

*Vision Research*, **128**, 68-82 (2016). Green Open Access. Accepted 12 Sept 2016.  
(no fee, 12 month embargo, CC-BY-NC-ND user license, satisfies RCUK/BBSRC requirements)  
Colour figures online only. Final version published online: 14-OCT-2016

## **Binocular functional architecture for detection of contrast-modulated gratings**

Mark A. Georgeson<sup>1</sup> & Andrew J. Schofield<sup>2</sup>

1. School of Life & Health Sciences, Aston University, Birmingham, B4 7ET, UK
2. School of Psychology, Birmingham University, Birmingham, B15 2TT, UK

*Running head:* Binocular summation for contrast modulation

*Corresponding author:* [m.a.georgeson@aston.ac.uk](mailto:m.a.georgeson@aston.ac.uk)

### **Highlights**

- Detection of contrast modulation (CM) shows full binocular summation
- Similarity or dissimilarity of the carriers has no effect on summation
- Out-of-phase envelopes don't cancel, but are a bit less detectable than monocular ones
- Model: monocular envelope extraction, then contrast-weighted binocular summation
- Monocular CM outputs in parallel with binocular ones; largest response wins

*Key words:* Binocular summation; Gratings; luminance modulation; contrast modulation; second-order filters

*Disclosure:* The authors report no conflicts of interest and have no proprietary interest in any of the materials mentioned in this article.

*Acknowledgement:* This work was supported by a BBSRC grant (BB/H00159/X1) to Mark Georgeson & Tim Meese, and EPSRC grant (EP/F026269/1) to Andrew Schofield.

To access the research data underlying this publication, please see:  
<http://dx.doi.org/10.17036/cb87e020-4fdc-4f77-a187-ec00ff934d20>

DOI: 10.1016/j.visres.2016.09.005

## Abstract

Combination of signals from the two eyes is the gateway to stereo vision. To gain insight into binocular signal processing, we studied binocular summation for luminance-modulated gratings (L or LM) and contrast-modulated gratings (CM). We measured 2AFC detection thresholds for a signal grating (0.75 c/deg, 216msec) shown to one eye, both eyes, or both eyes out-of-phase. For LM and CM, the carrier noise was in both eyes, even when the signal was monocular. Mean binocular thresholds for luminance gratings (L) were 5.4dB better than monocular thresholds - close to perfect linear summation (6dB). For LM and CM the binocular advantage was again 5-6dB, even when the carrier noise was uncorrelated, anti-correlated, or at orthogonal orientations in the two eyes. Binocular combination for CM probably arises from summation of envelope responses, and not from summation of these conflicting carrier patterns. Antiphase signals produced no binocular advantage, but thresholds were about 1-3dB higher than monocular ones. This is not consistent with simple linear summation, which should give complete cancellation and unmeasurably high thresholds. We propose a three-channel model in which noisy *monocular* responses to the envelope are binocularly combined in a contrast-weighted sum, but also remain separately available to perception via a *max* operator. Vision selects the largest of the three responses. With in-phase gratings the binocular channel dominates, but antiphase gratings cancel in the binocular channel and the monocular channels mediate detection. The small antiphase disadvantage might be explained by a subtle influence of background responses on binocular and monocular detection.

# 1. Introduction

The analysis of spatial information in vision unfolds over successive stages of the retino-cortical pathways, and involves combination of signals from the two eyes. A good deal is known from both psychophysics and neurophysiology about the spatial analysis and binocular combination of signals derived from spatial variations of luminance in the retinal image - often called *first-order* information - but rather less is known about mechanisms supporting the representation of *second-order* information, arising from spatial variation in higher-order image properties such as local contrast, local orientation or texture density (Cavanagh & Mather, 1989; Chubb & Sperling, 1989; Landy & Bergen, 1991). Over an ensemble of natural images, spatial variations in local luminance and local contrast amplitude (first- and second-order structure) were found to be uncorrelated (Schofield, 2000), while in the laboratory the two kinds of structure can be usefully isolated in computer-generated synthetic images (Fig. 1). This experimental approach has yielded much evidence for the idea of separate pathways encoding first- and second-order motion (see literature summary in Table 1 of Clifford & Vaina, 1999), with these paths perhaps converging to produce an integrated perception (e.g. Lu & Sperling, 1995; Scott-Samuel & Georgeson, 1999; Wilson, Ferrera, & Yo, 1992). Our studies of first- and second-order grating detection, and perceptual aftereffects, revealed a similar picture of separate encoding pathways responding to the spatial structure of luminance modulation (LM) and contrast modulation (CM) (Georgeson & Schofield, 2002; Schofield & Georgeson, 1999). In this paper we focus specifically on contrast modulation (CM) as a second-order property (see Fig. 1, bottom row), and we ask some basic questions about the binocular processing of CM signals.

In first-order vision, stereo disparity is encoded by populations of binocular neurons that combine input from monocular receptive fields that have similar size and position, and similar selectivity for orientation and direction, but are driven separately by the left and right eyes (e.g. Hubel & Wiesel, 1968; Ohzawa, DeAngelis, & Freeman, 1996). This neural binocular summation leads to a behavioural *binocular advantage*: contrast detection thresholds for luminance gratings of the same orientation, SF and spatial phase shown to both eyes are markedly better than for gratings shown to one eye (Anzai, Bearnse, Freeman, & Cai, 1995; Campbell & Green, 1965; Meese, Georgeson, & Baker, 2006; Simmons & Kingdom, 1998).

Since stereo depth discrimination is possible on the basis of CM disparity alone (Edwards, Pope, & Schor, 2000; Hess & Wilcox, 2008; Langley, Fleet, & Hibbard, 1999; Wilcox & Hess, 1996), it is natural to ask whether there is a corresponding binocular advantage for detection of CM. We test this by measuring detection thresholds for CM gratings presented to one eye or to both eyes, when the carrier pattern (dynamic noise texture, Fig. 1) is shown to both eyes. If a clear binocular advantage is found for CM, then we can ask whether it arises from binocular summation of second-order contrast envelope (CM) signals, or whether it might be inherited from summation of the first-order carrier signals. We did this in three experiments by assessing the binocular advantage for CM with pairs of noise carriers that were the same (perfectly correlated) in the two eyes, and comparing it with conditions where the carriers were (1) uncorrelated or anti-correlated, (2) uncorrelated but with the same or orthogonal orientations, and (3) uncorrelated but with the same or opposite contrast polarity in the two eyes. If matching carrier signals are important for summation then binocular CM performance should be better when the carriers are the same or similar than when they are very dissimilar in spatial correlation, orientation or polarity. The use of oriented carriers (Experiment 2) and arrays of light vs dark blobs (Experiment 3) might also be informative in light of the suggestion that second-order channels specific to carrier orientation, and channels specific to light/dark polarity, both contribute to CM detection (Motoyoshi & Kingdom, 2007). In a fourth experiment we examined the nature of the combination process for CM: whether it might be better described as a simple linear sum of each eye's modulation, or as a contrast-weighted sum in which the

contribution made by each eye is driven by the carrier contrast visible to that eye (Zhou, Georgeson, & Hess, 2014).

We might also learn a good deal about the combination process by comparing detection of in-phase and out-of-phase ('antiphase') binocular inputs. For example, if binocular summation were strictly linear, then antiphase inputs should cancel each other, and be undetectable. In this way, and guided by computational modeling, we aim to build a picture of the functional architecture for binocular CM processing.

FIG 1 HERE

## 2. Methods

In all experiments a two-alternative forced-choice (2AFC) staircase method was used to estimate thresholds for detection of contrast-modulated (CM) gratings, and in experiment 1 for luminance gratings (Lum) and luminance-modulated noise (LM) gratings as well. Details of image generation, display and procedure are given here for Experiment 1. Changes in conditions and procedure for Experiments 2-4 will be noted in the Results.

### 2.1. Experiment 1

#### *Image display*

Image arrays were generated in *Matlab* on a Macintosh G4 computer and displayed using *PsychToolbox* software (Brainard, 1997) on a Clinton Monoray CRT monitor with a fast-decay yellow-green phosphor, calibrated and gamma-corrected using a Minolta LS110 computer-controlled digital photometer. A Cambridge Research Systems Bits++ box was used in Mono++ mode to render image intensity with high accuracy (14-bit greyscale resolution; 16384 grey levels). Images were viewed through frame-interleaving FE1 goggles (CRS Ltd.) to present separate images to the two eyes with very little crosstalk (less than 1%). The high frame rate (150 Hz; 75 Hz per eye) ensured that the alternating display appeared as a steady image with no visible flicker. Mean luminance as measured through the frame-interleaving goggles was about 25 cd/m<sup>2</sup>.

#### *Grating images*

The variety of binocular images used in experiment 1 is illustrated in Fig. 1. Horizontal, sinusoidal gratings were viewed foveally in a 5 deg square aperture whose outer border contained a high contrast binary random texture that was always the same for both eyes. Its purpose was to act as a 'fixation lock' to stabilize binocular convergence. The modulation of luminance or carrier contrast that defined the grating was confined to a circular, smooth-edged window 4 deg in diameter at half-height, while the carrier (when present) filled the whole of the square aperture (see Fig. 1).

- The luminance grating was defined by:

$$Lum(x, y) = L_0 \{1 + m \cdot \sin(2\pi f y - \phi) \cdot win(x, y)\} \quad (1)$$

where  $L_0$  is mean luminance,  $m$  is the signal modulation to be detected,  $f$  is spatial frequency (0.75 c/deg),  $y$  is vertical position in deg (0 at the centre),  $\phi$  is spatial phase in radians (0 or  $\pi$ ), and  $win(x, y)$  is the smooth circular window (0 on the outside, rising to 1 on the inside, following a smooth, half-cycle raised-cosine transition 0.5 deg wide).

