Supplementary Figure 1: A comparison of the adaptation kernels for the full population of cells, with those obtained by fitting the model to neurons adapted within 120 degrees of their preferred directions. Related to Figure 4.

Supp. Figure 1A-C shows the kernels plotted in Figure 4 of the main text. The kernels for PDS (Supp. Figure 1A) and CDS (Supp. Figure 1C) cells differ strongly. As shown in the main text and in Supplementary Figure 3, the disinhibition for stimuli similar to the adapter is critical for the loss of pattern selectivity in MT after adaptation. It is also necessary for explaining the maintained responsivity of preferred-adapted PDS cells to their preferred grating, and the attractive shifts in grating tuning for flank-adapted PDS cells.

The disinhibition was particularly prominent in PDS cells adapted within 120 degrees of their preferred directions (Supp. Figure 1D); cells adapted further from their preferred direction (not shown) showed little evidence of disinhibition. This is likely because the value of the kernels at the adapted direction is not well constrained for these null-adapted cells. For instance, a PDS cell preferring 180 deg motion will have essentially no response to stimuli moving in the adapted direction (0 degrees). The loss of pattern selectivity in such a neuron is driven by the loss of responsivity in V1 to stimulus directions offset from the adapter (i.e. the trough of the kernel in Supp. Figure 1A), and thus closer to the preferred direction of these cells. Thus, an adaptation kernel consisting only of a stimulus-specific loss of responsivity can explain the loss of pattern selectivity in null-adapted PDS cells. Critically, although the tuned disinhibition component of the kernel is not required for explaining effects in null-adapted PDS cells, its presence does not preclude a loss of pattern selectivity in these cells (as shown in Figure 6 of the main text, and in more detail in Supplementary Figure 3). This is important because it means that a single set of V1 adaptation effects can explain the loss of pattern selectivity for PDS neurons with a range of preferences relative to the adapter.

Supplementary Figure 2: Effects of adaptation on V1 population responses. Related to Figures 4 and 5.

To test whether the differences between kernels for CDS and PDS cells could be explained by input from V1 neurons with differing surround suppression, we measured how adaptation altered population responses of V1 neurons which either had or lacked strong surround suppression. These data were previously reported in Patterson et al. (2013a), but analyzed differently. In brief, the recordings used multielectrode array inserted into V1 of anesthetized monkeys, sampling neurons within 1 mm of the pial surface. We measured the spatial receptive fields of all units and then presented small gratings (1.3 deg diameter, full contrast, 1 cyc/deg, drift rate of 6.25 Hz), at 5 orientations varying from 0-90 degrees. We only analyzed sites for which the small gratings covered at least 50% of the spatial receptive field. Full methods are provided in Patterson et al. (2013a).

For each neuron, we first determined the degree of surround suppression, by comparing responses to the preferred orientation of the small grating, with that evoked by the same stimulus when it was surrounded by a large grating annulus (2 degree inner annulus, 7.4 degree outer diameter; otherwise identical parameters to the center grating). This allowed us to divide the

population into cells that were either strongly suppressed (defined as more than 40% suppression by the annulus) or not (less than 20% suppression).

We then proceeded with adaptation experiments. To determine the effects induced by an adapter within the receptive field, we adapted with small gratings and measured population responses to test stimuli offset by 0-90 degrees. Supplementary Figure 2 shows the ratio of these population responses (after adaptation compared to before) for cells with surround suppression (Supp. Figure 2A) and without (Supp. Figure 2B). In both subpopulations, the strongest reduction in population response occurs for stimuli matched to the adapter (a ratio near 0.5) with weaker effects for offset stimuli. However, the tuning of the adaptation-induced suppression was broader for cells with little surround-suppression: the population response for stimuli offset by 90 from the adapter was reduced by about 35% in non-suppressed cells (Supp. Figure 2B) but slightly facilitated in cells with strong surround suppression (see Dhruv et al., 2011; Patterson et al., 2013a for further quantification of effects induced by orthogonal adapters).

