Supplementary Information

Our physiological conclusions are robust against variations in the data analysis. There were two key criteria in our analysis. (1) Contour criterion: For a given RF heat map we measured, we set a percentage of the peak response to mark a contour around the peak for calculating the RF center of mass. The contour criterion was set to 85% in the main text. (2) Completeness criterion: We required that the measured RF heat map included a minimum percentage of the contour defined by the contour criterion. This completeness criterion was set to 80% in the main text. We did extensive additional data analysis to demonstrate that our physiological conclusions are robust against variations in these criteria. We focused on Fig. 4 of the main text as it contained the main results on the RF shift directions in the delay and perisaccadic periods for both LIP and FEF. In Figs. S1 and S2, we kept the contour criterion at 85%, but set the completeness criterion to 90% and 70%, respectively (instead of 80% in Fig. 4). In Figs. S3 to S5, we changed the contour criterion to 75%, a value used by Zirnsak et al.'s [15], and set the completeness criterion to 90%, 80%, and 70%, respectively. These figures all show results similar to those in Fig. 4 of the main text.

686

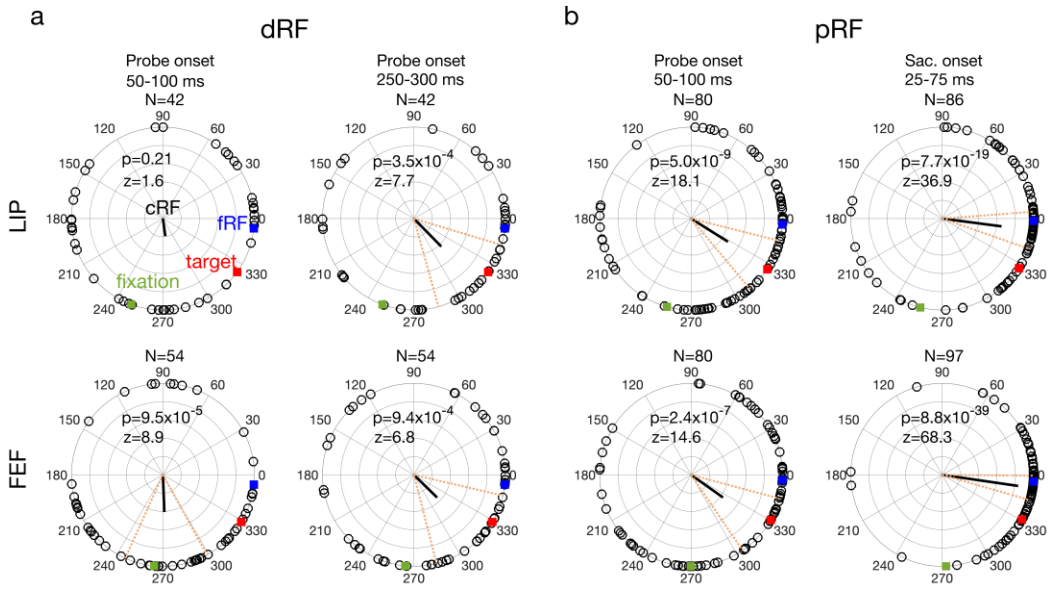

Fig. S1. The delay (a: dRF) and perrisaccadic (b: pRF) shift directions of all LIP (top row) and FEF (bottom row) cells from different time periods (columns). The contour criterion was 85% and the completeness criterion was 90%. The format was identical to that of Fig. 4 of the main text. The mean directions changed significantly across time in both LIP ($p = 2.5\times10^{-4}$, $F_{3,246}=6.6$) and FEF ($p = 1.9\times10^{-9}$, $F_{3,281}=15.7$), with Watson-Williams multi-sample test.