- The LM grating was the same as the Lum grating, but with carrier noise added:

$$LM(x, y) = L_0 \{1 + m \cdot \sin(2\pi f y - \phi) \cdot win(x, y) + c \cdot N(x, y)\} \quad (2)$$

where  $c$  is the RMS contrast of the carrier, and  $N$  is 2D, isotropic, bandpass spatial noise with a mean of 0 and standard deviation of 1. The noise centre frequency was 6 c/deg, and the log Gaussian filter used to produce it had a bandwidth of  $\pm 0.5$  octaves at half-height.

- The CM grating was analogous to the LM grating, but with signal and carrier noise multiplied rather than added:

$$CM(x, y) = L_0 \{1 + (1 + m \cdot \sin(2\pi f y - \phi)) \cdot c \cdot N(x, y)\}. \quad (3)$$

This definition creates second-order, contrast modulation (CM) but with no first-order Fourier component at the signal frequency. For CM, we can also define the *amplitude*  $a$  of contrast variation as the difference between peak and mean contrast. Just as contrast is a relative measure of luminance variation, so contrast modulation depth  $m$  is a relative measure of contrast variation (range 0 - 1) expressed as a proportion of the carrier contrast. Thus,  $m = a/c$ , and  $a = m \cdot c$ .

### *Timing & Binocularity*

On each trial, the two observation intervals lasted for 216 ms each, separated by a blank (mean luminance) interval of 500 ms. When present, the carrier noise was dynamic (it 'twinkled' over time). More specifically, each 216 ms interval showed a sequence of four different, uncorrelated noise samples with identical modulations, for 54 ms each. The texture border, and a 2x2 pixel, black, central fixation point were present throughout the trial. The texture border was randomly re-sampled between trials.

There were three types of signal presentation: Mon, Bin, Anti (see Fig. 1). (i) Signal modulation was presented to one eye (*Mon*; with the eye (left or right) and phase (0 or  $\pi$ ) randomly chosen on each Mon trial), (ii) signal to both eyes in-phase (*Bin*; both phases 0, or both  $\pi$ , randomly chosen), or (iii) signal to both eyes in anti-phase (*Anti*; left and right eye phases (0,  $\pi$ ) or ( $\pi$ , 0), randomly chosen). For both LM and CM, the carrier was present in both eyes, even when signal modulation was monocular. These left and right eye carriers could be identical (correlation  $r = 1$ ), or independent, uncorrelated samples ( $r = 0$ ), or anti-correlated between the eyes ( $r = -1$ ). When  $r = -1$ , the two noise samples were the same, but with positive contrast ( $c$ ) in one eye and negative ( $-c$ ) in the other eye.

### *Procedure & Data analysis*

On each trial, the observer's 2AFC task was to press a key to indicate whether the first or second interval contained the signal modulation. Auditory feedback about correctness was given after each response. Performance level was controlled by an up-down staircase procedure in which signal modulation depth ( $m$ ) was reduced by 2dB after 3 correct responses, and increased by 2dB after one error. Within a session, a pair of randomly interleaved staircases (40 trials each) was run for each signal type (Mon, Bin, Anti) and trials for these three signal types were also randomly interleaved. Importantly, this meant that observers were very unlikely to know whether the signal was monocular or binocular on any given trial. Lum, LM and CM, and the three levels of interocular correlation  $r$  were tested in separate sessions, and the whole procedure was repeated twice for each observer. Session orders were counterbalanced across observers. Frequencies of correct responses at each signal level were pooled over the two staircases and two sessions (total 160 trials) and a Weibull psychometric function:

$P(\text{correct}) = 1 - 0.5 \exp\{-(m/m_{thd})^\beta\}$  was fitted by the method of maximum likelihood (ML) to derive a modulation detection threshold ( $m_{thd}$ , defined at 81.6% correct) and psychometric slope,  $\beta$ , for each observer and condition. The staircase procedure delivers a very variable number of trials at different stimulus levels, but the ML method automatically and correctly allows for this. We often express luminance or contrast modulation thresholds on a log scale in decibels (dB), defined as  $20 \cdot \log_{10}(100m_{thd})$ , thus making a modulation of 1% = 0dB, 10% = 20dB and 100% = 40dB. A difference of 6dB is a factor of two difference in threshold.

Seven observers were tested (4 undergraduate students with little experience of psychophysics, and 3 experienced including the two authors). Observers wore spectacle corrections where necessary. Practice sessions were given. Informed consent was obtained and the work was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

FIG 2 HERE

## 2.2. Model

### *Conceptual framework*

An outline of our model for binocular detection of contrast modulation is sketched in Fig. 2. Following a standard filter-rectify-filter (FRF) approach to second-order vision (e.g. Chubb & Sperling, 1989; Lu & Sperling, 1995; Schofield & Georgeson, 1999; Solomon & Sperling, 1994; Zhou & Baker, 1993) the first stage is assumed to recover the monocular contrast envelopes by spatial filtering at multiple orientations, followed by full-wave rectification. Summation of these responses over first-order filter orientations (Dakin & Mareschal, 2000; Mareschal & Baker, 1999; Motoyoshi & Nishida, 2004) and over time (Manahilov, Calvert, & Simpson, 2003; Schofield & Georgeson, 2000) should improve the quality of the recovered envelope signal. Recent experiments on the detection of other forms of second-order structure (modulations of orientation or contrast polarity) imply that first-order filters map onto second-order filters in several distinct 'streams' (Motoyoshi & Kingdom, 2007), but for detection of CM the pooling over first-order orientations shown in Fig. 2 should be sufficient (Motoyoshi & Nishida, 2004). Fig. 3 (column *iii*) illustrates how the rectification introduces energy at the modulation frequency. At stage 2, lower frequency filters respond to this horizontal modulation signal and reject much of the carrier noise. Independent Gaussian noise is added to these two monocular response images to represent neural noise and other processing inefficiencies, bundled into a single source. Monocular envelope extraction is followed by binocular combination (Tanaka & Ohzawa, 2006; Wilcox & Hess, 1996).

We shall see that four new, less orthodox, features of our model (Fig. 2) are important in capturing the experimental behaviour. They are that (a) a steady background response ( $r_0$ ) is added to the two monocular signal amplitudes, (b) the monocular signals (L,R) are preserved alongside their binocular sum (B), (c) the contrast-driven gains (weights  $w_L$ ,  $w_R$ ) that control the binocular sum (Zhou, Georgeson, & Hess, 2014) are also applied to the monocular signals, and (d) the observer's trial-by-trial decision is based on whichever of the three response arrays has the greatest *amplitude* ( $r_L$ ,  $r_B$ ,  $r_R$ ) at the signal frequency: a *max* operator. [This use of response amplitude, irrespective of spatial phase, is equivalent to assuming that the observer has full knowledge of the signal orientation and spatial frequency, but is completely uncertain about its spatial phase, and therefore has to rely on amplitude alone. This seems reasonable, given that in the experiments signal orientation and SF were fixed but signal phase was randomly 0 or 180° from trial to trial.] Fig. 3 illustrates the noisy output of the L and R channels (column *iv*) and the B channel (column *v*) for the four signal conditions of interest (Binoc, Monoc, Antiphase and No modulation). Fig. 3 (column *vi*) illustrates graphically our simplifying assumption that signal amplitude is computed for each channel (L,B,R) as a single number for the whole image area, and that the *max* operator then selects the largest of the three, rather than conducting a piecemeal or point-by-point comparison.

FIG 3 HERE

### *Implementation to predict modulation thresholds*

To gain speed and efficiency in fitting the model to experimental data we simplified the implementation in two key ways, exploiting the fact that modulation was sinusoidal. Firstly, the full image-processing simulation (Fig. 3(*iii*)) confirms that we can treat the output of stage 1 as a

noisy sinusoidal waveform (allowing us to drop the computations of stage 1). Secondly, the noisy 2D image of a sinusoid can be efficiently represented by just two numbers: a two-element vector (or complex number) representing the even (cosine) and odd (sine) Fourier component amplitudes of the signal. The noise simply adds a Gaussian (zero mean) random number to each of these components, and thus introduces fluctuation in the amplitude and phase of the resultant vector. We ran numerical simulations to check that in its perturbation of signal amplitude and phase this efficient method was exactly equivalent to adding 2D Gaussian noise to the sine-wave image itself. Interestingly, when amplitude is considered irrespective of phase (see above), the mean Fourier amplitude at the signal frequency is above zero even when the signal amplitude is zero. This is so because even in 'pure' noise (no signal) there is a non-zero Fourier component at the signal frequency. Its phase is random, but its vector amplitude is by definition non-negative, and so has a positive mean. Importantly, for the phase-uncertain observer who computes amplitude irrespective of phase, this leads to a nonlinear relation between response amplitude and signal modulation depth (Fig. A1(A,B)). The *max* operator contributes further to the nonlinearity of mean output (*Resp*) against signal strength (Pelli, 1985; Tyler & Chen, 2000).

