

## Supplemental Data

### Probabilistic Population Codes

#### for Bayesian Decision Making

Jeffrey M. Beck, Wei Ji Ma, Roozbeh Kiani, Tim Hanks, Anne K. Churchland, Jamie Roitman, Michael N. Shadlen, Peter E. Latham, and Alexandre Pouget

### Supplemental Note: Probabilistic Population Codes for Bayesian Decision Making

#### Optimality of Global Inhibition

In the main text, we showed that a simple summation of MT spike counts leads to perfect information preservation in LIP. Thus, one way to preserve information would be to connect each MT neuron to a single LIP neuron which then perfectly integrates its inputs. One problem with this approach is that LIP neurons are likely to saturate very quickly to their maximum firing rate. This saturation can be prevented by using global recurrent inhibition, that is, by subtracting at each time step a term proportional to the total activity in the LIP layer.

Importantly, this type of inhibition does not result in any information loss. More specifically, inhibition modeled in this way maintains the Poisson-like statistics and leaves the associated posterior unaltered. To show this, suppose that at each time step the activity in LIP,  $\mathbf{r}$ , is Poisson-like with stimulus-dependent kernel  $\mathbf{h}(s)$ , and that inhibition acts by subtracting a vector  $\mathbf{n}$ , which is drawn from some conditional probability distribution  $p(\mathbf{n}|\mathbf{r})$ . In this case, the inhibited population pattern of activity is given by  $\mathbf{z}=\mathbf{r}-\mathbf{n}$  and we are interested in obtaining the conditions for which  $\mathbf{z}$  is Poisson-like with kernel  $\mathbf{h}(s)$  and contains the same amount of information as  $\mathbf{r}$ . As it turns out, one simple condition for which this is the case is when  $\mathbf{n}$  is in the null space of the derivative of  $\mathbf{h}(s)$  with respect to  $s$ , i.e., when  $\mathbf{h}'^T(s)\mathbf{n}=0$ . Specifically, we observe that,

$$\begin{aligned}
p(\mathbf{z}|s) &= \iint \delta(\mathbf{z}-\mathbf{r}+\mathbf{n}) p(\mathbf{n}|\mathbf{r}) p(\mathbf{r}|s) d\mathbf{r}d\mathbf{n} \\
&= \iint \delta(\mathbf{z}-\mathbf{r}+\mathbf{n}) p(\mathbf{n}|\mathbf{r}) \phi(\mathbf{r},c) \exp(\mathbf{h}^T(s)\mathbf{r}) d\mathbf{r}d\mathbf{n} \\
&= \int p(\mathbf{n}|\mathbf{z}+\mathbf{n}) \phi(\mathbf{z}+\mathbf{n},c) \exp(\mathbf{h}^T(s)\mathbf{z}+\mathbf{h}^T(s)\mathbf{n}) d\mathbf{n} \quad (\text{A1}) \\
&= \int p(\mathbf{n}|\mathbf{z}+\mathbf{n}) \phi(\mathbf{z}+\mathbf{n},c) f(\mathbf{n}) d\mathbf{n} \exp(\mathbf{h}^T(s)\mathbf{z}) \\
&= \phi^*(\mathbf{z},c) \exp(\mathbf{h}^T(s)\mathbf{z})
\end{aligned}$$

where we have used the fact that if  $\mathbf{n}$  is in the null space of  $\mathbf{h}'(s)$ , then  $f(\mathbf{n})=\exp(\mathbf{h}^T(s)\mathbf{n})$  is independent of  $s$ . Since  $p(\mathbf{r}|s)=\phi(\mathbf{r},c)\exp(\mathbf{h}^T(s)\mathbf{r})$ , it is clear from Eq. A1 that  $p(\mathbf{r}|s)$  and  $p(\mathbf{z}=\mathbf{r}-\mathbf{n}|s)$  (as a function of  $s$ ) only differ by a multiplicative constant. This in turn implies that the posterior distributions  $p(s|\mathbf{r})$  and  $p(s|\mathbf{z}=\mathbf{r}-\mathbf{n})$  are identical, and thus subtracting  $\mathbf{n}$  induced no information loss. Of course, this result is only useful if  $\mathbf{h}'(s)$  has a non-trivial null space. Fortunately, given the symmetry conditions which are enforced both on MT and the connectivity to LIP, it is easy to show that the vector comprised of only ones,  $\mathbf{1}=[1,1,\dots,1]^T$ , has the desired property. As a result, we can conclude that any global inhibition in LIP of the form,  $\mathbf{z}=\mathbf{r}-f(\mathbf{r})\mathbf{1}$ , when  $f(\mathbf{r})$  is a scalar (and possibly random) function of  $\mathbf{r}$ , will not affect either the information content or the Poisson-like statistics of LIP.