687

688

689

690

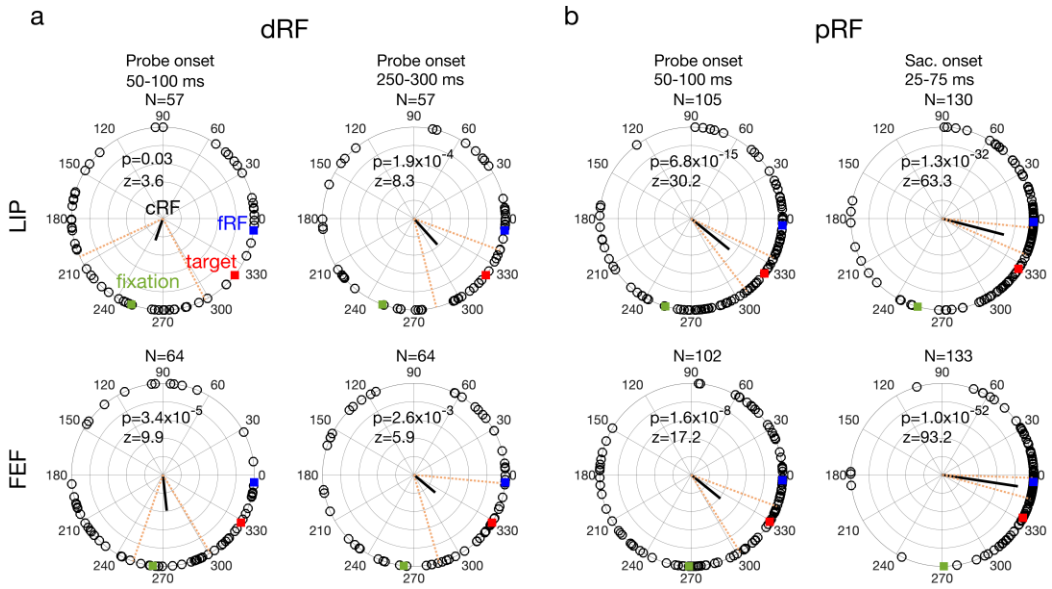

Fig. S2. The delay (a: dRF) and perrisaccadic (b: pRF) shift directions of all LIP (top row) and FEF (bottom row) cells from different time periods (columns). The contour criterion was 85% and the completeness criterion was 70%. The format was identical to that of Fig. 4 in the main text. The mean directions changed significantly across time in both LIP ($p = 3.6\times10^{-9}$, $F_{3,345}=14.9$) and FEF ($p = 1.7\times10^{-10}$, $F_{3,359}=17.3$), with Watson-Williams multi-sample test.

691

692

693

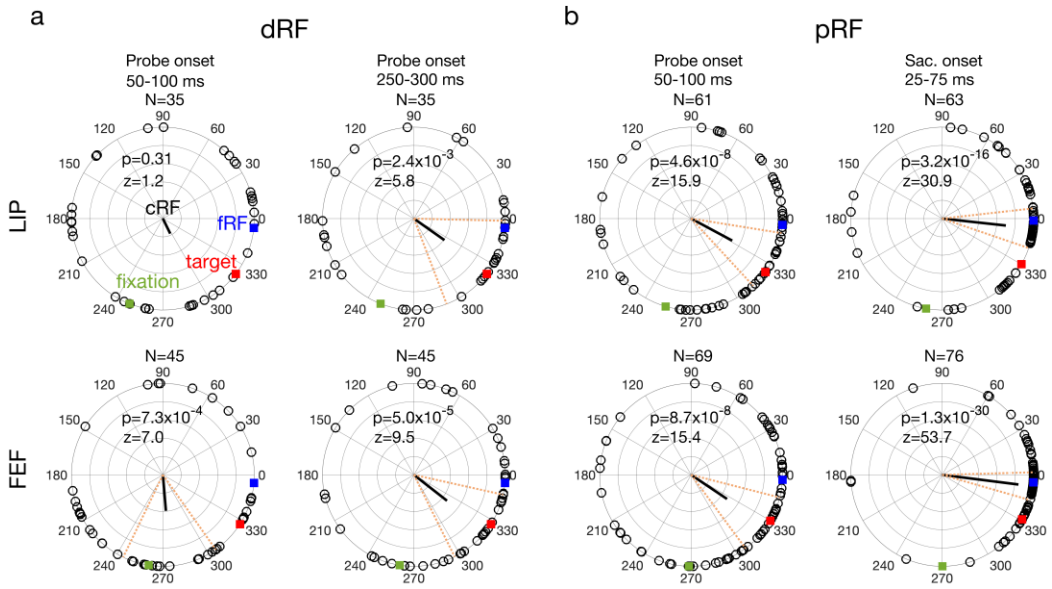

Fig. S3. The delay (a: dRF) and perrisaccadic (b: pRF) shift directions of all LIP (top row) and FEF (bottom row) cells from different time periods (columns). The contour criterion was 75% and the completeness criterion was 90%. The format was identical to that of Fig. 4 in the main text. The mean directions changed significantly across time in both LIP ($p = p = 0.021$, $F_{3,190}=3.3$) and FEF ($p = 1.4\times10^{-7}$, $F_{3,231}=12.5$), with Watson-Williams multi-sample test.