More formally, the weighted responses of the two monocular pathways in this Fourier vector [sine, cosine] form are:

$$\vec{r}_L = w_L \cdot [a_L \sin(\phi_L) + N(0, \sigma), a_L \cos(\phi_L) + N(0, \sigma)] \quad (4)$$

$$\vec{r}_R = w_R \cdot [a_R \sin(\phi_R) + N(0, \sigma), a_R \cos(\phi_R) + N(0, \sigma)] \quad (5)$$

where  $a_L, a_R$  are the left and right eye amplitudes of contrast modulation (equal to modulation depth  $m$  times carrier contrast  $c$ ; see Sec. 2.1/*Grating images*),  $\phi_L, \phi_R$  are the signal phases,  $N(0, \sigma)$  is an independent sample of Gaussian noise (mean = 0, variance =  $\sigma^2$ ), uncorrelated between eyes and across trials. Following Zhou, Georgeson & Hess (2014), the second-order weights  $w_L, w_R$  are driven by first-order contrast:

$$w_L = \frac{c_L}{c_L + c_R}, \quad w_R = \frac{c_R}{c_L + c_R} \quad (6)$$

where  $c_L, c_R$  are the left and right carrier contrasts. Note that the left eye weight, and hence response amplitude, is strongly affected by the presence or absence of carrier contrast in the right eye, and vice-versa. In this scheme, even the 'monocular' second-order responses are subject to a form of interocular suppression, and we tested for this in experiment 4. The binocular response is the vector sum of the two noisy monocular vectors

$$\vec{r}_B = \vec{r}_L + \vec{r}_R \quad (7)$$

The corresponding response amplitudes (just three numbers) are:

$$r_L = r_0 + |\vec{r}_L|, \quad r_R = r_0 + |\vec{r}_R|, \quad r_B = 2r_0 + |\vec{r}_B| \quad (8)$$

where  $r_0$  is a constant background level of response, unrelated to signal modulation depth. We make no assumption about the source of  $r_0$ , but it is plausible that the background response might be driven by mean luminance, or carrier contrast, or both. We may think of it as a psychophysical counterpart to the 'spontaneous firing rate' or background discharge exhibited by neurons at many stages of the visual pathway. Note that we also assumed binocular summation of background responses in  $r_B$  (eqn. 8). The final output *Resp* from one interval of a given trial is given by the channel with the greatest response amplitude:

$$Resp = \max\{r_L, r_B, r_R\} \quad (9)$$

and in the 2AFC detection task the observer chooses the interval that gave the greater *Resp.* In our Monte Carlo simulation the proportion of correct responses for each stimulus condition was computed across 50,000 trials at each of 32 levels of modulation from 0 to 0.5. A Weibull function was fitted in the same way as for the experimental data to derive a threshold, and psychometric slope  $\beta$ . Models were fitted to data using the simplex algorithm (*fminsearch* in *Matlab*) to minimize the squared error between predicted and observed CM thresholds in dB, usually with adjustment of two free parameters,  $r_0$  and  $\sigma$ . With this model architecture (Fig. 2; eqns 4-9), noise standard deviation  $\sigma$  controls the overall level of sensitivity, but the pattern of relative thresholds depends on the model architecture, and is fine-tuned by  $r_0$ .

FIG 4 HERE

### 3. Results

#### 3.1. Experiment 1. Detection of LM & CM with 2-D noise carriers

The first main result was strikingly simple: group mean thresholds for in-phase binocular modulation were close to half the monocular thresholds, in all cases. Fig. 4 shows the group mean thresholds in dB, relative to a monocular baseline. The group mean *binocular advantage* for luminance gratings (Lum) was  $5.42 \text{ dB} \pm 0.54 \text{ s.e.m.}$ ; for gratings in noise (LM,  $c=0.2$ ) it was  $5.31 \text{ dB} \pm 0.22$ , and for contrast-modulated gratings (CM)  $6.09 \text{ dB} \pm 0.31$  for  $c=0.2$ , and  $5.40 \text{ dB} \pm 0.12$  for  $c=0.1$ . These were close to the 6dB (factor of 2) improvement (grey band in Fig. 4) that could be expected from perfect linear summation of in-phase signals across the eyes, with fixed noise.

The antiphase condition, however, showed emphatically that linear binocular summation is not the whole story. With linear summation, out-of-phase signals should completely cancel, and so antiphase thresholds should be unmeasurably high. But this was not so: the mean antiphase thresholds (Fig. 4) were about the same as the monocular thresholds (for Lum), or around 1 to 3dB higher (for LM and CM). The slightness of this *antiphase disadvantage* must put strong constraints on any model of binocular combination.

Comparing CM performance for  $c=0.2$  vs  $c=0.1$ , we found that the relation between Mon, Bin and Antiphase thresholds was almost identical at the two carrier contrasts (Fig. 4C,D), except that thresholds were on average 1.4 dB lower when the carrier contrast was doubled (Fig. 8). This was very similar to the shallow improvement of CM sensitivity with increasing carrier contrast described by Schofield & Georgeson (1999). It is quite important that this observed improvement is small in the present experiment, because it rules out early distortion (compressive nonlinearity in the luminance response of the display or the retina) as the basis for CM detection here. With such nonlinearity the distortion amplitude increases linearly with modulation depth, but as the *square* of carrier contrast (Scott-Samuel & Georgeson, 1999) and so doubling  $c$  should increase sensitivity by a factor of 4, i.e. 12 dB, not the 1.4 dB observed. We also ran a numerical simulation using our modulated bandpass-noise carriers, passed through a Naka-Rushton photoreceptor nonlinearity, that confirmed these arguments for the present experiment. We conclude that early distortion does not drive CM detection in our experiments.

The binocular advantage for both LM and CM (Fig. 4) was much the same for the 3 levels of interocular correlation of the carrier. This was confirmed by a selective 3-factor repeated-measures ANOVA testing *Ocularity* (Mon,Bin) x *Modulation type* (LM,CM) x *Carrier correlation* (1,0,-1) x *Subjects* (7), using the results of Fig. 4B,C excluding the Antiphase condition. Input data were detection thresholds in dB re 1% (not the relative thresholds of Fig. 4). Not surprisingly, the effect of *Ocularity* (binocular advantage) was very significant [ $F(1,6)=633, p<0.0001$ ], as was the effect of *Modulation type* (mean LM thresholds were 21.5

dB lower than CM) [ $F(1,6)=1195$ ,  $p<0.0001$ ], but the main effect of *Carrier correlation* was not significant [ $F(2,12)=0.936$ ,  $p=0.42$ ], nor was the interaction of *Carrier correlation* with *Modulation type* [ $F(2,12)=1.55$ ,  $p=0.25$ ].

First-order correlation (of the carriers) thus did not influence the first- or second-order binocular advantage for in-phase signals. This strongly suggests that binocular summation for CM does not rest on first-order binocular matching, and instead is likely to occur after recovery of the envelope information, when the carrier has been discarded (cf. Zhou, Georgeson, & Hess, 2014; Zhou, Liu, Zhou, & Hess, 2014). Our model (Fig. 2) is built on this foundation. We have inferred from Experiment 1 that binocular summation of CM does not require similar carriers in the two eyes. Experiment 2 therefore tested this idea in a more extreme case - where the carrier orientations were orthogonal between the eyes - a classic recipe for binocular rivalry.

FIG 5 HERE

### 3.2. Experiment 2. Detection of CM with oriented noise carriers

Procedure for Experiment 2 was similar to Experiment 1 in most respects, except that the dynamic carrier noise (centre frequency 6 c/deg) was filtered to be narrowband in orientation (Fig. 5A,B), with Gaussian orientation tuning (s.d. =  $15^{\circ}$ ). Only CM was tested, at r.m.s. carrier contrast  $c = 0.2$  in both eyes, with interocular correlation  $r = 0$ . Modulation was horizontal (0.75 c/deg; Mon, Bin or Antiphase) as before. Carrier orientation was either the same in both eyes ( $+45^{\circ}$  or  $-45^{\circ}$  from vertical) or orthogonal ( $\pm 45^{\circ}$ ) between the eyes (Fig. 5A,B). Each orientation was assigned equally often to the left and the right eye. Thresholds were derived from fitted psychometric functions (160 trials per condition per subject, as before). Only the three experienced subjects (MAG, AJS, SAW) were tested.

*Result:* The binocular advantage (Fig. 6A) was again close to 6dB (grey band) for parallel carriers (mean 5.5 dB) and for orthogonal carriers (mean 5.6 dB). This further supports our conclusion that the use of mis-matched, uncorrelated and potentially rivalrous carriers does not impede binocular combination of the envelope waveforms that they carry.

As in Experiment 1, there was rather little or no cancellation for Antiphase CM. The antiphase disadvantage was no more than about 2 dB, and there was a hint (not statistically significant) that this small disadvantage may occur only when the carriers have the same orientation (Fig. 6A). There was an analogous hint in Experiment 1 that the antiphase disadvantage was slightly greater with identical carriers ( $r = 1$ ) than dissimilar carriers ( $r = 0, -1$ ), but without more data we make no strong claim about this.

FIG 6 HERE

### 3.3. Experiment 3. Detection of CM with light and dark blob carriers

Procedure for Experiment 3 was almost the same as Experiment 2, except that the oriented carrier noise was replaced by random arrays of  $16 \times 16$  'blobs' with light (L) or dark (D) centres (Fig. 5C; see legend for details). Average centre frequency of the Fourier spectrum of these textures was 6 c/deg as before, and r.m.s. contrast  $c = 0.1$ . Blob positions were uncorrelated between the eyes, and the blobs had the *same* (L/L, D/D) or *opposite* polarity (L/D, D/L) across the (left/right) eyes. The 3 observers from Experiment 2 were tested again.

*Result:* The pattern of CM thresholds (Fig. 6B) was similar to Experiment 2 (Fig. 6A). Mean binocular advantage was 6.04 dB for textures of the same polarity and 5.24 dB for those of opposite polarity between the eyes. Mean antiphase threshold was only 0.3 dB higher than the

mean monocular threshold but, like experiment 2, there was a hint that an antiphase disadvantage occurred for similar carriers (same polarity) but not for dissimilar ones (opposite polarity).

We also note that there was a consistent effect of polarity *per se*. CM thresholds, averaged over the 3 observers and 3 test conditions [Mon, Bin, Anti], were significantly lower for dark blob carriers in both eyes (D/D) than for light blobs in both eyes (L/L) (mean difference = 2.9 dB; SEM = 0.7 dB,  $n=9$ ). This difference was evident for Mon, Bin and Antiphase test conditions (Fig. 6B). By comparison, there was a trivial difference between CM thresholds for carriers whose polarity differed between eyes in opposite ways (L/D versus D/L) (mean difference = 0.2 dB; SEM = 0.6 dB,  $n=9$ ). This advantage for seeing contrast variation in dark blobs may be closely related to the analogous advantage observed for dark targets (rather than light ones) in contrast discrimination tasks (Kingdom & Whittle, 1996; Legge & Kersten, 1983; McIlhagga & Peterson, 2006). Such asymmetries in response to contrast polarity probably have their origin in an early compressive response to luminance, perhaps driven by local light adaptation at the photoreceptor level. Such a nonlinearity gives rise to polarity asymmetries in a great variety of perceptual outcomes (Lu & Sperling, 2012), including perceived size and perceived location of edges (Georgeson & Freeman, 1997; Mather & Morgan, 1986), and reaction times (Komban, Alonso, & Zaidi, 2011). It is also reflected in neural responses at the retinal, lateral geniculate and cortical levels (Kremkow et al., 2014).

