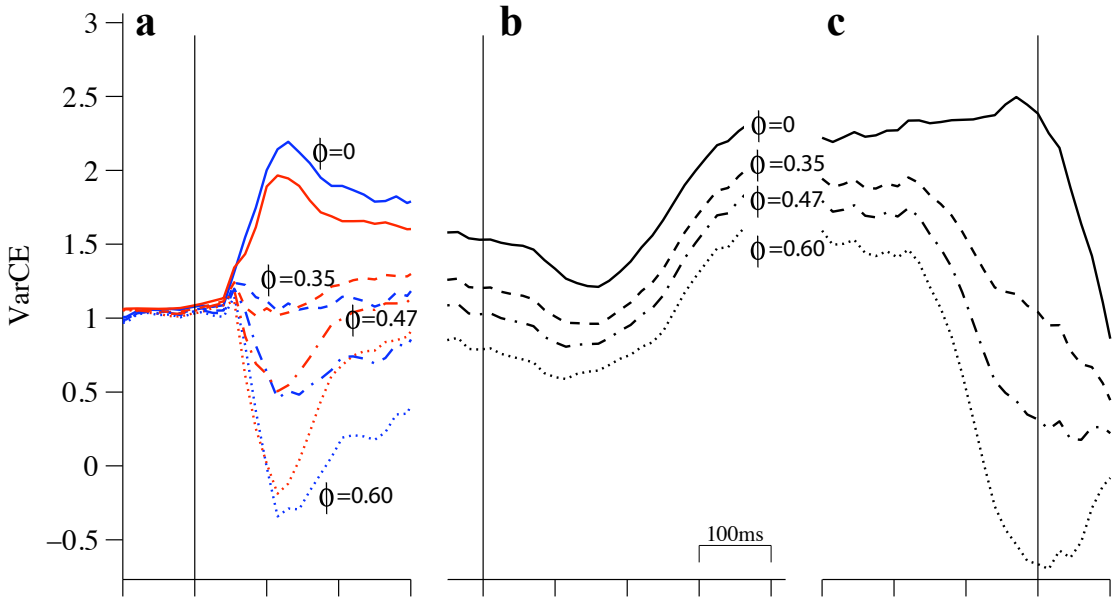
Inventory of supplementary material:

This supplementary material contains:

- Supplemental data: 5 figures and their accompanying legends. The figures in the main text to which they relate are listed in each title.
- Supplemental references