### Optimal Action Selection

As stated in the main text, a line attractor network in the superior colliculus (SC) layer of the network can recover the maximum likelihood estimate of saccade direction from LIP activity if the LIP spike counts follow ‘Poisson-like’ distributions (the LIP analog of Eq. 4 in the main text) and the dynamics in the SC layer is such that

$$\mathbf{v}^{\dagger\text{SC}}(s) \propto \mathbf{h}'^{\text{LIP}}(s), \quad (\text{A2})$$

where  $\mathbf{v}^{\dagger\text{SC}}(s)$  is the left null eigenvector of the Jacobian in the superior colliculus layer evaluated on the attractor and  $\mathbf{h}'^{\text{LIP}}(s)$  is the derivative of the kernel with respect to  $s$  in the LIP layer (Eq. 5 in the main text).

This result is a corollary of our previous work on optimal estimation with line attractor networks<sup>1,2</sup> and Bayesian inference with population codes<sup>3</sup>. In that work, we showed that the variance of an unbiased estimate obtained from a line attractor is minimized when the left null eigenvector of the attractor is given by

$$\mathbf{v}^{\dagger\text{SC}}(s) \propto \left(\boldsymbol{\Sigma}^{\text{LIP}}(s, c)\right)^{-1} \mathbf{f}'^{\text{LIP}}(s, c). \quad (\text{A3})$$

Here,  $\boldsymbol{\Sigma}^{\text{LIP}}(s)$  is the stimulus-dependent covariance matrix of the LIP activity and  $\mathbf{f}'^{\text{LIP}}(s)$  is the derivative with respect to the stimulus  $s$  of the stimulus-dependent mean, also called the tuning curve. Taking the derivative of the tuning curve associated with Poisson-like distribution yields,

$$\begin{aligned} \mathbf{f}'(s, c) &= \frac{d}{ds} \int \mathbf{r} \phi(\mathbf{r}, c) \exp(\mathbf{h}(s) \cdot \mathbf{r}) d\mathbf{r} \\ &= \int (\mathbf{r} \mathbf{r}^T \mathbf{h}'(s)) \phi(\mathbf{r}, c) \exp(\mathbf{h}(s) \cdot \mathbf{r}) d\mathbf{r} \\ &= \int (\mathbf{r} \mathbf{r}^T \mathbf{h}'(s) - \mathbf{r} \mathbf{f}^T(s, c) \mathbf{h}'(s)) \phi(\mathbf{r}, c) \exp(\mathbf{h}(s) \cdot \mathbf{r}) d\mathbf{r} \\ &= \left\langle \mathbf{r} (\mathbf{r} - \mathbf{f}(s, c))^T \mathbf{h}'(s) \right\rangle_s \\ &= \boldsymbol{\Sigma}(s, c) \mathbf{h}'(s) \end{aligned} \quad (\text{A4})$$

where we have used the fact that  $\mathbf{f}^T(s, c) \mathbf{h}'(s) = 0$ , which results from the fact that the probability distribution integrates to 1. Eq. A4 implies that

$$\mathbf{h}'^{\text{LIP}}(s) = \left(\boldsymbol{\Sigma}^{\text{LIP}}(s, c)\right)^{-1} \mathbf{f}'^{\text{LIP}}(s, c) \quad (\text{A5})$$

and comparison with Eq. (A3) yields the relationship described by Eq. (A2) and Eq. (8) in the main text.

We have shown in general that the variance of the estimate obtained from a line attractor network is minimized when Eq. (A3) is satisfied<sup>1,2</sup>. This would seem to imply that the SC layer provides the best possible unbiased estimator, in the sense that variance is minimal. However, all we have shown is that is that the SC layer provides the best

estimator *that can be obtained by a line attractor*. This last qualifier is crucial, since response distributions exists for which the line attractor network is suboptimal<sup>1,2</sup>. Fortunately, the Poisson-like family is not one of these. This can be seen by comparing the variance of the line attractor network estimate with the Cramer-Rao bound associated with the Poisson-like family of distributions.