### 3.4. Experiment 4. Detection of CM: binocular, monocular and unocular

In this final experiment, we tested the idea that the ocular gains (weights) in binocular combination for CM are not fixed, but are controlled by the relative strengths (contrasts) of the carriers in the two eyes (Zhou, Georgeson & Hess, 2014). Eqn. 6 in our model implies that sensitivity to modulation in one eye should be higher when the carrier contrast is low or zero in the other eye, because the weight for the tested eye then switches from 0.5 to 1.0, while the other eye switches from 0.5 to 0. Thus we compared CM detection in the *monocular* condition (carrier in both eyes, modulated in one eye) with a new *unocular* condition (carrier and modulation only in one eye), and we predicted that unocular thresholds would be lower than monocular ones.

Procedure was similar to Experiment 1, using the 2-D isotropic noise carriers, but with CM only, and  $c = 0.2$ . The antiphase condition was replaced by the *unocular* condition, in which one eye saw a modulated carrier while the other saw a blank test image (zero-contrast at mean luminance). The textured border and fixation point were present in both eyes as usual. Two observers (MAG, AJS) were tested. Thresholds for Binocular, Monocular and Unocular modulation were derived from Weibull fits to data from individual sessions, done separately for test modulation in the left and right eyes (40 trials for each within-session threshold estimate; N=6 sessions for AJS, with carrier correlation  $r = 0$ ; N=18 sessions for MAG, 6 sessions each for  $r = 1, 0, -1$ ). This allowed us to compute descriptive statistics for each observer. Each mean threshold in Fig. 7 was thus derived from 240 trials (AJS) or 720 trials (MAG)

*Result:* Mean CM thresholds in dB  $\pm 1$  s.e. are plotted in Fig. 7A,B. As expected, binocular thresholds (Bin1, Bin2; tested twice in each session) were significantly lower than monocular ones. The binocular advantage was again close to 6dB: mean  $\pm$  s.e was 6.4 dB  $\pm$  0.38, N=18, for MAG, and 5.3 dB  $\pm$  0.57, N=6, for AJS.

Our key finding in Experiment 4 was that *unocular* thresholds for the left and right eyes (UniL, UniR) were significantly lower than the corresponding *monocular* thresholds (MonL, MonR) by an average of 2.6 dB for MAG ( $t=5.20$ ,  $df=35$ ,  $p<0.00001$ ) and 2.2 dB for AJS ( $t=4.04$ ,  $df=11$ ,  $p<0.001$ ). [This difference co-existed with small differences in sensitivity between the eyes. CM thresholds were an average of 3dB higher in the right eye for MAG, but 1.8dB lower in the right

eye for AJS. For MAG we confirmed that this was accompanied by lower detectability of the carrier in the right eye: 2AFC contrast thresholds for detecting the carrier alone were 2dB higher in the right eye than the left. In modelling the results (red circles in Fig. 7, discussed later) we take these small ocular asymmetries into account. ]

FIG 7 HERE

## 4. Discussion

In four experiments we tested for the presence of binocular summation of contrast modulation (CM) at detection threshold, and found it to be very robust. Using carriers of the same contrast in both eyes, we found a consistently high level of binocular advantage (around 5.5 to 6 dB), meaning that binocular thresholds were 1.88 to 2 times better (lower) than monocular thresholds. In experiments 1-3 similarity or dissimilarity in the spatial structure of the carriers (in terms of interocular correlation, orientation, or contrast polarity) had little or no effect on the binocular advantage. Orthogonal orientations (in experiment 2) and opposite contrast polarities (in experiments 1 and 3) are potentially strongly rivalrous, but this did not prevent binocular combination of the contrast envelopes. This could mean that envelope information is extracted at a low level of processing, before the stage at which rivalry suppression is expressed. But it is also consistent with evidence that after stimulus onset there is a brief (or sometimes prolonged) period of 'false fusion' where potentially rivalrous inputs are both represented, before dominance of one and suppression of the other are established (Liu, Tyler, & Schor, 1992; Song & Yao, 2009; Wolfe, 1983). Either way, it is clear that matching and fusion of the carriers are not necessary for full binocular summation of the envelope signals. Whether this remains true at shorter or longer durations than the one we used, or for static rather than dynamic carriers, is an interesting question for future work.

The near-doubling of sensitivity with modulation in both eyes strongly suggests simple binocular summation: linear summation of left and right eye responses that are themselves proportional to contrast modulation depth. But if the summation were truly linear - preserving the sign of the envelope signal - then we should expect almost complete cancellation between out-of-phase modulations. Antiphase thresholds should be unmeasurably high, but they were not. They were at most about 2dB higher than monocular thresholds. These results for CM were very similar to those with luminance-modulated noise (LM, Experiment 1). An explanation is needed that captures both the high degree of summation, and the lack of cancellation.

### *A model for binocular detection of contrast modulation*

We propose a 3-channel model (Fig. 2; eqns. 4-9) in which a binocular channel computes a weighted linear sum of the second-order (contrast envelope) signals from the two eyes, but access to the two monocular responses is maintained in parallel. The weight or gain assigned to each eye depends on the carrier contrast in that eye relative to the other eye. Response selection across the 3 noisy channels is achieved by winner-take-all (a *max* operator). It follows that the monocular channels will mediate detection at large phase disparities (around  $180^\circ$ ), where the binocular channel response is nulled. We now show how well this model can account for our data, and discuss the influence of noise and background response levels on the behaviour of the *max* operator, and on the predicted thresholds.

FIG 8 HERE

Fig. 8 compares the group mean CM thresholds from experiment 1 with the best-fitting model thresholds. The horizontal axis plots phase disparity in deg, to emphasize that the binocular and antiphase conditions are just two key points ( $0^\circ$ ,  $180^\circ$ ) on the continuum of phase disparities. Dashed horizontal line marks the model's monocular threshold, while open circles show how the model's threshold varied with phase disparity. The least-squares fit with 2 free parameters ( $r_0$ ,  $\sigma$ )

was excellent (Fig. 8A,B). Antiphase signals cancel each other in the binocular channel, and so a model that used only the binocular channel would fail badly (Fig. 8A, dash-dot curve). In our 3-channel model, the monocular channels mediate detection in the antiphase condition (at  $180^\circ$ , blue diamond) and, as Fig. 8A shows, are expected to do so for disparities  $90 - 270^\circ$ . Thus the 3-channel architecture of this model yields both the required summation and lack of substantial cancellation. Fig. 9C shows, in summary, how the fitted model describes the observed thresholds (horizontal axis) and how well it predicted the average psychometric slopes ( $\beta$ ), which played no part in the fitting procedure.

#### *Possible role of background response, $r_0$*

The model predictions (Fig. 8) are fine-tuned by the background response  $r_0$ . With no background response the model fit was poorer (r.m.s. error = 1.7 dB; not shown). Binocular advantage was under-estimated (4.6dB instead of 6dB), and the predicted antiphase threshold was 1dB better than monocular, rather than about 2dB worse. Our results are therefore not consistent with  $r_0 = 0$ . Fig. 9A illustrates more broadly how model thresholds varied with increasing  $r_0$  for a fixed noise level  $\sigma$ . Binocular thresholds were least affected, showing a slight decrease (improvement in sensitivity); monocular thresholds increased by about 2dB, then levelled off; antiphase thresholds increased without limit as  $r_0$  increased. Thus the predicted binocular advantage quickly asymptoted at 6dB for  $r_0 > 0.5\sigma$ , but the antiphase disadvantage was very sensitive to the relative strength of the background response ( $r_0/\sigma$ ). For  $r_0 \leq 0$ , antiphase thresholds fell *below* the monocular ones, giving a small antiphase advantage. We therefore need to understand the apparent paradox that a non-informative background response, present in both intervals, can influence signal detectability.

FIG 9 HERE

We looked carefully at how the background response controls the Mon and Bin thresholds, and hence the predicted level of binocular advantage. We describe the main influences in the Appendix. In brief, increasing  $r_0$  in our model pushes up  $r_B$  more than  $r_L$  or  $r_R$  and so favours the contribution of the B channel in both the Mon and Bin conditions. With binocular summation in the B channel this promotes the full factor-of-two (6dB) binocular advantage. Conversely, antiphase signals cancel in the B channel, and so reliable detection requires the monocular channels to deliver the *max* output. Favouring the B channel by increasing  $r_0$  prevents them from doing this, and so entails an increasing antiphase disadvantage that has no upper limit (Fig. 9A, blue diamonds). Colour-filled symbols and vertical dashed line in Fig. 9A show that the results of Experiment 1 are consistent with a small, positive, background response that supports the full binocular advantage yet creates only a small antiphase disadvantage. The model fit to these group data implies that the background response level is about half the internal noise level ( $r_0 = 0.6\sigma$ ). When there is no background response, the monocular channels contribute more to monocular detection, and this lowers the monocular threshold and reduces the predicted binocular advantage to well below 6dB.

The data of experiments 1-3 gave hints that the antiphase disadvantage increased with the interocular similarity of the carriers, and differed reliably between individuals. These (relatively weak) effects in the data (Appendix, Sec. 2) could mean that the background response  $r_0$  is higher for some observers than others, and higher when the carriers are correlated or at least similar (same orientation, or same polarity) than when they are not. See Appendix Sec. 2 for details.