694

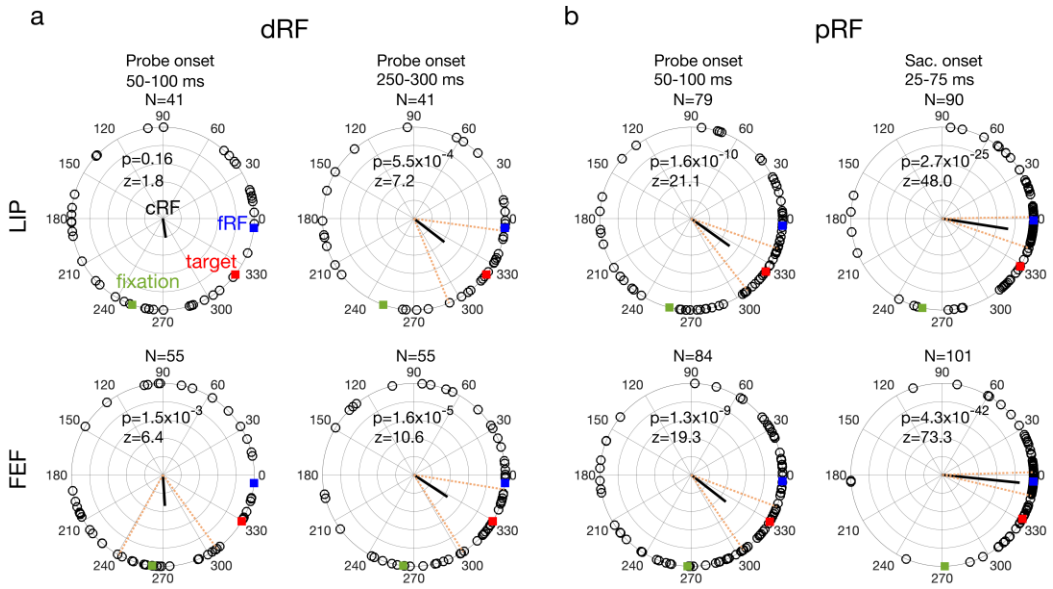

Fig. S4. The delay (a: dRF) and perrisaccadic (b: pRF) shift directions of all LIP (top row) and FEF (bottom row) cells from different time periods (columns). The contour criterion was 75% and the completeness criterion was 80%. The format was identical to that of Fig. 4 in the main text. The mean directions changed significantly across time in both LIP ($p = 1.9\times10^{-4}$, $F_{3,247}=6.8$) and FEF ($p = 2.5\times10^{-9}$, $F_{3,291}=15.4$), with Watson-Williams multi-sample test.