To determine how adaptation effects on the surround influence population responses, we adapted with an annular grating (2 deg inner diameter, 7.4 degree outer diameter). The annulus did not encroach on the cells' receptive fields (i.e. we required that it did not evoke a response above the mean spontaneous rate plus 0.5 times its standard deviation). The effects of annular adapters on population responses differed between the two groups of cells. In surround-suppressed cells (Supp. Figure 2A), an annular adapter caused responses to increase by nearly 50% (ratio of 1.46, $p<0.05$ based on a bootstrap test). However, responses to offset test stimuli were not significantly elevated, indicating that the tuning of this disinhibition was more stimulus-specific than the suppressive effects of adapting within the receptive field. For non-surround suppressed cells (Supp. Figure 2B), an annular adapter did not facilitate V1 population responses, but caused a weak loss of responsivity (mean ratio varied from 0.85 to 0.98). Note that, for both the surround-suppressed and non-suppressed populations, the number of cells differs between the two adaptation conditions (i.e. small gratings and annular adapters). These data were recorded from the same implants; the deviation is because the adaptation experiments were run in separate blocks, providing a slightly different yield.

Assuming these effects combine in a straightforward way, the predicted effect of adapting with a large grating (7.4 deg diameter) is a weak loss of responsivity for stimuli matched to the adapter (ratio of 0.91, obtained by multiplying the 0 deg offset point for the solid and dotted lines in Supp. Figure 2A). This is surrounded by a region of stronger adaptation-induced suppression for offset stimuli (ratios 0.58 and 0.75). For non-suppressed cells, there is substantial suppression for stimuli matched to the adapter (0.55) and offset from it (0.50 and 0.57, falling to 0.70 for orthogonal stimuli).

The assumption that effects of large grating adapters would involve the simple combination of the center and surround adapted separately (i.e. that the gain factors in Supplementary Figure 2 can simply be multiplied) is supported by the different effects of adapting with large and small stimuli on V1 tuning (Wissig and Kohn, 2012; Patterson et al., 2013). Small adapters cause tuning to shift away, whereas large gratings cause attractive shifts in preference. These attractive shifts require a broadly tuned loss of responsivity, and a sharply tuned disinhibition in the neighborhood of the adapter.

The V1 population ratios agree well in both magnitude and shape with the kernels derived from fitting a model of pattern selectivity to our measured MT responses. The primary point of deviation is the narrower tuning for the measured V1 effects in the surround-suppressed population. These differences could be due to our relatively limited sampling of neurons; our ignorance as to which, if any, of these neurons project to MT; and our presumed mixing of direction-selective and non-direction-selective neurons in our sampled V1 population.

In summary, the key components of the adaptation kernels of PDS and CDS cells are evident in V1 population spiking responses, when these are divided into suppressed (for PDS kernels) and non-suppressed (for CDS) populations. Namely, adaptation of surround suppressed V1 cells involves a tuned loss of responsivity and a sharply-tuned disinhibition or facilitation, corresponding to the kernels shown in Figure 4A. Adaptation of non-surround suppressed cells gives rise to a weakly tuned loss of responsivity, corresponding to the kernels in Figure 4C.

Supplementary Figure 3: Effect of adaptation on MT tuning and pattern selectivity, for a broad range of weighting profiles. Related to Figures 4, 5, and 6.

Fitting our model to the measured neuronal responses resulted in different adaptation kernels for PDS and CDS neurons (referred to here as PDS kernels and CDS kernels, respectively). Here we address two inter-related issues concerning this finding, and our explanation for the loss of MT pattern selectivity:

(1) Is it possible to account for MT adaptation effects in CDS and PDS cells using a single adaptation kernel? That is, can manipulations of V1-MT connectivity—the weighting profile—generate all of the observed MT effects from a single set of V1 adaptation effects?

(2) If it is not possible, which aspects of the kernels determine the effects measured in MT? What is the relative importance of the tuned disinhibition for stimuli similar to the adapter (0 deg) and the effect for stimuli far offset from the adapter (e.g. 180 deg), for which the PDS and CDS kernels also differ?

We addressed these issues by running simulations. First, we created a broad range of connectivity profiles by altering their bandwidth and the level of inhibition (Supp. Figure 3A; the dotted black lines indicate weights of zero. We illustrate the effects of adaptation for a subset of these profiles, outlined in the colored boxes. The PDS connectivity profiles (blue to green) produce PDS cells with pattern index values ranging from 3.6 to 1.3. The CDS connectivity (cyan to magenta) profiles produce CDS cells with pattern index values ranging from -3.4 to -2.0.