#### *Contrast-weighted summation*

We found that monocular CM thresholds (Fig. 7) were significantly higher than unocular thresholds for the same eye by an average of 2.4dB (32%). This is striking, because the two conditions had the same test images in one eye and differed only in the presence or absence of a uniform (hence uninformative) carrier contrast in the other eye. A form of interocular

suppression is at work here, and our model embodies that interaction in the weights for each eye (eqn. 6). The weight for one eye goes up when contrast in the other eye goes down (Baker, Meese, & Georgeson, 2007; Ding & Sperling, 2007; Meese et al., 2006). Uniocular thresholds are lower than monocular because without a carrier in the other eye the weight for the tested eye is higher.

To test this interpretation we fitted the model to data of experiment 4, while allowing for the small between-eye differences in sensitivity to the carrier. The latter was done simply by adjusting the model carrier contrast  $c_R$ , decreasing it by 2dB for MAG and increasing it by 1dB for AJS. This influenced both the weights (via eqn. 6) and the signal amplitude  $a_R$  (because  $a_R = m.c_R$ , where  $m$  is modulation depth). With this minor elaboration, the model thresholds are shown as red circles in Fig. 7, and they captured the experimental data very accurately for both observers (see figure legend).

To prove that this success depended on the contrast-dependent weighting scheme (eqn. 6), we ran the model with the same parameters and contrast asymmetry as before but with fixed weights  $w_L = w_R = 0.5$ . The resulting fit was notably worse than in Fig. 7, and rms errors increased to 1.56 dB (MAG) and 1.36dB (AJS). The predicted *uniocular* thresholds were 2-3dB too high, reflecting the incorrect under-weighting of the tested eye. Re-fitting the parameter values ( $r_0, \sigma$ ) improved the fit a little, but the residual errors were of a similar kind. In particular, the model with fixed weights predicts the *same* thresholds for uniocular and monocular, and so cannot predict our finding that uniocular thresholds were significantly lower than monocular. The close fit seen in Fig. 7 is strong evidence for contrast-weighted averaging of CM signals rather than simple averaging with fixed weights.

#### *Suprathreshold perception of binocular CM*

In a recent study, Zhou, Georgeson & Hess (2014) showed that the perception of suprathreshold contrast modulation also exhibits binocular summation and, as here, they found that summation was the same for correlated and uncorrelated carriers. In a perceptual matching experiment, binocular test modulation of 0.8 was matched by *uniocular* comparison modulation of 0.8 (a veridical match), but a monocular test of 0.8 appeared halved: it was matched by a uniocular comparison close to 0.4. This halving suggests averaging of test modulations across the two eyes, and indeed dichoptic tests intermediate between these two cases all fell onto the curve predicted by linear averaging. But to explain both this averaging behaviour *and* the veridical matching between binocular and uniocular CM, Zhou *et al* proposed the contrast-weighting scheme that we have used here (eqn. 6). This contrast-dependent mechanism of combination implements averaging (weights of 0.5,0.5) when the carriers are in both eyes, but winner-take-all (weights of 1,0) when the carrier is only in one eye. The present model incorporates this interactive form of CM combination and elaborates the model for the forced-choice detection task.

Zhou *et al* (2014) found that perceived modulation depth fell with increasing CM phase disparity from 0 to 90<sup>0</sup>, and this too was well predicted by linear averaging of the two disparate sinewaves. The binocular summing channel alone was sufficient to explain all their matching results, whereas we have used the same binocular mechanism along with parallel, monocular channels and selection by the *max* operator in the detection task. This is a major difference, but is not contradictory: our model is a generalization of Zhou *et al*'s model. We tested phase disparities of 0 and 180<sup>0</sup>, and found that monocular channels in the model play a key role in preventing cancellation at large phase disparities (90-270<sup>0</sup>), but no role for disparities of 0-90<sup>0</sup> (Fig. 8A), the range tested by Zhou *et al*. Thus Zhou *et al* did not need to include monocular CM channels to explain their data, but we expect that the monocular channels will be needed to explain CM matching in the disparity range 90-180<sup>0</sup>, as follows.

### *A prediction*

Our model for CM predicts that above threshold both monocular and antiphase modulations (e.g.  $m=0.5$ ) will appear to have about half the modulation of a binocular CM test image (Fig. 9B). This is already known to be correct for monocular CM (Zhou *et al*, 2014) but has not been tested for antiphase. It's an interesting prediction, because for luminance gratings it certainly does *not* hold. At contrasts above about 20%, monocular, binocular and antiphase luminance gratings *all* appear to have the same contrast (Baker, Meese, & Georgeson, 2007; Baker, Wallis, Georgeson, & Meese, 2012; Huang, Zhou, Zhou, & Lu, 2010). The difference in prediction arises very simply, because monocular channels in our CM model have a contrast-weighted gain of 0.5 (when the carrier is in both eyes; eqn. 6), but in the first-order model (Zhou *et al*, 2014) the monocular channels are not influenced by the other eye, and always have a weight of 1.0. The idea of computing the *max* over monocular and binocular channel responses has also been successfully applied to explain dichoptic (first-order) contrast discrimination and contrast matching (Georgeson, Wallis, Meese, & Baker, 2016).

### *Relation to binocular summation of luminance contrast*

The literature on binocular summation for luminous targets and luminance contrast is large, with a long history (reviewed by Blake, Sloane, & Fox, 1981; Blake & Fox, 1973; Howard & Rogers, 1995). There is a great variety of models for binocular summation, but with a consensus that the observed binocular advantage is (a) greater than one would expect from independent detection by the two eyes and (b) implies some form of central summation of signals from each eye. There is, surprisingly, no equally clear consensus on just how much improvement occurs. Expressing binocular advantage as the Mon/Bin contrast threshold ratio, in dB, the estimates in the literature vary from about 3 dB to 6 dB (ratios from  $\sqrt{2}$  to 2). This lack of consensus probably springs, in part, from the smallness of these effects - just a few dB - coupled with the considerable difficulty of getting reliable, unbiased threshold estimates to the required accuracy. Over the last 30 years, computer-controlled, forced-choice psychophysical methods with better display technology, and much greater numbers of trials, have improved the reliability of threshold estimation, and with these methods estimates of the binocular advantage tend to cluster around 4-5dB (Legge, 1984; Maehara & Goryo, 2005; Meese & Baker, 2011; Meese *et al.*, 2006; Meese & Summers, 2009; Simmons & Kingdom, 1998; Simmons, 2005) - as if forming a compromise between the theoretically salient values of 3 and 6 dB (which represent an ideal observer with independent noise in the two eyes *versus* linear summation with fixed noise). In a retrospective analysis, however, we found that the variation between studies may not be random. Instead there appears to be an unsuspected systematic relation between the binocular advantage and the antiphase threshold, discussed next.

FIG 10 HERE

Sherrington (1904) was a pioneer of studies on binocular summation, and his work on binocular flicker included tests of in-phase and antiphase summation. But in a search for published studies that had (i) used sinewave gratings, (ii) compared in-phase with antiphase, (iii) used reliable forced-choice methods and (iv) gave data that could be expressed as a threshold ratio, we found just two (Legge, 1984; Simmons, 2005). Their results and ours are re-plotted in Fig. 10. The data points come from a variety of conditions (luminance, chromatic, LM and CM gratings; see legend for details), but appear to reveal a common trend: higher binocular advantage is associated with higher antiphase thresholds. The grey curve (Fig. 10) shows that this trend is what we should expect from our 3-channel model if the background response  $r_0$  varied across studies and across stimulus conditions. Moving up the curve represents an increase in  $r_0$  which (as we saw above) increasingly favours the binocular channel over the monocular ones. This bias exposes the full factor-of-two summation in the binocular channel, but at the expense of reduced detectability for the antiphase stimulus which is nulled by that channel. We should be suitably cautious about correlation evidence across different studies, and we do not yet know what might cause  $r_0$  to vary between stimulus conditions or between observers. But the 3-

channel model implies that there should be a constraint on the relation between the two threshold measures (grey curve in Fig. 10). They are not free to vary independently, and that constraint is evident in the pooled experimental data. In short, we argue that binocular advantage and antiphase threshold are correlated across studies because the relative contribution made by monocular and binocular channels varies, and this affects both measures in a predictable way - they both go up or down together. In our model that is driven by variation in  $r_0$ , but the same idea could probably be implemented in other ways, e.g. variation in the number or relative sensitivity of monocular and binocular neurons that contribute to detection (Anderson & Movshon, 1989).

### *Cortical physiology*

Five decades of research on the visual cortex of cats and monkeys (for review see Cumming & DeAngelis, 2001) has established very clearly that the neural basis for binocular combination and stereo disparity coding lies in the convergence of inputs from both eyes onto binocular cortical cells whose left- and right-eye receptive fields are closely matched in size, orientation, spatial frequency (SF) tuning and direction selectivity, but which vary in phase disparity and/or spatial disparity (eg. Ohzawa, DeAngelis, & Freeman, 1996). Complex cells inherit their disparity tuning from simple cells, or simple-like pairs of sub-units, but by pooling across different input pairs they generalize across contrast polarity and across a wider range of spatial locations (Ohzawa, DeAngelis, & Freeman, 1997; Tanabe & Cumming, 2008). Preferred disparity and the degree of disparity selectivity vary greatly between cells but, surprisingly, both these aspects of disparity tuning appear to be independent of ocular dominance (the extent to which a cell responds more to one eye than the other when tested monocularly) (Kara & Boyd, 2009; LeVay & Voigt, 1988).