As is easy to show based on previous work<sup>2</sup>, the variance of the unbiased network estimate in the SC layer, denoted  $\sigma_{\text{SC}}^2$ , is given by

$$\frac{1}{\sigma_{\text{SC}}^2} = \mathbf{f}'^{\text{LIP}}(s, c) \cdot (\boldsymbol{\Sigma}^{\text{LIP}}(s, c))^{-1} \mathbf{f}'^{\text{LIP}}(s, c). \quad (\text{A6})$$

To compute the Cramer-Rao bound in LIP, denoted  $\sigma_{\text{CR}}^2$ , we use the standard relation<sup>4</sup>

$$\frac{1}{\sigma_{\text{CR}}^2} = \left\langle \left( \frac{\partial}{\partial s} \log p(\mathbf{r}^{\text{LIP}} | s, c) \right)^2 \right\rangle \quad (\text{A7})$$

where the angle brackets represent an average with respect to  $p(\mathbf{r}^{\text{LIP}} | s)$ . If LIP neurons fire with Poisson-like statistics, then it is easy to show that

$$\frac{\partial}{\partial s} \log p(\mathbf{r}^{\text{LIP}} | s, c) = \mathbf{h}'(s) \cdot (\mathbf{r} - \mathbf{f}^{\text{LIP}}(s, c)) \quad (\text{A8})$$

where we have once again used the fact that  $\mathbf{h}'(s) \cdot \mathbf{f}^{\text{LIP}}(s, c) = 0$  for Poisson-like distributions. Using Eq. (A8),  $\sigma_{\text{CR}}^2$  becomes

$$\frac{1}{\sigma_{\text{CR}}^2} = \mathbf{h}'(s) \cdot \boldsymbol{\Sigma}^{\text{LIP}}(s, c) \mathbf{h}'(s) \quad (\text{A9})$$

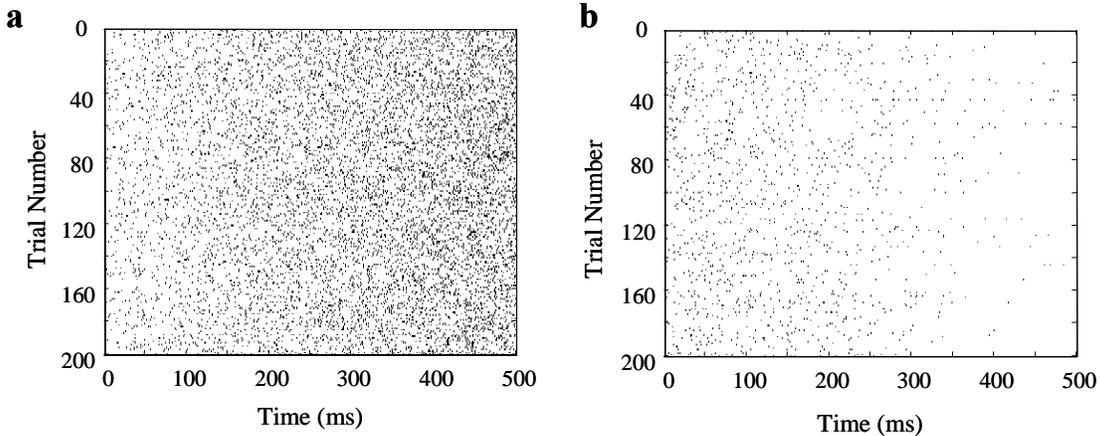
Recalling that  $\mathbf{h}'(s) = (\boldsymbol{\Sigma}^{\text{LIP}}(s, c))^{-1} \mathbf{f}'^{\text{LIP}}(s, c)$ , we conclude that

$$\frac{1}{\sigma_{\text{CR}}^2} = \mathbf{f}'^{\text{LIP}}(s, c) \cdot (\boldsymbol{\Sigma}^{\text{LIP}}(s, c))^{-1} \mathbf{f}'^{\text{LIP}}(s, c). \quad (\text{A10})$$

Comparing Eqs. (A6) and (A10), we see that they are identical. Thus, the line attractor network achieves the Cramer-Rao bound when the variability in LIP is Poisson-like. Note that if the response variability in LIP is not Poisson-like, there is no guarantee that the line attractor network can reach the Cramer-Rao bound, which means there is no guarantee that there is a set of parameters for which the line attractor network is optimal.