Supplemental data

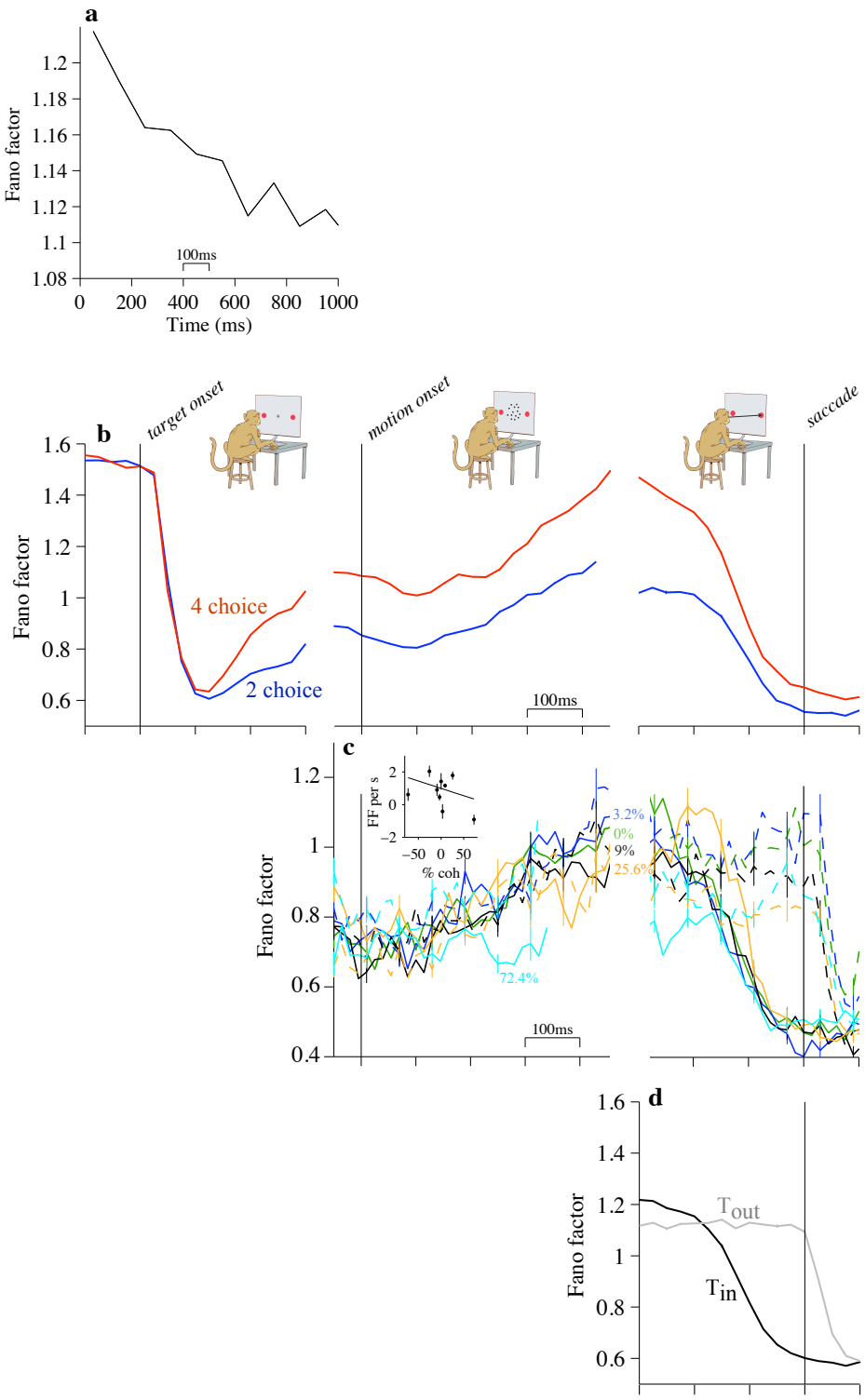
Churchland et al. Supp. Fig. 1



Supplementary figure 1 (Related to Methods and Figures 3,4, and 5). Robustness of main findings against variation in the value of ϕ. The values illustrated for these calculations span a credible range. $\phi = 0.6$ represents a rough upper bound, because larger values would imply a negative VarCE at some time points, which is impossible. $\phi = 0.35$ represents a rough lower biological bound based on current injection experiments performed in cortical slice (Nawrot et al., 2008). Note that when $\phi = 0$ (solid traces), there is no PPV, and the quantity depicted is the sample variance of the spike count. Except for the change in ϕ, calculations are identical to the corresponding figures in the main text. **a**. Effect of number of choices on VarCE in the pre-

decision epoch (same data as in Fig. 3). Number of choices is indicated by color. The critical finding that VarCE is larger for 4-choice trials is supported for a large range of ϕ. The relationship reverses when $\phi < 0.3$, which is not surprising because the higher firing rates on 2-choice trials are associated with larger total variance. **b**. The rise in VarCE during early motion viewing is present for all values of ϕ. **c**. The decline in VarCE preceding T_{in} choices is apparent for a wide range of ϕ. It is statistically reliable for $\phi \geq 0.15$. The decline is not apparent when $\phi=0$, which is unsurprising: the high presaccadic firing rates for trials ending in T_{in} choices ensure a large amount of total variance. In the main text, we point out that the late decline in VarCE is less apparent for T_{out} choices (Fig. 5a, *bottom,* gray vs. black traces). This difference is statistically reliable for $\phi \geq 0.30$ (data not shown).