We explored the effects produced by the CDS (thick line; Supp. Figure 3B) and PDS (Supp. Figure 3C) kernels. To determine the importance of the kernel value for stimuli far offset from the adapter, we also considered two simple variants: for the CDS kernel, we shifted the kernel upward so that the value for stimuli offset from the adapter was similar to that seen in the PDS kernel (a value near 1; thin line in Supp. Figure 3B); for the PDS kernel, we shifted the kernel

downward so that its value for stimuli far offset from the adapter was similar to the original CDS kernel (a value near 0.75; thin line Supp. Figure 3C).

Supp. Figure 3D shows how applying the CDS kernels alters pattern selectivity in PDS cells, created by a range of weighting profiles (blue to green line color, corresponding to the profiles highlighted in Supp. Figure 3A). The data are plotted as a function of each model neuron's preference relative to the adapter. CDS kernels can reduce pattern selectivity, particularly for neurons whose preference is offset by 45 degrees or more from the adapter. However, preferred-adapted PDS neurons show an increase in pattern index, inconsistent with the physiology and the strong loss of perceptual coherence for plaids matched to the adapter direction. The effects on pattern selectivity are not dependent on the mean offset of the kernels: the original CDS kernel (thick lines) and its shifted version (thin lines) produce similar effects.

Supp. Figure 3E shows the corresponding simulations for the PDS kernels. These reduce pattern selectivity in PDS neurons, as shown in Figure 6 of the main text. For a range of connectivity profiles, the loss of pattern selectivity is robust for preferred- and null-adapted cells, with weaker effects for cells in between (e.g. offsets of 60-90 degrees). This is consistent with our perceptual and physiological data. The downward shift of the kernels has a minimal effect on the predicted loss of pattern selectivity (thick vs. thin lines). Note that weighting profiles without strong inhibition (darkest blues) can produce strong losses of pattern selectivity. The intuition provided by Figure 6 of the main text still applies in these cases—the loss of pattern selectivity simply involves a differential reduction in responsivity across plaids, rather than enhancement of responses to some plaids and reduction of responses others.

Supp. Figure 3F shows similar simulations, but for CDS cells (cyan to magenta colors, corresponding to the weighting profiles highlighted in Supp. Figure 3A). The component kernel causes an increase in the pattern index for cells with offsets of less than 90 degrees, and a small decrease for cells with a greater offset. Overall, there is a moderate increase in pattern selectivity (i.e. less component-like) for CDS cells, as we also observed in MT. Supp. Figure 3G shows the corresponding simulations for PDS-kernels. These also predict that CDS cells become less component-like after adaptation.

The adaptation kernels must not only account for the change in pattern selectivity, but also for the change in MT tuning for gratings (Figure 5 of the main text). The CDS and PDS kernels also make distinct predictions for these effects, across a wide range of connectivity profiles. For CDS kernels, there is a strong loss of responsivity for preferred-adapted PDS cells (green to blue colors in Supp. Figure 3H) and CDS cells (cyan to magenta colors). Similarly these kernels produce strong repulsive shifts for flank-adapted neurons (positive values in Supp. Figure 3J). The magnitude of these effects is influenced by the shift in the kernels (thick vs. thin symbols).

In contrast, the PDS kernels result in maintained responsivity in both PDS and CDS cells (symbols with thick lines in Supp. Figure 3I, clustered around a ratio of 1), unless the kernels are shifted downward (thin symbols in 3I, corresponding to the thin lines in 3C). These kernels also produce attractive shifts in flank-adapted PDS and CDS cells (Supp. Figure 3K).

To summarize, we relate this broad set of observations back to our original questions:

(1) Our simulations strongly suggest that it is not possible to account for the observed MT effects with a single common kernel. The effects of kernels on tuning for gratings are straightforward and qualitatively similar across a broad range of connectivity profiles. CDS kernels result in a strongly loss of responsivity in preferred-adapted MT cells (3H) and either no change in preference or repulsive shifts (3J). We observed these effects in MT in CDS but not PDS cells. We found no weighting profile that could reproduce the effects observed in PDS cells in MT, using CDS kernels. Only PDS kernels could produce these effects observed in PDS cells in MT—maintained responsivity for preferred-adapted cells and attractive shifts in preference for flank-adapted ones. In addition, only PDS kernels result in a strong loss of pattern selectivity for preferred- and null-adapted MT cells (3E).