### Network Simulations with LNP Neurons

Fig. S1 shows two examples of raster plots for 200 trials with the coherence set to 51.2%. The top neuron has its receptive field in the direction of the motion in the display, while the bottom neuron has a receptive field 180° away.



**Figure S1:** Raster plots for two neurons in the LIP layer of the model. The direction of motion is the same on all trials with the coherence set to 51.2%. The stopping bound was deactivated to allow the neurons to integrate for the full duration of the trial (500ms). **a.** Neurons with a receptive field corresponding to the direction of motion in the display. **b.** Same as in **a** but for a neuron with a receptive field 180° away from the direction in the display.

### Analysis of LIP Data

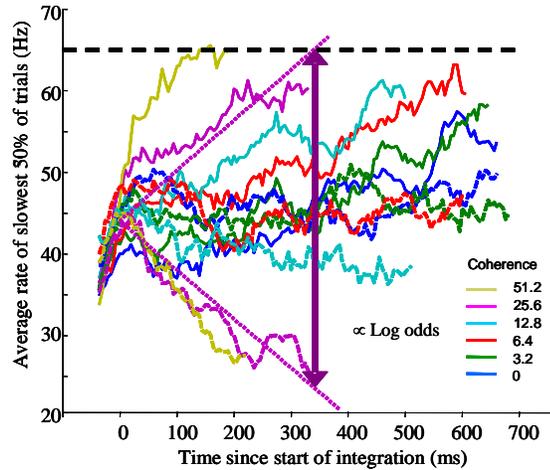
In the main text, we show that the LIP layer of our model encodes a probability distribution which reflects both the reliability of the sensory evidence (Fig. 3a-b) and the performance of the animal (Fig. 5a). To test quantitatively whether these predictions hold

in vivo, we would need to estimate both  $p(s|\mathbf{r}^{\text{LIP}})$  and the posterior used by the animal to make a decision. This would require multi-unit recordings in LIP, preferably in a continuous decision making task. Such data are not currently available. Instead, we use a data set from Roitman and Shadlen<sup>5</sup>, in which the monkey had to decide between two directions, 180° apart, and one from Churchland et al.<sup>6</sup>, in which the monkey had to decide between four directions, 90° apart. Only two types of neurons were recorded, those with preferred saccade directions corresponding to one of the two choice targets. These neurons were recorded on different trials.

Given these limitations, we cannot test our predictions quantitatively, but we can test them qualitatively. With only two targets and two types of neurons, we can estimate the log odds, averaged over many trials of the same coherence, from the average activity of the two groups of neurons. Using Eq. (6) in the main text (applied to LIP activity), and assuming a left-right symmetry, it can be shown that:

$$\left\langle \log \left( \frac{p(s = 0^\circ | r_0^{\text{LIP}}, r_{180}^{\text{LIP}})}{p(s = 180^\circ | r_0^{\text{LIP}}, r_{180}^{\text{LIP}})} \right) \right\rangle \propto \langle r_0^{\text{LIP}} \rangle - \langle r_{180}^{\text{LIP}} \rangle \quad (\text{A11})$$

where  $r_0^{\text{LIP}}$  and  $r_{180}^{\text{LIP}}$  are the activities of the LIP neurons tuned to 0° and 180°, respectively. Therefore, if we can measure a quantity proportional to  $\langle \Delta r \rangle = \langle r_0^{\text{LIP}} \rangle - \langle r_{180}^{\text{LIP}} \rangle$  in LIP, we can test our predictions qualitatively. Specifically, we can ask whether the average log odds in LIP grows with time and whether it grows faster for higher coherence (as in the case of the model, see Fig. 3a). In addition, we can ask whether the average log odds grows with coherence (as shown for the model in



**Figure S2:** Log odds in LIP. For each coherence, the average firing rate of the slowest 50% of trials at that coherence is plotted, both for neurons which contain the chosen target (solid lines) and neurons which contain the other target (dashed lines). Each average is computed only up to the termination time of the fastest trial in the average, minus 60 ms to account for the saccade. This is meant to ensure that neither saccadic bursts nor terminated trials are included in the average. Straight lines starting from a common point are fitted through both the upward and downward moving traces, as illustrated by the dotted line for 25.6% coherence. When the upper trace hits 65 Hz, the difference between the upper and the lower line is computed. This difference is approximately proportional to the average log odds of choice.