695

696

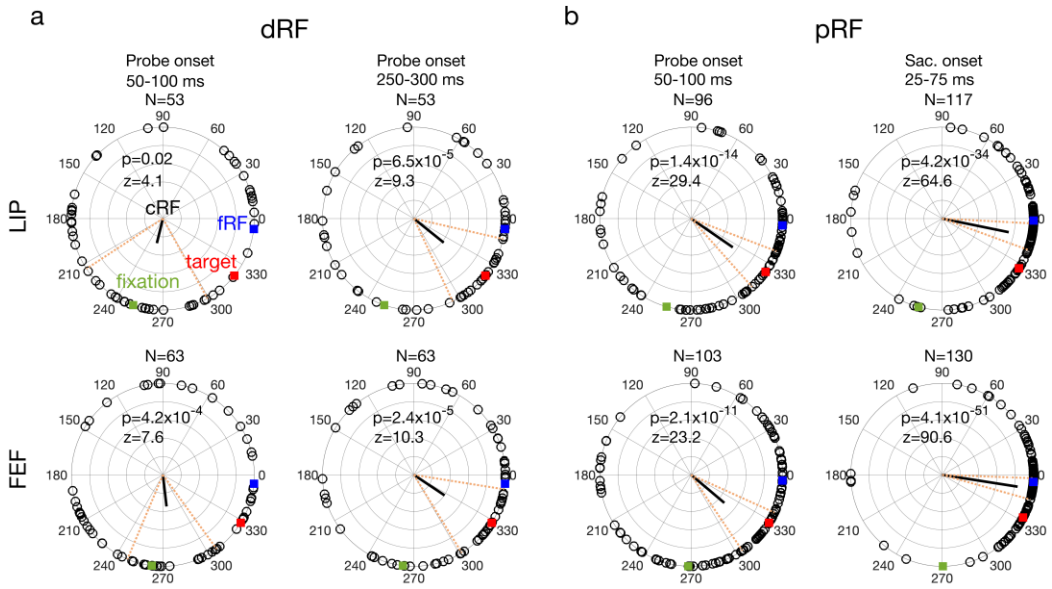

Fig. S5. The delay (a: dRF) and perrisaccadic (b: pRF) shift directions of all LIP (top row) and FEF (bottom row) cells from different time periods (columns). The contour criterion was 75% and the completeness criterion was 70%. The format was identical to that of Fig. 4 in the main text. The mean directions changed significantly across time in both LIP ($p = 6.7\times10^{-9}$, $F_{3,315}=14.5$) and FEF ($p = 7.0\times10^{-10}$, $F_{3,355}=16.2$), with Watson-Williams multi-sample test.

697

698 In the main text, we showed the distributions of the RF shift directions at four
699 time points (Fig. 4). For completeness, we show in Fig. S6 the distributions of the shift
700 vectors (both the directions and amplitudes) [21].

701

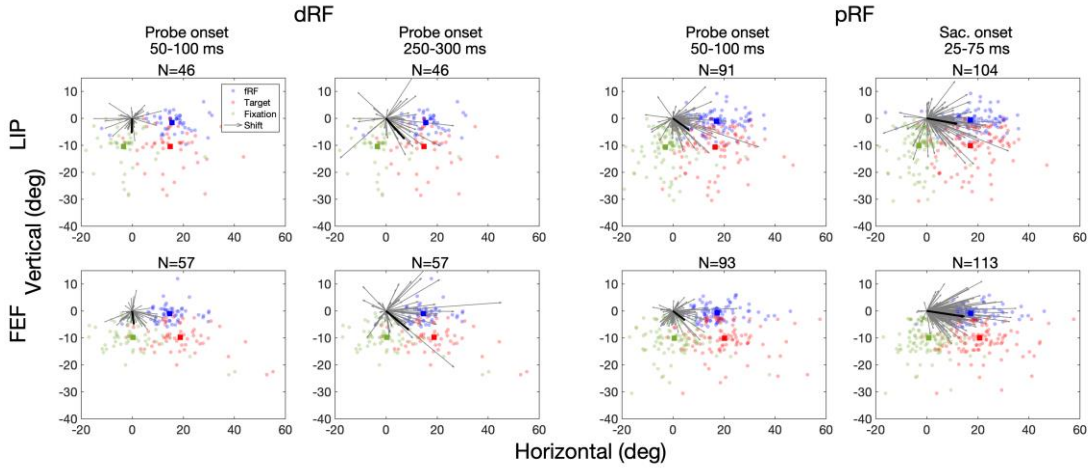

Fig. S6. The delay (a: dRF) and perrisaccadic (b: pRF) shift vectors of all LIP (top row) and FEF (bottom row) cells from different time periods (columns). This figure corresponds to Fig. 4 of the main text but shows both the shift direction and amplitude of each cell. In each panel, we align the cells' cRF centers at (0, 0) and saccade directions along positive horizontal. The cells' fRF centers, the targets, and the initial-fixation points are shown as blue, red, and green dots, respectively, and their mean positions as the blue, red, and green squares, respectively. Gray arrows indicate the cells' RF shift vectors and the black line is the vector determined by calculating the mean direction and mean amplitude of the individual vectors.

