Details of the model fitting

We fitted a model to the mean neuronal response; that is, to the normalized firing rate, D_{ij}, where i,j indexes the stimulus present in left, right eyes. The six bar positions are indexed by i=1…6, and we use i=0 to indicate that no bar was present. We try and make the model response M_{ij} as close to the observed D_{ij} as possible.

The full model has 14 parameters. L_i , R_i is the input from the left, right eye when there is a bar at position i (i=1…6); b is the tonic (background) input; and γ is the output exponent.

The model response M_{ij} is thus as follows. The background response of the model neuron in the absence of visual stimulus is

$$M_{00} = b^{\gamma};$$

the response to a monocular bar at the i^{th} position in the left, right eye is

$$M_{i0} = \lfloor L_i+b \rfloor^{\gamma} \text{ , } M_{0i} = \lfloor R_i+b \rfloor^{\gamma};$$

and the response to binocular bars at the i^{th} position in the left eye and the j^{th} position in the right eye is

$$M_{ij} = \lfloor L_i+R_j+b \rfloor^{\gamma},$$

where $\lfloor x \rfloor = x$ if x>0 and 0 otherwise (i.e. a threshold at 0). The sum of squared errors between model and data is

$$\epsilon = \sum_{i=0}^{6} \sum_{j=0}^{6} \left(D_{ij} - M_{ij}\right)^2$$

where D_{ij} is the mean normalised firing rate of the neuron, and bar position 0 means no bar present.

We also included a regularisation term intended to keep parameter values close to zero except where they clearly improved the fit. This term was equal to one-thousandth of the summed squared parameters:

$$\Lambda = [b^2 + \gamma^2 + \sum_{i=1}^{6}(L_i^2 + R_i^2)]$$

The 13 parameters L_i , R_i and γ were adjusted so as to minimise $\varepsilon+\Lambda$ across all data. The parameter b was not fitted as a free parameter, but was constrained so as to account for the background firing rate of the cell given the fitted output exponent γ, i.e. we set

$$b = D_{00}^{(1/\gamma)}$$

In practice, this constraint makes little difference compared to fitting all 14 parameters together freely.

Fitting was carried out by the Matlab routine FMINSEARCH. Convergence to local optima can be a problem in such multi-parameter optimisation, and the choice of initialisation is often critical. We started by doing a 12-parameter fit with γ constrained to 1 and b set to D_{00}. We explored two initialisations: a flat initialisation (L_i=1, R_i=1 for all i), and an initialisation reflecting the monocular responses ($L_i=D_{i0}-D_{00}$, $R_i=D_{0i}-D_{00}$), and selected whichever gave the lowest fit error $\varepsilon+\Lambda$. We then removed the constraint on γ, and performed the full 13-parameter fit. We again explored two initialisations: the L_i,R_i found by the 12-parameter fit, and the monocular-response initialisation $L_i=D_{i0}-D_{00}$, $R_i=D_{0i}-D_{00}$, both with γ=1, and again chose whichever yielded the lowest fit error. We found that with these methods, the optimisation converged rapidly and reliably.

To compute percentage of variance explained, we computed the total variance of the mean observed response in each stimulus condition:

$$T = \frac{1}{48}\sum_{i=0}^{6}\sum_{j=0}^{6}\left(D_{ij} - \overline{D}\right)^2$$

and the residual variance of the difference between the mean observed response and the model fit in each stimulus condition:

$$R = \frac{1}{48}\sum_{i=0}^{6}\sum_{j=0}^{6}\left(D_{ij} - M_{ij} - \overline{D} + \overline{M}\right)^2;$$

the percentage of variance explained is then PV = 100(T-R)/T.