Fig. 5a). Since coherence is monotonically related to performance (through the psychometric curve), the finding that the average log odds grows with coherence would also show that the average log odds at decision time reflects the animal's performance).

We first describe how we estimated  $\langle \Delta r \rangle$  at decision time. Estimating the firing rates is in principle straightforward, but in practice complicated by the fact that different trials terminate at different times and that LIP neurons exhibit a saccade-related burst at the end of each trial. This means we cannot perform a straightforward average of all trials. What we do instead is consider, at each motion strength, the 50% of the trials that take the longest to terminate, and note that our results are robust to the fraction of trials included. To eliminate the saccadic burst, which is not part of the integration, we remove the last 60 ms of all trials. We then average separately the firing rates for  $T_{\text{in}}$  (the neuron whose response field corresponds to the chosen target) and for  $T_{\text{out}}$  (the neuron whose response field corresponds to a saccade  $180^\circ$  away from the chosen target) – in other words, we are conditioning neural responses on the direction of the saccade (note that for this particular analysis, we ignore  $T_{90}$  in the four-choice data). Both averages were computed only for the duration of the shortest trial in the average. We then perform a linear regression, assuming that the integration starts at 190 ms and 44 spikes/s (for the Churchland data, 155 ms and 27 spikes/s). We extrapolate the linear fits to 65 spikes/s (see Fig. S2) (for the Churchland et al data, 73 spikes/s). At the time the upper trace hits 65 spikes/s (73 spikes/s, respectively), we measure the difference between the fits to both traces. Fig. 5c in the main text shows that the difference increases with coherence. Because the fits are linear, the value of the bound is irrelevant to this result – a different bound would only give an overall scale factor.

Given the various approximations we have made (e.g. linear fits, a percentage of the trials excluded, no correlations), the difference in firing rates obtained with this method,  $\langle \Delta r \rangle$ , is not the true difference in firing rates at decision time, but a number monotonically related to this variable. Ultimately, this is not a problem since we do not need to know the absolute value of the log odds; only their relative values matter for our prediction. Therefore, the scaling between  $\langle \Delta r \rangle$  and log odds can be set arbitrarily. In

Figs. 5c and d, when we plot the log odds, we chose a scaling (and a shift) to match Figs. 5a and b.

To obtain the error bars in Fig. 5, we assumed that the rates increase linearly with time and are perturbed by noise,

$$r^{\text{LIP}}(t_i) = at_i + b + \eta_i,$$

where  $t_i = 190 + i \cdot 60$  are the times in ms at which the rates are computed ( $i = 0, 1, 2, \dots$ ). Integration is taken to start at 190 ms after stimulus onset. Rates are computed by dividing the average number of spikes across all trials in an interval of length 60 ms centered at  $t_i$  by the length of that interval.  $b$  is the rate at the start of integration, which is taken to be fixed at  $44 \text{ s}^{-1}$ . The noise  $\eta_i$  is assumed to follow a normal distribution of unknown variance. Then this is a linear regression without a constant term. The best

estimate of the slope is  $\hat{a} = \frac{\sum_{i=1}^n (r_i^{\text{LIP}} - b)t_i}{\sum_{i=1}^n t_i^2}$  and the error in this estimate is given by

$$\sigma_a = \sqrt{\frac{\sum_{i=1}^n (r_i^{\text{LIP}} - \hat{a}t_i - b)^2}{(n-1)\sum_{i=1}^n t_i^2}}.$$

The decision time (w.r.t. the start of integration) is  $t_{\text{decision}} = \frac{r_{\text{bound}} - b}{a_0}$ , where  $a_0$  is the integration slope for the  $T_{\text{in}}$  neuron. The log odds at decision time can now be rewritten as

$$\text{LO} = \langle r_0 \rangle_{t_{\text{decision}}} - \langle r_{180} \rangle_{t_{\text{decision}}} = (r_{\text{bound}} - b) \left( 1 - \frac{a_{180}}{a_0} \right).$$

It is estimated by substituting the estimates of the slopes,  $a_0$  and  $a_{180}$ . Since we know the errors in both slopes, the error in the log odds at decision time can be approximated as

$$\sigma_{\text{LO}} = (r_{\text{bound}} - b) \sqrt{\frac{\sigma_{a_{180}}^2}{\hat{a}_0^2} + \frac{\hat{a}_{180}^2}{\hat{a}_0^4} \sigma_{a_0}^2}$$

To test the other prediction (average log odds as a function of time and coherence, Fig. 3a), we performed the same analysis with two modifications. Instead, of conditioning the activity of the neurons on direction of saccade, we conditioned them on direction of motion. Also, when plotting log odds in Fig. 3d, we use the following scaling:  $\log \text{odds} = 0.6 \times \langle \Delta r \rangle$ .