Cortical cells can be responsive to a variety of second-order patterns that would not be detected directly by first-order linear filters (for review see Baker & Mareschal, 2001). For CM, Zhou & Baker (1993) found that about half the cells in area 18 of the cat were selectively responsive to the low SF envelopes of CM gratings as well as to low SF luminance gratings. Such cells tended to exhibit 'cue invariance', having preferred orientations and SFs that were similar for luminance gratings and CM gratings, while the preferred carrier SF was usually an order of magnitude higher, both in cat area 18 (Zhou & Baker, 1993) and monkey V2 (Li et al., 2014). These neural responses to CM are consistent with processing by a two-stage FRF mechanism (cf. Sec 2.2). Recently, however, surround suppression has been identified as a second mechanism in both cats and monkeys through which both V1 and V2 cells may exhibit selectivity for second-order spatial structure (Hallum & Movshon, 2014; Tanaka & Ohzawa, 2009). The carrier and envelope response characteristics are very different for this type of CM response, and so it appears that FRF and surround suppression are two distinct routes by which CM information is encoded in the visual cortex (Hallum & Movshon, 2014; Tanaka & Ohzawa, 2009).

Only one physiological study, however, has examined whether cortical cells responding to CM show binocular summation, and whether they are selective for second-order (CM) disparity. Tanaka & Ohzawa (2006) showed that many cells in cat area 18 (36% of their sample) were significantly sensitive to the interocular phase disparity of the CM envelope, showing a roughly sinusoidal modulation of response with phase disparity. The preferred disparity for CM was usually similar to that for luminance gratings of the same SF. Rather like our results with anti-correlated carriers, the response to binocular CM was unaffected by phase disparity of the carriers. Tanaka & Ohzawa (2006) showed that disparity tuning for CM envelopes in FRF models could arise from binocular convergence at the first-stage (carrier-tuned filters) or the second-stage (enveloped-tuned filters) (their Figs. 5A, 5B). But they argued that the second-stage convergence model was much more plausible, because to respond to large envelope disparities, the first-stage model would require implausibly large disparities between the relatively small receptive fields that are driven by the carrier. In sum, our proposed model for



## Appendix

### 1. How can a steady background response $r_0$ affect signal detectability ?

If the background response were the same for L, B and R channels, it would simply add a constant to all three responses in both the signal and non-signal intervals, and therefore have no influence at all on the behaviour of the *max* operator. All predicted thresholds would remain unchanged with or without a background response, and this expectation was confirmed by running such a modified model. But in our model architecture (Fig. 2) binocular summation occurs for the input signals, the internal noise, *and* the background responses. In consequence the influence of the background response on signal detection via the *max* operator is fairly complex.

FIG A1 HERE

We examined the way the mean and variance of  $r_L$ ,  $r_R$ ,  $r_B$  and *Resp* varied with modulation depth, and with  $r_0$ , and looked at the contribution made by each of the 3 channels to the *Resp* - that is, the proportion of trials on which the L, R, or B channels provided the *max* response. In all cases (Fig. A1) binocular summation of  $r_0$  shifts the mean response of the B channel upwards by twice as much as the L or R channels (compare top and bottom rows in Fig. A1). When this elevation of the B channel background is large enough (Fig. A1,D,E), the B channel comes to dominate the output of the *max* operator for both Monoc *and* Binoc modulation. In effect, the L,R channel outputs are silenced, while the B channel sums L and R inputs and so yields the full (6dB) binocular advantage that is observed. The background biases the system towards use of the B channel.

On the other hand, if the background response is small or zero then this bias is low or absent, and the L,R channels can play a larger role in detection of Monoc modulation (Fig. A1,A). The variance of the monocular channels is lower [not shown], because they do not sum the noise from both eyes, and this leads to lower Monoc thresholds when the background response is lower (Fig. 9A, green triangles). But for Binocular detection (Fig. A1,B) binocular summation favours the B channel even when there is no background response and so the L,R channels again have little influence, and Binoc thresholds are little affected by  $r_0$  in the range  $r_0 \geq 0$  (Fig. 9A, red circles).

Antiphase thresholds, however, increase dramatically with increase in  $r_0$ . This occurs because antiphase signals cancel in the B channel (Fig. A1,C,F). As we saw above, increasing  $r_0$  favours the B channel, but now the B channel carries no information: its mean is the same in the signal and non-signal intervals. Modulation depth therefore has to be increased until the L,R channels overcome the bias against them, to signal a reliable difference between intervals (e.g. for modulation depths  $> 0.2$  in Fig. A1,F). Antiphase thresholds rise markedly with increases in  $r_0$ . In the *max* operator, the larger but uninformative B response increasingly occludes the smaller, informative L,R responses (Fig. A1,F), reducing  $d'$  and raising thresholds.

### 2. Does the background response $r_0$ differ between conditions & between observers ?

#### *Dependence of $r_0$ on interocular similarity of the carriers ?*

We know little about the source of the proposed background response, except that in our model it must arise before binocular summation (Fig. 2). We treated  $r_0$  as an additive constant, but here we discuss some hints in the data that  $r_0$  might depend on the interocular similarity of the carriers. Fig. 9A showed that the model antiphase threshold increased markedly with  $r_0$ , implying that observed variations in antiphase thresholds, between conditions or between individuals, could reflect differences in  $r_0$ . Thus in experiment 1 we saw hints that the experimental antiphase disadvantage increased with interocular correlation (Fig. 4C,D) -