In these simulations, we treated the weighting profile as a "free parameter", in trying to account for adaptation effects in MT. It is important to note that in fact the weighting cannot be adjusted arbitrarily: it is strongly constrained by the need to relate the grating and plaid tuning of each cell, before adaptation. We confirmed this by comparing the profiles obtained by fitting just the pre-adaptation responses of each cell, with those obtained by fitting our full model (i.e. the pre- and post-adaptation data together, as in the main text). The median circular correlation coefficient between the two weighting profiles was 0.97. Thus, the weighting profile was essentially determined by the need to account for the pre-adaptation grating and plaid tuning and not modifiable to account for the adaptation data.

(2) Our simulations show that the essential difference between the CDS and PDS kernels is the tuned disinhibition for stimuli similar to the adapter. It is this kernel feature that generates the expected pattern of changes in pattern selectivity, responsivity for gratings, and shifts in preference. Shifting the kernels up and down (thin vs. thick lines in Supp. Figure 3) had a more muted effect on the predicted effects of adaptation.

Details of the simulations: The weighting profiles were obtained by varying the bandwidth parameter (the multiplicative factor in the exponent of the circular Gaussian function) from 0.1 to 1.85, in steps of 0.25. The inhibition in the weighting profile was varied from 0.6 to 0.9 in steps of 0.06. This was implemented as a uniform negative shift in the profiles. Note that the physiological data were fit with tuned inhibitory weights to account for the sometimes irregular shaped tuning curves (e.g. responses to one component being larger than the other in CDS cells). The scaling factor and exponent (k, and p of Eq. 4) of the non-linearity were fixed at 1 and 3.5, respectively. Changes in these parameters had only small quantitative effects on the reported results. To calculate the change in the pattern selectivity, we simulated 20 trials of responses, chosen from a Poisson distribution with mean rate defined by the simulated tuning curves. This was necessary because the tuning curves, uncorrupted by noise or variability, can generate pattern index values not seen physiologically (PI values >10).

Supplementary Figure 4: Analysis of data arising from most well-isolated single units. Related to Figure 2.

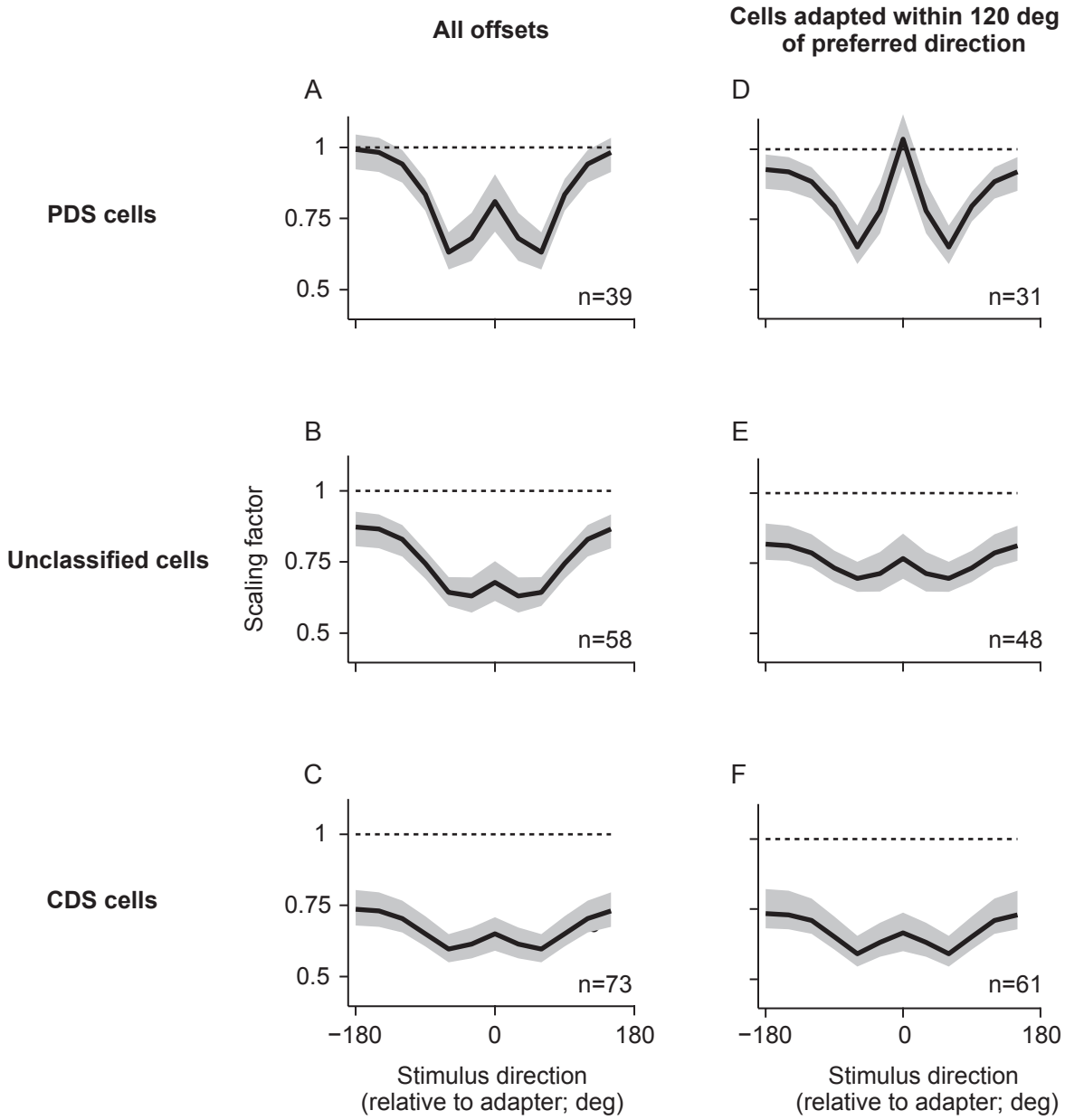
We used an array of high-impedance electrodes and tetrodes to record responses in MT. We sorted these data offline, using standard software (Plexon Offline Sorter) and criteria. Although