### Optimality in the Presence of the Bound

It is important to note that the presence of a bound does not affect the optimality of the action selection mechanism described in the main text. This may seem counterintuitive as the exact mechanism which implements the bound has a strong effect on the shape of the noise distribution  $p(\mathbf{r}^{\text{LIP}}(t)|s)$ . However, it is possible to show that any stopping criterion which depends only upon the pattern of activity  $\mathbf{r}^{\text{LIP}}(t)$ , does not affect the information content of LIP at stopping time. This can be demonstrated by simply showing that the posterior distribution associated with the complete history of activity in MT can be reconstructed from the activity in LIP. However, it is perhaps more informative to demonstrate this fact by showing the similarity in the actual computations of the noise distributions for LIP in the presence and absence of a bound. Therefore, consider the time history of activity from MT,  $\mathbf{r}^{\text{MT}}(t_1:t_N)$ , which is supposed to be Poisson-like:

$$p(\mathbf{r}^{\text{MT}}(t_1:t_N)|s) = \phi(\mathbf{r}^{\text{MT}}(t_1:t_N)) \exp\left(\mathbf{h}(s) \cdot \sum_{n=1}^N \mathbf{r}^{\text{MT}}(t_n)\right). \quad (\text{A12})$$

When LIP is just the temporal sum of MT activity, we can conclude LIP activity at time  $t_N$  is distributed according to

$$\begin{aligned}
p(\mathbf{r}^{\text{LIP}}(t_N) | s) &= \int \dots \int \delta\left(\mathbf{r}^{\text{LIP}}(t_N) - \sum_{n=1}^N \mathbf{r}^{\text{MT}}(t_n)\right) p(\mathbf{r}^{\text{MT}}(t_1:t_N) | s) d\mathbf{r}^{\text{MT}}(t_1) \dots d\mathbf{r}^{\text{MT}}(t_N) \\
&= \int \dots \int \delta\left(\mathbf{r}^{\text{LIP}}(t_N) - \sum_{n=1}^N \mathbf{r}^{\text{MT}}(t_n)\right) \prod_{n=1}^N p(\mathbf{r}^{\text{MT}}(t_n) | s) d\mathbf{r}^{\text{MT}}(t_n) \\
&= \int \dots \int \delta\left(\mathbf{r}^{\text{LIP}}(t_N) - \sum_{n=1}^N \mathbf{r}^{\text{MT}}(t_n)\right) \prod_{n=1}^N \phi(\mathbf{r}^{\text{MT}}(t_n)) \exp(\mathbf{h}(s) \cdot \mathbf{r}^{\text{MT}}(t_n)) d\mathbf{r}^{\text{MT}}(t_n) \\
&= \int \dots \int \delta\left(\mathbf{r}^{\text{LIP}}(t_N) - \sum_{n=1}^N \mathbf{r}^{\text{MT}}(t_n)\right) \exp\left(\mathbf{h}(s) \cdot \sum_{n=1}^N \mathbf{r}^{\text{MT}}(t_n)\right) \prod_{n=1}^N \phi(\mathbf{r}^{\text{MT}}(t_n)) d\mathbf{r}^{\text{MT}}(t_n) \\
&= \int \dots \int \delta\left(\mathbf{r}^{\text{LIP}}(t_N) - \sum_{n=1}^N \mathbf{r}^{\text{MT}}(t_n)\right) \exp(\mathbf{h}(s) \cdot \mathbf{r}^{\text{LIP}}(t_N)) \prod_{n=1}^N \phi(\mathbf{r}^{\text{MT}}(t_n)) d\mathbf{r}^{\text{MT}}(t_n) \\
&= \left( \int \dots \int \delta\left(\mathbf{r}^{\text{LIP}}(t_N) - \sum_{n=1}^N \mathbf{r}^{\text{MT}}(t_n)\right) \prod_{n=1}^N \phi(\mathbf{r}^{\text{MT}}(t_n)) d\mathbf{r}^{\text{MT}}(t_n) \right) \exp(\mathbf{h}(s) \cdot \mathbf{r}^{\text{LIP}}(t_N)) \\
&= \left( \int \dots \int \prod_{\substack{\sum_{n=1}^N \mathbf{r}^{\text{MT}}(t_n) = \mathbf{r}^{\text{LIP}}(t_N) \\ n=1}}^N \phi(\mathbf{r}^{\text{MT}}(t_n)) d\mathbf{r}^{\text{MT}}(t_n) \right) \exp(\mathbf{h}(s) \cdot (\mathbf{r}^{\text{LIP}}(t_N))) \tag{A13} \\
&= \phi^{\text{LIP}}(\mathbf{r}^{\text{LIP}}(t_N)) \exp(\mathbf{h}(s) \cdot \mathbf{r}^{\text{LIP}}(t_N))
\end{aligned}$$