Churchland et al. Supp. Fig. 2

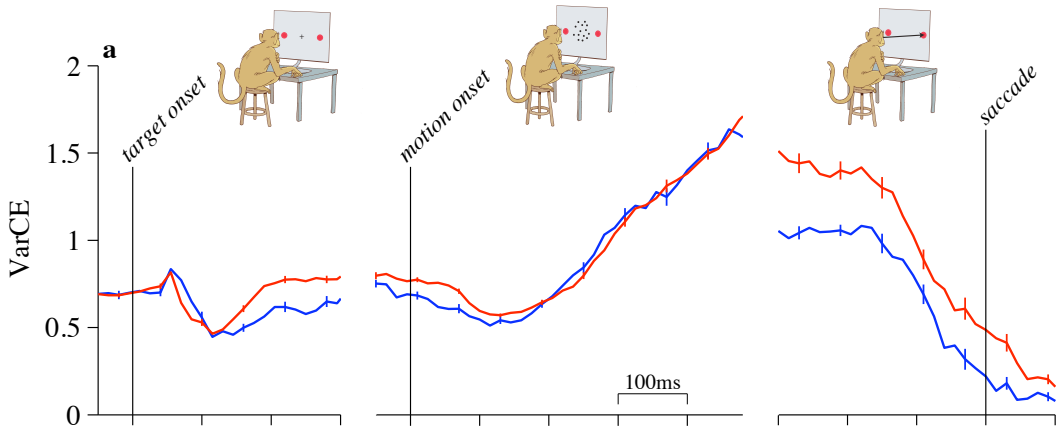


Supplementary figure 2 (Related to Figures 3,4,5) Additional remarks on the Fano factor. The Fano factor is the ratio of variance to mean spike count over repetitions of trials in a fixed counting window. It is constant for a rate-modulated (or non-homogeneous) renewal process

(Cox and Isham, 1980; Nawrot et al., 2008). However, a time dependent change in the Fano factor does not imply that a process is not a rate-modulated renewal (cf., Teich et al., 1997). It could instead reflect a time dependent change in VarCE. Thus, in many instances, variation in the Fano factor parallels the changes in VarCE. **a**. Example of a dissociation between the Fano factor and VarCE. The process depicted in Figure 2c of the main text shows a linear rise in firing rate with noise perturbation instantiated as a random constant offset. This random component is captured by a contant VarCE (black horizontal line in Fig. 2f). The Fano factor for the same example decreases as a function of time. It gives the misleading impression that the stochastic point process becomes less variable as a function of time, whereas the simulation is in fact a nonstationary Poisson point process with a different offset to the rate on each repetition. The example underscores a potential advantage of the VarCE over the Fano factor when investigating the variance associated with firing rates that change over time. **b-d.** Fano factor computed from data. Columns indicate trial epoch (icons). *Left column:* Responses are aligned to the onset of the choice targets. All trials are included (as in Fig. 3). *Middle column*: Responses are aligned to motion onset. *Right column*: Responses are aligned to the time of the saccade. Only correct choices are included (as in Fig. 5). **b**. Effect of number of choices. *Left panel*: In the pre-decision epoch, the Fano factor follows the same pattern as the VarCE (same data as Fig. 3). *Middle panel*: In the early motion-viewing epoch, the time dependent increase in Fano factor mimics the VarCE, but unlike the VarCE (Fig. 4c, *bottom*), the Fano factor is larger for the 4-choice data (same data as in Fig. 4c). *Right panel*: At the end of the decision, the Fano factor undergoes a precipitous decline preceding T_{in} choices for both 2- and 4-choice conditions (same data as Fig. 5c). **c**. Fano factor for different motion strengths and directions. Motion strength is indicated by color (labels); motion direction is indicated by line style (solid lines, T_{in} motion; dashed lines, T_{out} motion). Data from 2- and 4-choice conditions are combined (using residuals; see Methods). Mean firing rates for these conditions are displayed in Supp. Fig. 4 (*top*). *Left*: responses aligned to motion onset. Like the calculation of VarCE in this epoch, trials leading to all possible choices comprise the mixture of trials from which the variance (and mean) are calculated. As in the main text, data within 100ms of saccade initiation are not included in the calculations. 3000-9000 trials contribute to each curve. *Inset* shows rate of rise of the Fano factor from 190-290 ms after motion onset as a function of signed motion strength, where sign indicates direction. The rate of rise of the Fano factor did not depend on motion strength and direction ($p=0.35$, weighted regression). *Right*: responses aligned to saccade initiation. Like the calculation of VarCE in this epoch, only trials that have identical motion strength and lead to the same choice (e.g., T_{in})