isolation quality was good, we wished to ensure that there was no relationship between recording quality and the disruption of MT pattern selectivity we report. We therefore calculated, for each unit adapted with the common stimulus ensemble (n=133), a measure of tuning quality—namely, the waveform signal-to-noise ratio, or SNR (Kelly et al., 2007). Supplementary Figure 1 shows the distribution of SNR values for the recorded population. We then quantified adaptation effects in units whose data fell in the top 1/3 of the SNR distribution—corresponding to a threshold of 3.73. Note that SNR values for tetrode waveforms tend to be lower than for single electrodes of comparable quality (e.g. of multielectrode array recordings, for which we have performed similar analysis; Wissig and Kohn, 2012). This is because well-isolated units on tetrodes often include tetrodes sites (sometimes several) with no detectable waveforms, effectively lowering the SNR. Thus, a threshold of 3.73 is exceedingly strict. Nevertheless, effects were similar in this subset of the data to the results reported in the main text. For PDS cells (n=13), there was a significant loss of pattern selectivity (PI fell from 3.53±0.56 to 0.99±0.47, p=0.003). For unclassified cells, the PI fell slightly from -0.01±0.18 to -0.15±0.35 (p=0.67). For component cells (n=16), the PI increased slightly, from -2.41±0.21 to -1.77±0.37 (p=0.08).