On the second-to-last line, we just rewrote the first term as an integral over all paths of  $\mathbf{r}^{\text{MT}}(t_1:t_N)$  that have the property that  $\sum_{n=1}^N \mathbf{r}^{\text{MT}}(t_1:t_N) = \mathbf{r}^{\text{LIP}}(t_N)$ . Thus, in the absence of the bound,  $\mathbf{r}^{\text{LIP}}(t_N)$  is also Poisson-like. This is because the sum over all possible paths of  $\mathbf{r}^{\text{MT}}(t_1:t_N)$  which lead to a particular  $\mathbf{r}^{\text{LIP}}(t_N)$  only affects the measure function  $\phi^{\text{LIP}}$ , which is independent of  $s$ . When a decision boundary is crossed for the first time at time  $t_N$  we no longer sum over all paths of  $\mathbf{r}^{\text{MT}}(t_1:t_N)$  which lead to a particular  $\mathbf{r}^{\text{LIP}}(t_N)$ , but rather only over those paths which did not cross the decision boundary prior to time  $t_N$ . When the boundary function is only a function of  $\mathbf{r}^{\text{LIP}}(t_n)$  and thus only a function of  $\mathbf{r}^{\text{MT}}(t_1:t_N)$  the only thing that changes in Eq. (A13) is the number of terms in the sum on the second to last line, i.e. we now only integrate over the paths that led to  $\mathbf{r}^{\text{LIP}}(t_N)$  but did not cross the bound. Therefore, when the decision bound is implemented by any function which is only a function of LIP or MT activity, i.e. independent of  $s$ , we can conclude that the

presence of the bound only affects the measure function, and is therefore Poisson-like with

$$p(\mathbf{r}^{\text{LIP}}(t_N) | s) = \phi^{\text{LIP+BOUND}}(\mathbf{r}^{\text{LIP}}(t_N)) \exp(\mathbf{h}(s) \cdot \mathbf{r}^{\text{LIP}}(t_N)) \quad (\text{A14})$$

where

$$\phi^{\text{LIP+BOUND}}(\mathbf{r}^{\text{LIP}}(t_N)) = \int \dots \int \prod_{n=1}^N \phi(\mathbf{r}^{\text{MT}}(t_n)) d\mathbf{r}^{\text{MT}}(t_n)$$

$\sum_{n=1}^N \mathbf{r}^{\text{MT}}(t_n) = \mathbf{r}^{\text{LIP}}(t_N)$   
 and bound not crossed

Finally, comparison of the associated MT and LIP posteriors, (Eqs. (2) and (6) in the main text) demonstrates that no information was lost in the MT-to-LIP transformation.

## Supplemental References

1. Deneve, S., Latham, P. & Pouget, A. Efficient computation and cue integration with noisy population codes. *Nature Neuroscience* **4**, 826-831 (2001).
2. Latham, P.E., Deneve, S. & Pouget, A. Optimal computation with attractor networks. *Journal of Physiology (Paris)* **97**, 683-694 (2003).
3. Ma, W.J., Beck, J.M., Latham, P.E. & Pouget, A. Bayesian inference with probabilistic population codes. *Nat Neurosci* **9**, 1432-1438 (2006).
4. Papoulis, A. *Probability, random variables, and stochastic process* (McGraw-Hill, inc., New York, 1991).
5. Roitman, J.D. & Shadlen, M.N. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J Neurosci* **22**, 9475-9489 (2002 <http://www.shadlen.org/sl/matlab/data/jamie/jamie.htm>).
6. Churchland, A.K., Kiani, R. & Shadlen, M.N. Decision-making with multiple alternatives. *Nat Neurosci* **11**, 693-702 (2008).