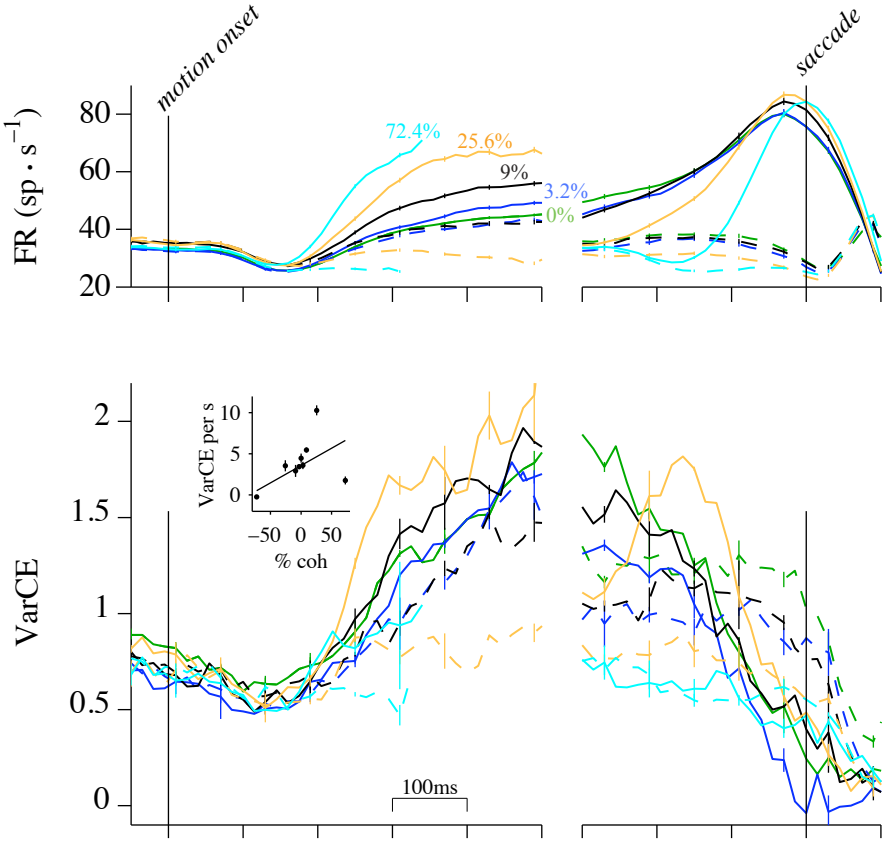
constitute the mixture of trials from which the variance (and mean) are calculated. Color indicates motion strength; line style indicates the saccadic choice. Error trials for nonzero motion strengths are excluded from this analysis. Trials with 0% motion strength are grouped by choice. 2000-3000 trials contribute to each curve. **d.** Fano factor for responses aligned to saccade initiation. The early decline for T_{in} choices and later decline for T_{out} choices is similar to the time course of VarCE in this epoch (same trials as in Fig. 5a).

Churchland et al. Supp. Fig. 3



Supplementary figure 3 (Related to Figures 3,4,5) Effect of number of choices on VarCE. This is simply a reorganization of figure panels from the main text to facilitate direct comparison. In the main text, we focused our analyses to expose novel features of neural computation in three epochs: pre-decision, early decision formation and decision termination. These analyses naturally enforce different trial groupings, as explained in the main text. Some readers may find it helpful to compare analyses displayed on a single graph. The left, middle and right panels are from Figures 3, 4c and 5c, respectively.