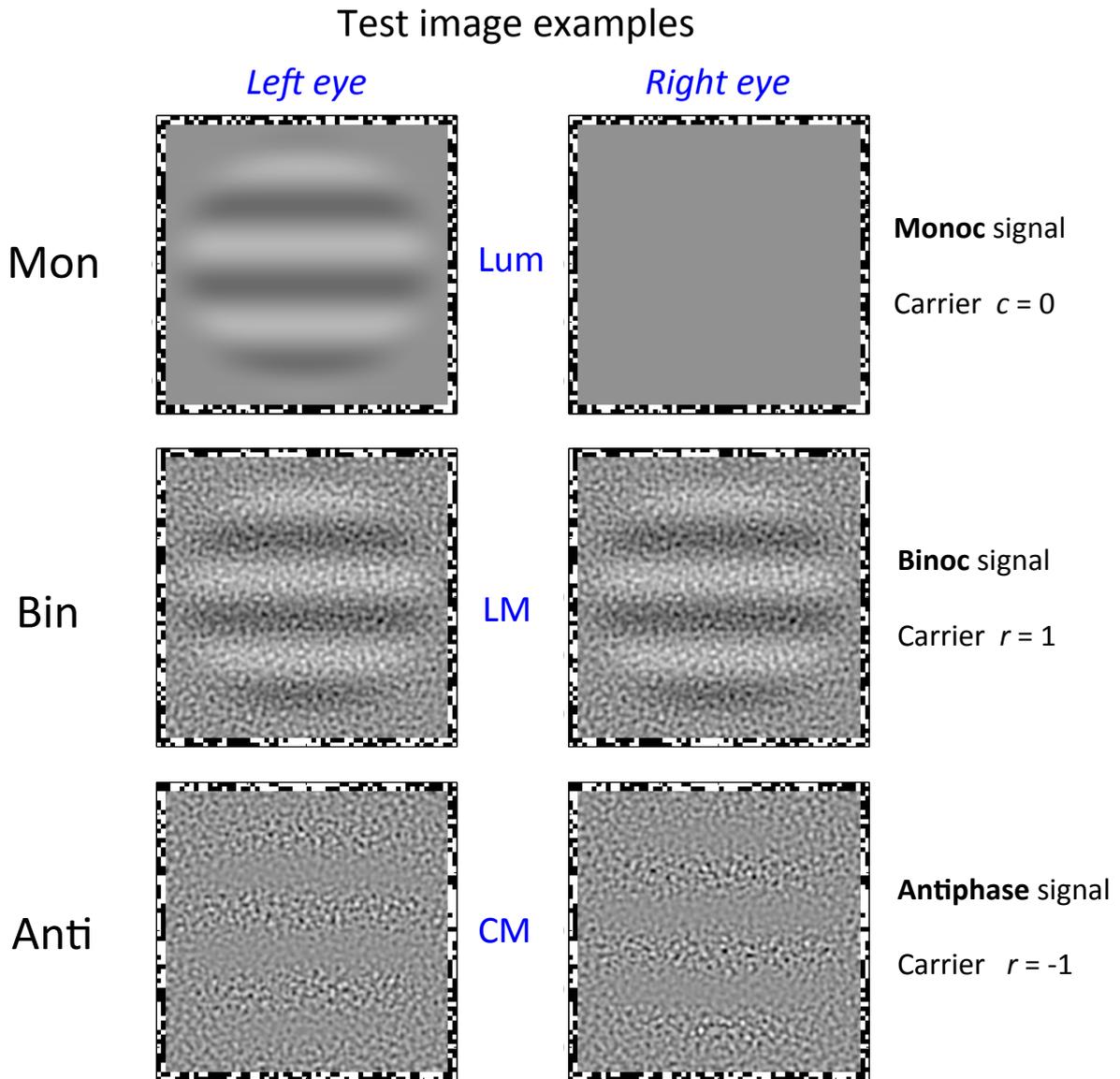


## References

- Anderson, P. A., & Movshon, J. A. (1989). Binocular combination of contrast signals. *Vision Research*, *29*, 1001–1016. [http://doi.org/10.1016/0042-6989\(89\)90060-6](http://doi.org/10.1016/0042-6989(89)90060-6)
- Anzai, A., Bearnse, M. A., Freeman, R. D., & Cai, D. (1995). Contrast coding by cells in the cat's striate cortex: monocular vs. binocular detection. *Visual Neuroscience*, *12*, 77–93. <http://doi.org/10.1017/S0952523800007331>
- Baker, C. L., & Mareschal, I. (2001). Processing of second-order stimuli in the visual cortex. *Progress in Brain Research*, *134*, 171–191. [http://doi.org/10.1016/S0079-6123\(01\)34013-X](http://doi.org/10.1016/S0079-6123(01)34013-X)
- Baker, D. H., Meese, T. S., & Georgeson, M. A. (2007). Binocular interaction: contrast matching and contrast discrimination are predicted by the same model. *Spatial Vision*, *20*(5), 397–413. <http://doi.org/10.1163/156856807781503622>
- Baker, D. H., Wallis, S. A., Georgeson, M. A., & Meese, T. S. (2012). The effect of interocular phase difference on perceived contrast. *PLoS ONE*, *7*(4), 1–6. <http://doi.org/10.1371/journal.pone.0034696>
- Blake, R., & Fox, R. (1973). The psychophysical inquiry into binocular summation. *Perception & Psychophysics*, *14*(1), 161–185.
- Blake, R., Sloane, M., & Fox, R. (1981). Further developments in binocular summation. *Perception & Psychophysics*, *30*(3), 266–76.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.
- Campbell, F. W., & Green, D. G. (1965). Monocular versus binocular visual acuity. *Nature*, *208*, 191–192.
- Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision*, *4*(2/3), 103–129. <http://doi.org/10.1007/s13398-014-0173-2>
- Chubb, C., & Sperling, G. (1989). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(8), 2985–2989. <http://doi.org/10.1073/pnas.86.8.2985>
- Clifford, C. W. G., & Vaina, L. M. (1999). A computational model of selective deficits in first and second-order motion processing. *Vision Research*, *39*(1), 113–130. [http://doi.org/10.1016/S0042-6989\(98\)00082-0](http://doi.org/10.1016/S0042-6989(98)00082-0)
- Cumming, B. G., & DeAngelis, G. C. (2001). The physiology of stereopsis. *Annual Review of Neuroscience*, *2*, 203–238.
- Dakin, S. C., & Mareschal, I. (2000). Sensitivity to contrast modulation depends on carrier spatial frequency and orientation. *Vision Research*, *40*(3), 311–29.
- Ding, J., & Sperling, G. (2007). Binocular combination: measurements and a model. In L. R. Harris & M. R. M. Jenkin (Eds.), *Computational vision in neural and machine systems* (pp. 257–305). Cambridge, UK: Cambridge University Press.
- Edwards, M., Pope, D. R., & Schor, C. M. (2000). First- and second-order processing in transient stereopsis. *Vision Research*, *40*(19), 2645–2651.
- Georgeson, M. A., & Freeman, T. C. (1997). Perceived location of bars and edges in one-dimensional images: computational models and human vision. *Vision Research*, *37*(1), 127–142.
- Georgeson, M. A., & Schofield, A. J. (2002). Shading and texture: separate information channels with a common adaptation mechanism? *Spatial Vision*, *16*(1), 59–76. <http://doi.org/10.1163/15685680260433913>
- Georgeson, M. A., Wallis, S. A., Meese, T. S., & Baker, D. H. (2016). Contrast and Lustre: a model that accounts for eleven different forms of contrast discrimination in binocular vision. *Vision Research*, (in press).
- Hallum, L. E., & Movshon, J. A. (2014). Surround suppression supports second-order feature encoding by macaque V1 and V2 neurons. *Vision Research*, *104*, 24–35. <http://doi.org/10.1016/j.visres.2014.10.004>
- Hess, R. F., & Wilcox, L. M. (2008). The transient nature of 2nd-order stereopsis. *Vision Research*, *48*(11), 1327–1334. <http://doi.org/10.1016/j.visres.2008.02.008>
- Howard, I. P., & Rogers, B. J. (1995). *Binocular vision and stereopsis*. Oxford: Oxford University Press.
- Huang, C.-B., Zhou, J., Zhou, Y., & Lu, Z.-L. (2010). Contrast and phase combination in binocular vision. *PLoS One*, *5*(12), e15075. <http://doi.org/10.1371/journal.pone.0015075>
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, *195*, 215–243.
- Kara, P., & Boyd, J. D. (2009). A micro-architecture for binocular disparity and ocular dominance in visual cortex. *Nature*, *458*(7238), 627–631. <http://doi.org/10.1167/7.15.21>
- Kingdom, F. A. A., & Whittle, P. (1996). Contrast discrimination at high contrasts reveals the influence of local light adaptation on contrast processing. *Vision Research*, *36*(6), 817–829. [http://doi.org/10.1016/0042-6989\(95\)00164-6](http://doi.org/10.1016/0042-6989(95)00164-6)
- Komban, S. J., Alonso, J.-M., & Zaidi, Q. (2011). Darks Are Processed Faster Than Lights. *Journal of Neuroscience*, *31*(23), 8654–8658. <http://doi.org/10.1523/JNEUROSCI.0504-11.2011>
- Kremkow, J., Jin, J., Komban, S. J., Wang, Y., Lashgari, R., Li, X., ... Alonso, J.-M. (2014). Neuronal nonlinearity explains greater visual spatial resolution for darks than lights. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(8), 3170–5. <http://doi.org/10.1073/pnas.1310442111>
- Landy, M. S., & Bergen, J. R. (1991). Texture segregation and orientation gradient. *Vision Research*, *31*, 679–691.
- Langley, K., Fleet, D. J., & Hibbard, P. B. (1999). Stereopsis from contrast envelopes. *Vision Research*, *39*(14), 2313–24.

