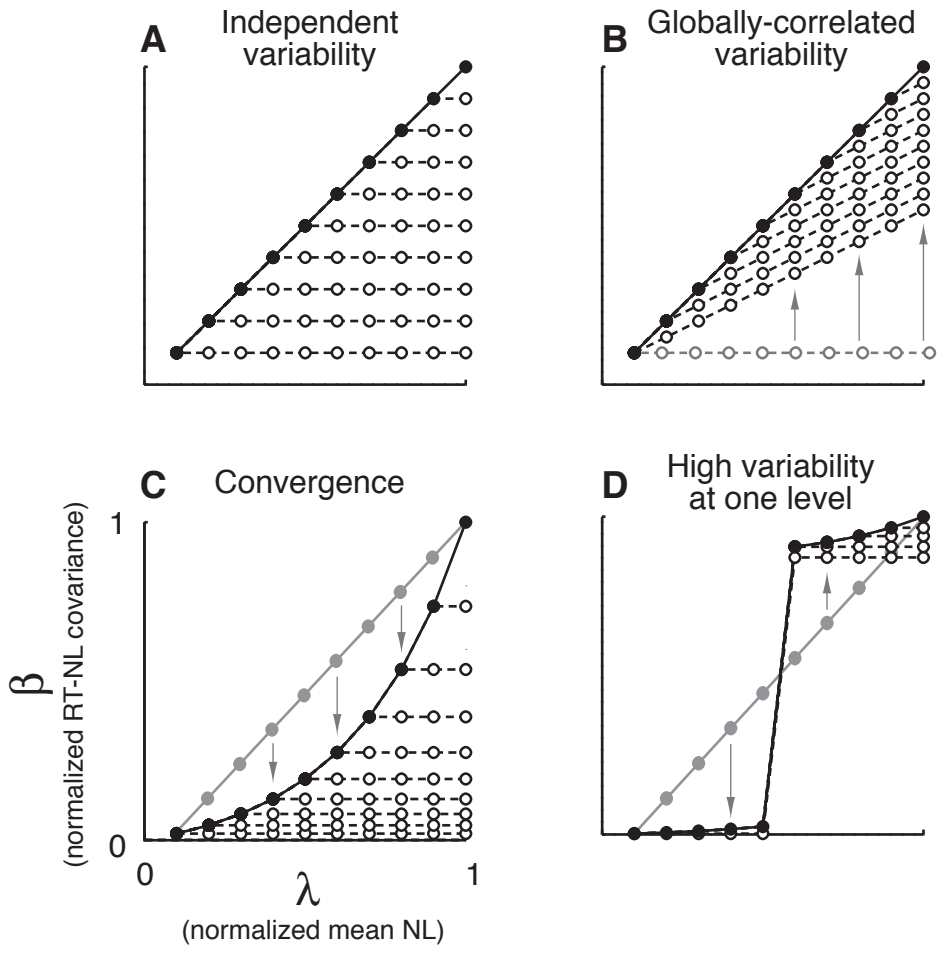
Robinson DL and Rugg MD. Latencies of visually responsive neurons in various regions of the rhesus monkey brain and their relation to human visual responses. *Biological Psychology* 26: 111-116, 1988.

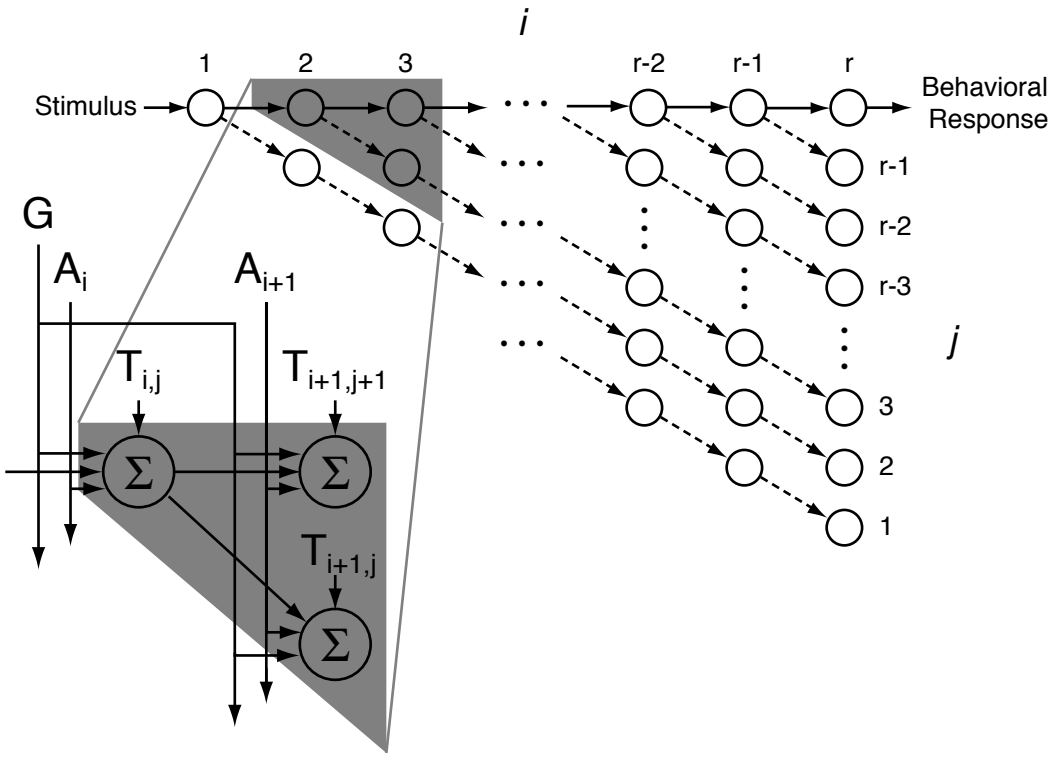
Schroeder CE, Mehta AD, and Givre SJ. A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. *Cerebral Cortex* 8: 575-592, 1998.

Zeki SM. Localization and globalization in conscious vision. *Ann Rev Neurosci* 24: 57-86, 2001.


A
'Sensory' neuron
B
'Motor' neuron
Trials sorted by reaction time
time
stimulus onset
reaction time
C
500
'motor' neuron
μ_{RT}
μ_{NL}
1.0
λ
β
'sensory' neuron
0
0.0
0
μ_{RT}
500
Neuronal latency (ms)
Normalized latency (λ)
Reaction time (ms)


A
Level: 1 2 3 n-2 n-1 n
Stimulus
Behavioral response
RT = ∑Ti
B
λ (normalized mean neuronal latency)
C
β (normalized RT-NL covariance)
Level
D
"On-chain" examples
"Off-chain" examples
λ: 0.3 β: 0.3
λ: 0.6 β: 0.6
λ: 0.9 β: 0.9
λ: 0.3 β: 0.2
λ: 0.6 β: 0.2
λ: 0.9 β: 0.2
Neuronal latency (ms)
Reaction time (ms)


A
Independent variability
B
Globally-correlated variability
C
Convergence
D
High variability at one level
β
(normalized RT-NL covariance)
1
0
0
1
λ
(normalized mean NL)


i
1
2
3
r-2
r-1
r
Stimulus
Behavioral
Response
r-1
r-2
r-3
j
3
2
1
G
A_i
A_i+1
T_i,j
T_i+1,j+1
T_i+1,j
Σ
Σ
Σ