702

703

704

In Fig. 10 of the main text, we showed the automatic emergence of both the attention-modulated center/surround connections and the CD-gated directional connections in artificial neural networks trained to predictively update, across saccades, the representation of retinal locations of briefly flashed stimuli. We ran additional simulations to show that the same was true under many other conditions, with two examples in Figs. S7 and S8. In Fig. S7, we trained a neural network on both brief input stimuli and persistent input stimuli. In Fig. S8, we repeated the simulation in Fig. 10 of the main text but without the attentional modulation at the stimuli. In both cases, we found similar connectivity patterns to those in Fig. 10. It is not surprising that attention at the stimuli is not important for learning the connectivity patterns. To perform the task of updating the stimulus retinal positions, a network had to develop the center/surround connectivity to maintain the attractor activity pattern and the CD-gated directional connectivity to move the attractor pattern appropriately [30]. These requirements do not depend on attentional modulation. Once the connections are learned, attention can modulate the center/surround connectivity to enhance processing at the attended location and cause convergent RF shifts.

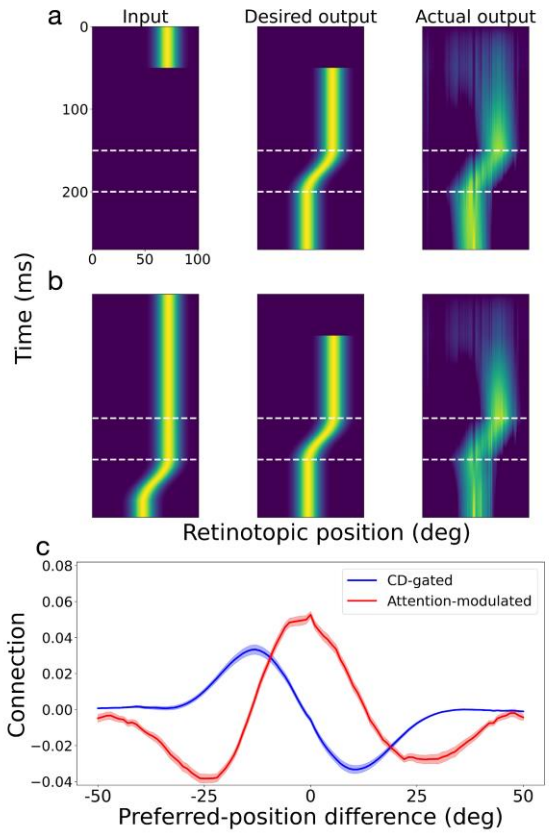

Fig. S7. Automatic generation of the required connectivity patterns in the circuit model by training neural networks to predictively update retinal positions of both brief (a) and persistent (b) input stimuli during saccades. The format of the figure was identical to that for Fig. 10 of the main text except that both an example of brief input (a) and an example of the persistent input (b) are shown.

731

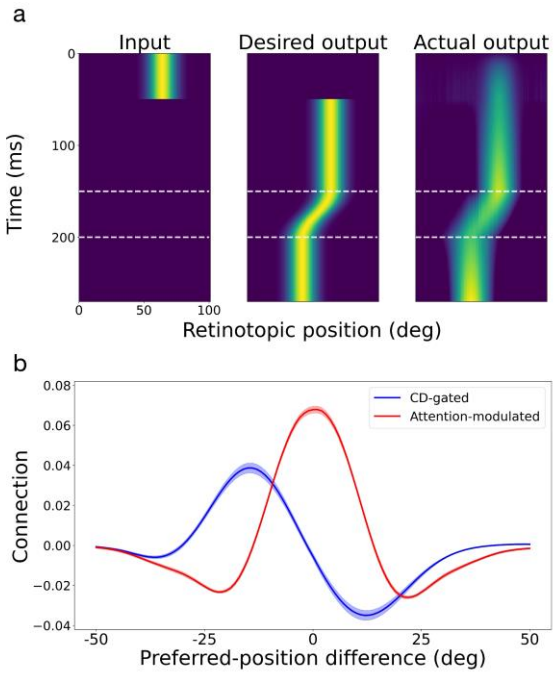

Fig. S8. Automatic generation of the required connectivity patterns in the circuit model by training neural networks to predictively update retinal positions of brief input stimuli during saccades without attentional modulation. We still labeled the symmetric connections (red) as attention-modulated for easy comparison with Figs. 10 and S7. The format of the figure was identical to that for Fig. 10 of the main text.