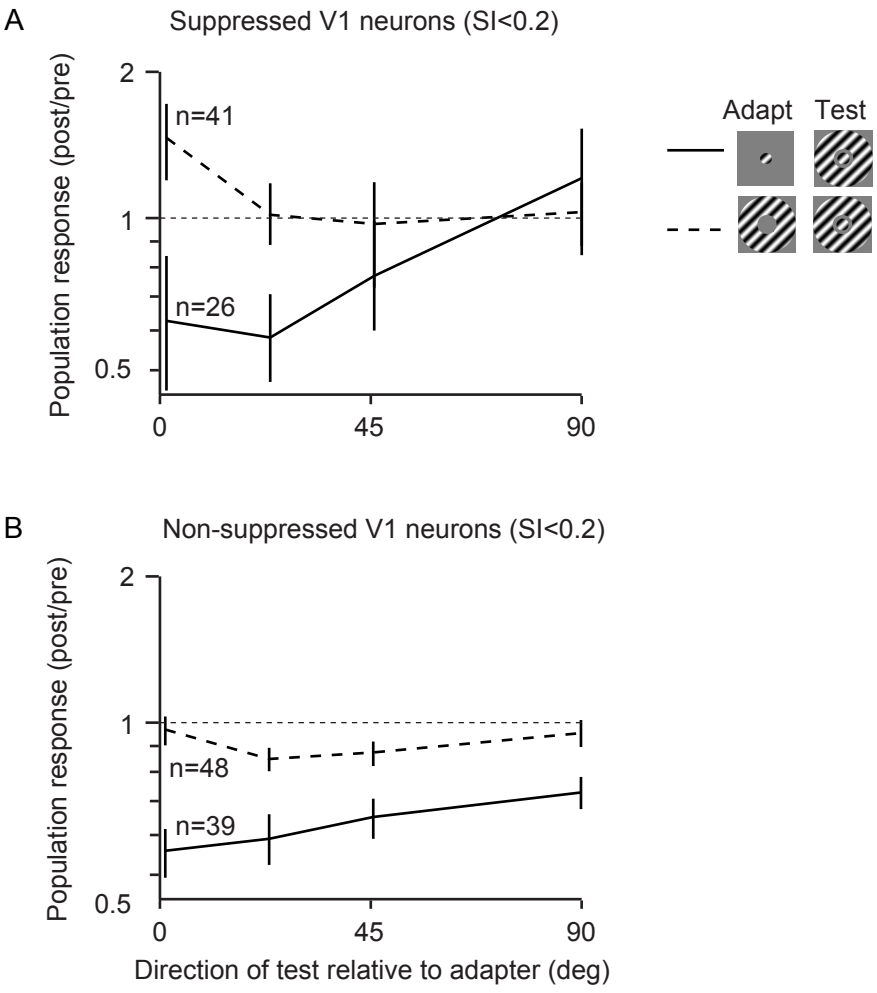
Supplemental References:

Dhruv NT, Tailby C, Sokol SH, Lennie P (2011) Multiple adaptable mechanisms early in the primate visual pathway. J Neurosci. 31: 15016-15025.

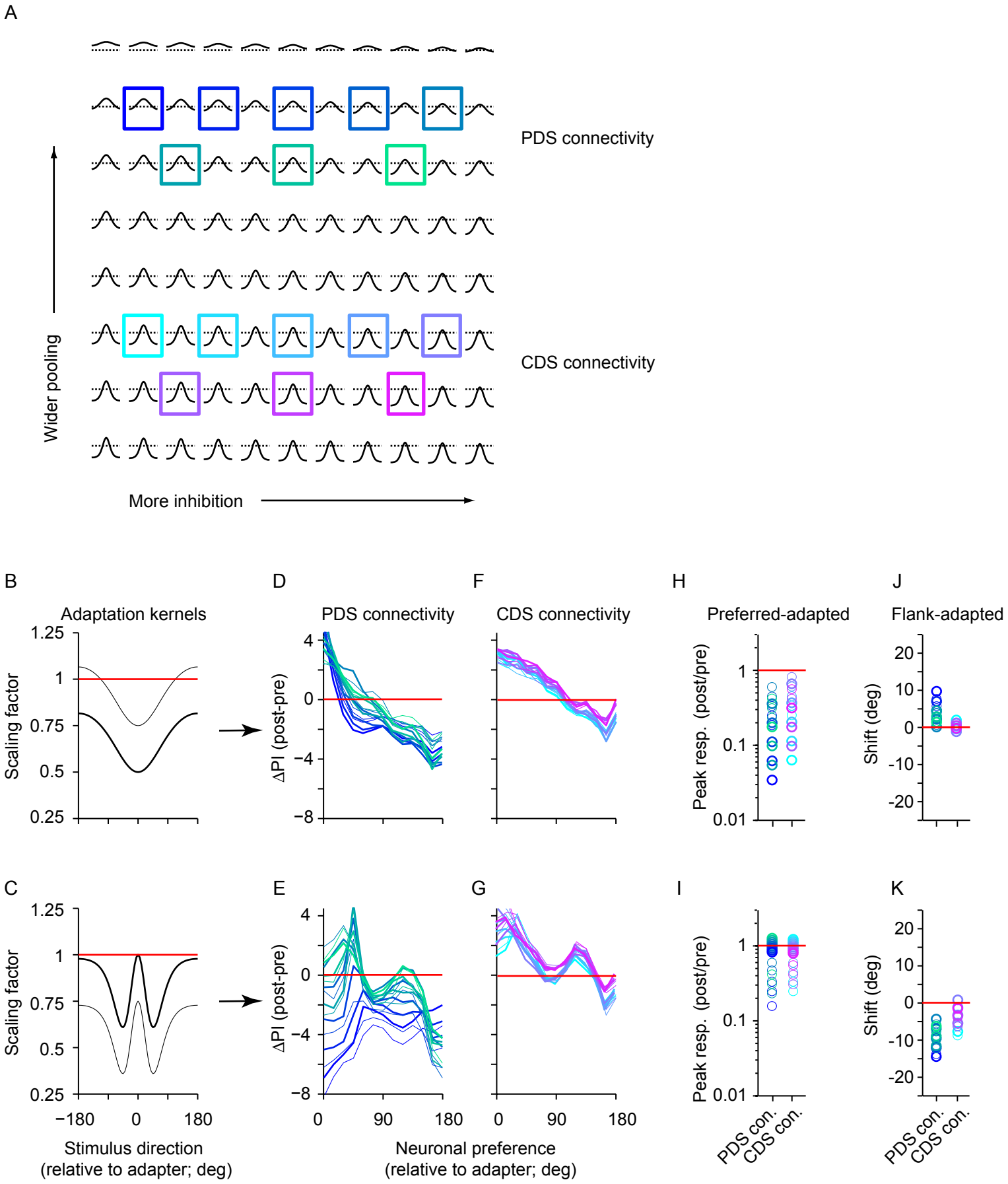
Kelly RC, Smith MA, Samonds JM, Kohn A, Bonds AB, Movshon JA, Lee TS (2007) Comparison of recordings from microelectrode arrays and single electrodes in the visual cortex. J Neurosci. 27: 261-264.