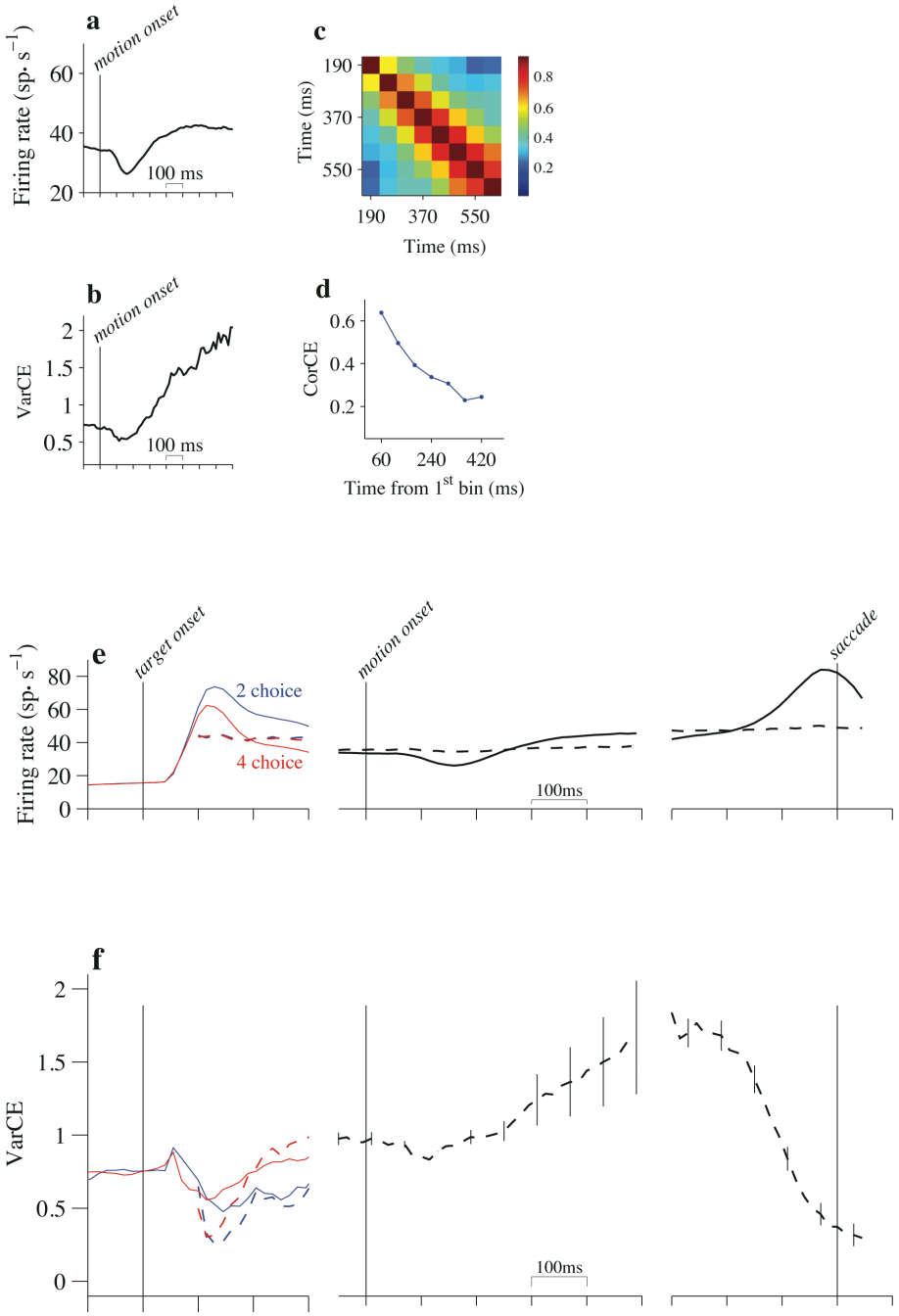
Churchland et al. Supp. Fig. 4



Supplementary Figure 4 (Related to Figures 4,5). Effect of motion strength and direction on mean firing rate (*top*) and VarCE (*bottom*) during motion viewing and around the time of the saccade. Here we furnish a more complete set of conditions to complement the examples shown in Figures 4a and 5a,b of the main text. Motion strength is indicated by color (labels); motion direction is indicated by line style (solid lines: T_{in} motion; dashed lines: T_{out} motion). Data from 2- and 4-choice conditions are combined (using residuals; see Methods). *Left*: VarCE for data aligned to motion onset, excluding epochs from trials that fall within 100 ms of saccade initiation. Trials leading to all possible choices constitute the mixtures from which the variance and mean are calculated. 3000-9000 trials contribute to each curve. *Inset* shows rate of rise of the VarCE from 190-290 ms after motion onset as a function of signed motion strength, where sign indicates direction. The rate of rise of the VarCE depends weakly on motion strength and direction (slope=0.042±0.017 units of VarCE per s per %coh, $p<0.03$, weighted regression). *Right*: VarCE for data aligned to saccade initiation. Only trials leading to the same choice (e.g.,

T_{in}) constitute the mixture of trials from which the variance is calculated. Error trials for nonzero motion strengths are excluded from this analysis. 2000-3000 trials contribute to each curve.

Churchland et al. Supp. Fig. 5



Supplementary figure 5 (Related to Figures 3,4,5). The time-dependent changes in VarCE are not explained by changes in mean firing rate or the value of ϕ. We applied two strategies to address this issue. Both effectively suppress changes in mean firing rate as a function of time. For the first strategy, we chose a subset of conditions in which the mean firing rate is constant as

a function of time. We analyzed responses during early decision formation on both 2- and 4-choice trials when motion was greater than 3.2% coherence and in the T_{out} direction. Analyses based on this strategy are in panels a-d.