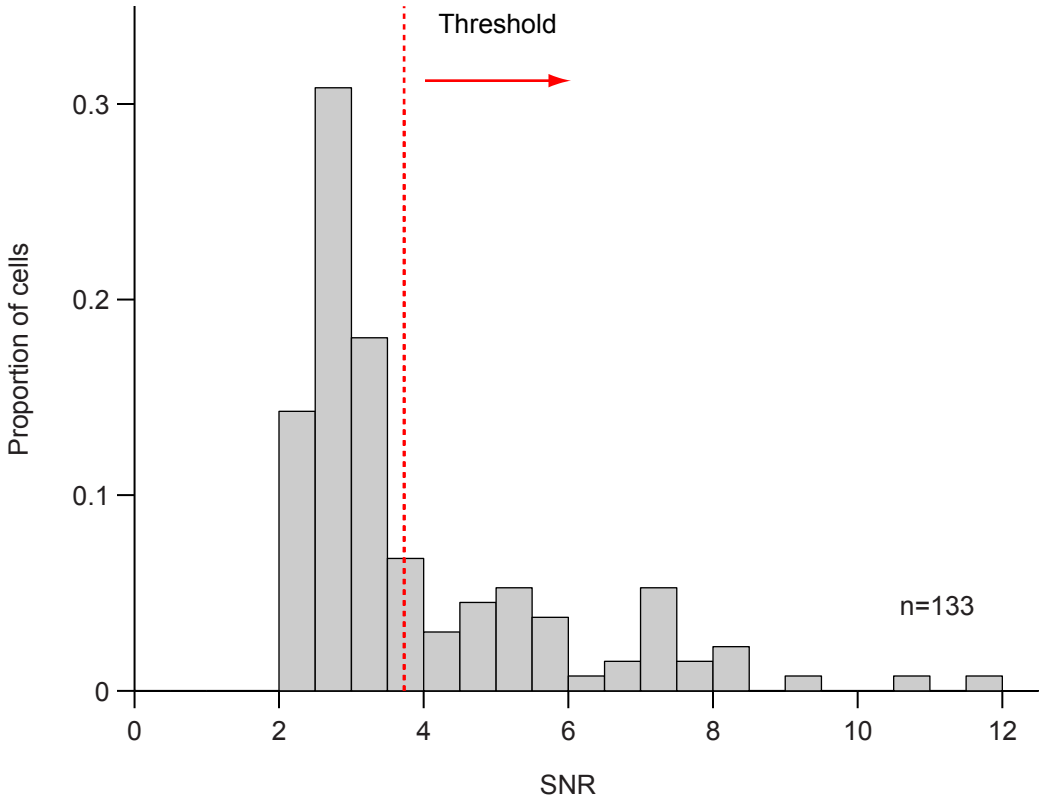
Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3



Supplementary Figure 4