For the second strategy, we applied a "mean matching" procedure to subsample the spike counts in each time window in a manner that ensures no change in the mean firing rate as a function of time (Churchland et al., 2010). To understand the technique, consider the 0% coherence condition during the early motion-viewing epoch (middle panel). For each of the 8 successive *time-bins*, we established a histogram of the 70 mean firing rates (one per neuron), using 10 *rate-bins*. For the i^{th} time-bin, this provides 10 histogram counts (i.e., number of observations): $k_{i1}, k_{i2}, ..., k_{i10}$, where $\sum_j k_{ij} = 70$ samples. For each of these rate-bins, we then identify the smallest count among the 8 time-bins: $K_j = \min\left\{k_{1j}, k_{2j}, ..., k_{8j}\right\}$. The 10 values, $K_1, K_2, ..., K_{10}$, constitute the "largest common distribution" (LCD) for which we are assured that at least K_j neurons contribute the same mean rate in each of the 8 time windows. These are not necessarily the same neurons, of course. We then constitute a population of responses in each time bin by sampling from a random subset of the K_j neurons that contribute to the j^{th} rate bin. The random sampling is necessary because $K_j \le k_{ij}$. Importantly, a sampled neuron contributes all of its trials (for that time bin) to the calculations of mean and variance that follow. This procedure is guaranteed to produce approximately the same mean rate in each time window ("approximate" is warranted because the rate-bins have finite width). Because the mean counts are not time dependent, the VarCE equals the sample variance minus a constant, regardless of the value of ϕ. For the analysis depicted in panels e and f, we combined data from neurons and stimulus conditions. Note that the "mean matching" procedure does not permit calculation of the CorCE matrix because each neuron contributes to only a subset of the time bins.

a-d, First strategy: selection by motion strength and direction.

a. Mean firing rate for all neurons during early decision formation, using only the set of trials with motion opposite to T_{in} (see above). Responses are aligned to the onset of random dot motion. After an initial dip, firing rates change by only ~8 sp/s from 190 ms after motion onset to the end of the analysis epoch. **b.** VarCE increases over the same time period. **c.** CorCE matrix. The pattern is similar to that observed in the analysis using all trials (Fig. 4d). **d.** CorCE between spike counts at the beginning of decision formation (160-220 ms after motion onset) and each subsequent time bin. This is the top row of the CorCE matrix in c. Note similarity to Fig. 4e.

e-f, Second strategy: mean matching.

e. Mean firing rates for responses aligned to different epochs in the trials indicated by the black vertical lines. Note that matched mean is the same throughout the duration of each epoch (dashed traces). Analyses were performed separately in each epoch to facilitate a comparison of effects that take place within an epoch. Solid curves show the raw mean firing rates, for the same conditions, before application of the mean-matching procedure. For responses aligned to target onset (left), we generated a distribution of spike counts that was common to both 2- and 4-choice trials to ensure that any differences in the VarCE did not arise from differences in the mean firing rates associated with those two conditions. The number of targets (for this epoch only) is indicated by color (labels). **f**. VarCE. The same trends are present in both the raw VarCE (bottom panels in Figs, 3, 4 and 5) and mean-matched VarCE. Error bars indicate standard error of the mean across conditions. This provides reassurance that changes in VarCE are not due to changes in the PPV, because the PPV, which depends on the mean, is the same at every time point in the epoch for the mean-matched data.

Supplemental References

Churchland, M.M., Yu, B.M., Cunningham, J.P., Sugrue, L.P., Cohen, M.R., Corrado, G.S., Newsome, W.T., Clark, A.M., Hosseini, P., Scott, B.B.*, et al.* (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nature neuroscience *13*, 369-378.

Cox, D.R., and Isham, V. (1980). Point processes (London ; New York: Chapman and Hall).

Nawrot, M.P., Boucsein, C., Rodriguez Molina, V., Riehle, A., Aertsen, A., and Rotter, S. (2008). Measurement of variability dynamics in cortical spike trains. Journal of neuroscience methods *169*, 374-390.

Teich, M.C., Heneghan, C., Lowen, S.B., Ozaki, T., and Kaplan, E. (1997). Fractal character of the neural spike train in the visual system of the cat. J Opt Soc Am A Opt Image Sci Vis *14*, 529-546.