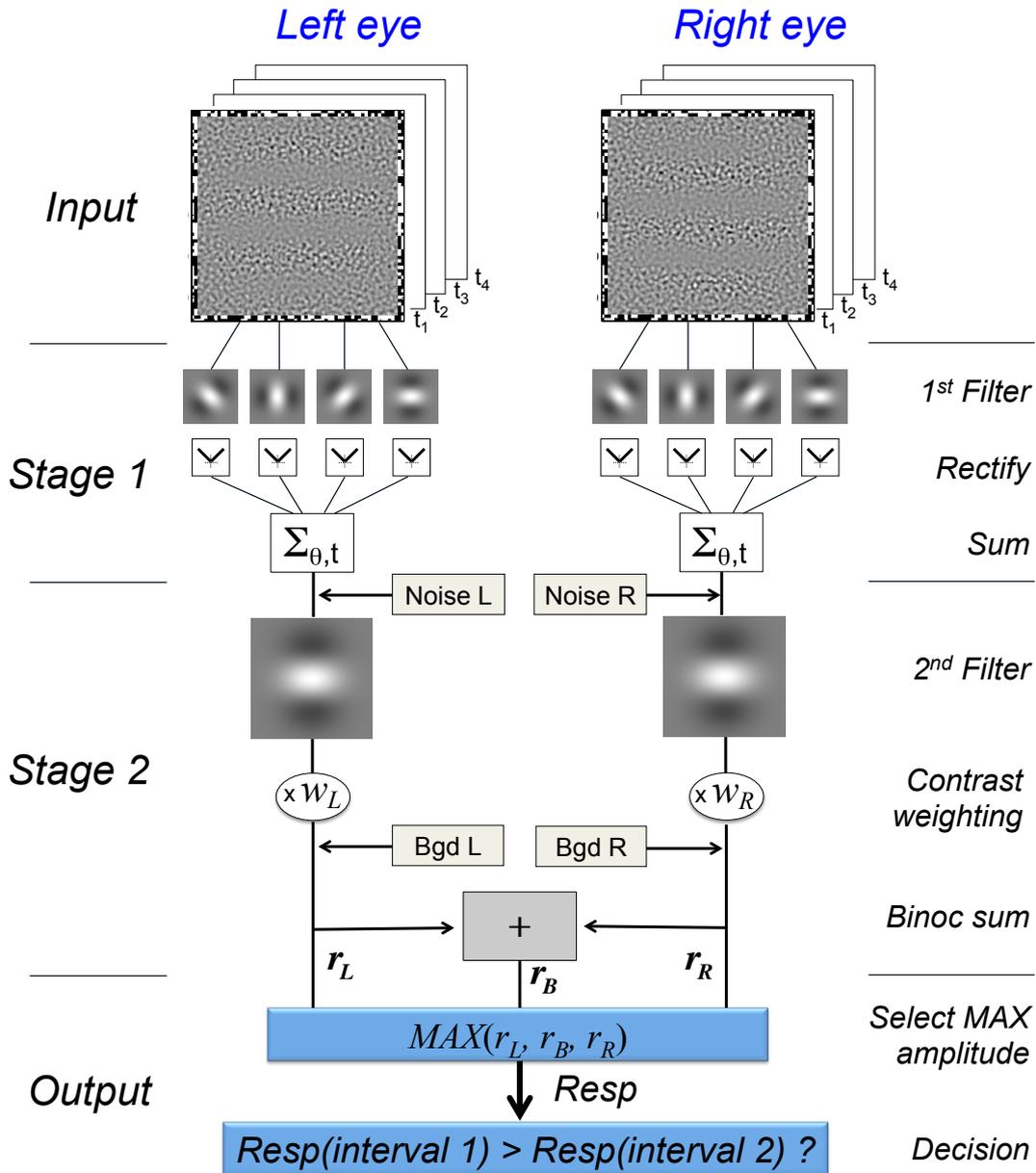
- Legge, G. E. (1984). Binocular contrast summation--I. Detection and discrimination. *Vision Research*, 24(4), 373–383.
- Legge, G. E., & Kersten, D. (1983). Light and dark bars; contrast discrimination. *Vision Research*, 23(5), 473–483. [http://doi.org/10.1016/0042-6989\(83\)90122-0](http://doi.org/10.1016/0042-6989(83)90122-0)
- LeVay, S., & Voigt, T. (1988). Ocular dominance and disparity coding in cat visual cortex. *Visual Neuroscience*, 1(4), 395–414. <http://doi.org/10.1017/S0952523800004168>
- Li, G., Yao, Z., Wang, Z., Yuan, N., Talebi, V., Tan, J., ... Baker, C. L. (2014). Form-cue invariant second-order neuronal responses to contrast modulation in primate area V2. *Journal of Neuroscience*, 34(36), 12081–92. <http://doi.org/10.1523/JNEUROSCI.0211-14.2014>
- Liu, L., Tyler, C. W., & Schor, C. M. (1992). Failure of rivalry at low contrast: evidence of a suprathreshold binocular summation process. *Vision Research*.
- Lu, Z.-L., & Sperling, G. (2012). Black-white asymmetry in visual perception. *Journal of Vision*, 12(10):8(2012), 1–21. <http://doi.org/10.1167/12.10.8.Introduction>
- Lu, Z. L., & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research*, 35(19), 2697–2722. [http://doi.org/10.1016/0042-6989\(95\)00025-U](http://doi.org/10.1016/0042-6989(95)00025-U)
- Maehara, G., & Goryo, K. (2005). Binocular, monocular and dichoptic pattern masking. *Optical Review*, 12(2), 76–82. <http://doi.org/10.1007/s10043-004-0076-5>
- Manahilov, V., Calvert, J., & Simpson, W. A. (2003). Temporal properties of the visual responses to luminance and contrast modulated noise. *Vision Research*, 43, 1855–1867. [http://doi.org/10.1016/S0042-6989\(03\)00275-X](http://doi.org/10.1016/S0042-6989(03)00275-X)
- Mareschal, I., & Baker, C. L. (1999). Cortical processing of second-order motion, 527–540.
- Mather, G., & Morgan, M. J. (1986). Irradiation: implications for theories of edge localization. *Vision Research*, 26(6), 1007–1015.
- McIlhagga, W., & Peterson, R. (2006). Sinusoid = light bar + dark bar? *Vision Research*, 46(12), 1934–1945. <http://doi.org/10.1016/j.visres.2005.12.004>
- Meese, T. S., & Baker, D. H. (2011). Contrast summation across eyes and space is revealed along the entire dipper function by a “Swiss cheese” stimulus. *Journal of Vision*, 11(1):23, 1–23. <http://doi.org/10.1167/11.1.23>
- Meese, T. S., Georgeson, M. A., & Baker, D. H. (2006). Binocular contrast vision at and above threshold. *Journal of Vision*, 6(11), 1224–43. <http://doi.org/10.1167/6.11.7>
- Meese, T. S., & Summers, R. J. (2009). Neuronal convergence in early contrast vision: binocular summation is followed by response nonlinearity and area summation. *Journal of Vision*, 9, 7.1-16. <http://doi.org/10.1167/9.4.7>
- Motoyoshi, I., & Kingdom, F. A. A. (2007). Differential roles of contrast polarity reveal two streams of second-order visual processing. *Vision Research*, 47(15), 2047–2054. <http://doi.org/10.1016/j.visres.2007.03.015>
- Motoyoshi, I., & Nishida, S. (2004). Cross-orientation summation in texture segregation. *Vision Research*, 44, 2567–2576. <http://doi.org/10.1016/j.visres.2004.05.024>
- Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1996). Encoding of binocular disparity by simple cells in the cat’s visual cortex. *Journal of Neurophysiology*, 75(5), 1779–1805.
- Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1997). Encoding of Binocular Disparity by Complex Cells in the Cat’s Visual Cortex. *Journal of Neurophysiology*, 77, 2879–2909.
- Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America. A, Optics and Image Science*, 2(9), 1508–32.
- Schofield, A. J. (2000). What does second-order vision see in an image? *Perception*, 29(9), 1071–86. <http://doi.org/10.1068/p2913>
- Schofield, A. J., & Georgeson, M. A. (1999). Sensitivity to modulations of luminance and contrast in visual white noise: Separate mechanisms with similar behaviour. *Vision Research*, 39(16), 2697–2716. [http://doi.org/10.1016/S0042-6989\(98\)00284-3](http://doi.org/10.1016/S0042-6989(98)00284-3)
- Schofield, A. J., & Georgeson, M. A. (2000). The temporal properties of first- and second-order vision. *Vision Research*, 40(18), 2475–87. [http://doi.org/10.1016/S0042-6989\(00\)00111-5](http://doi.org/10.1016/S0042-6989(00)00111-5)
- Scott-Samuel, N. E., & Georgeson, M. A. (1999). Does early non-linearity account for second-order motion? *Vision Research*, 39(17), 2853–65. [http://doi.org/10.1016/S0042-6989\(98\)00316-2](http://doi.org/10.1016/S0042-6989(98)00316-2)
- Sherrington, C. S. (1904). On binocular flicker and the correlation of activity of “corresponding” retinal points. *British Journal of Psychology*, 1(1), 26–60.
- Simmons, D. R. (2005). The binocular combination of chromatic contrast. *Perception*, 34, 1035–1042. <http://doi.org/10.1068/p5279>
- Simmons, D. R., & Kingdom, F. A. A. (1998). On the binocular summation of chromatic contrast. *Vision Research*, 38(8), 1063–1071. [http://doi.org/10.1016/S0042-6989\(97\)00272-1](http://doi.org/10.1016/S0042-6989(97)00272-1)
- Solomon, J. A., & Sperling, G. (1994). Full-wave and half-wave rectification in second-order motion perception. *Vision Research*, 34(17), 2239–2257. [http://doi.org/10.1016/0042-6989\(94\)90105-8](http://doi.org/10.1016/0042-6989(94)90105-8)
- Song, C., & Yao, H. (2009). Duality in binocular rivalry: distinct sensitivity of percept sequence and percept duration to imbalance between monocular stimuli. *PloS One*, 4(9), e6912. <http://doi.org/10.1371/journal.pone.0006912>
- Tanabe, S., & Cumming, B. G. (2008). Mechanisms underlying the transformation of disparity signals from V1 to V2 in the macaque. *Journal of Neuroscience*, 28(44), 11304–14. <http://doi.org/10.1523/JNEUROSCI.3477-08.2008>
- Tanaka, H., & Ohzawa, I. (2006). Neural basis for stereopsis from second-order contrast cues. *Journal of*



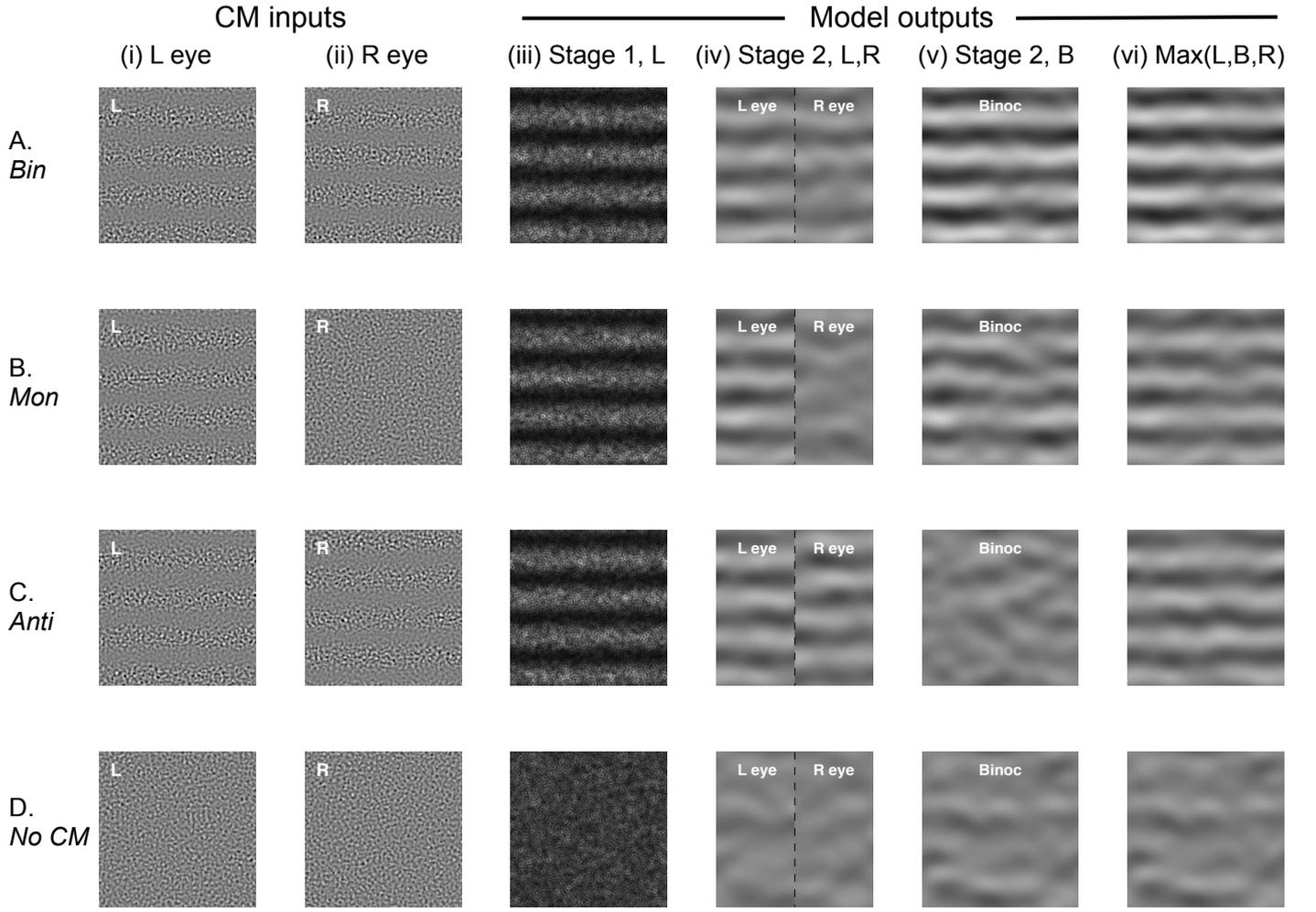


**Figure 1.** Experiment 1. Examples of left/right image pairs used in the various experimental conditions. Top row: luminance gratings (Lum), monocular signal (Mon), no noise carrier (contrast,  $c = 0$ ). Middle row: luminance modulation (LM) of a noise carrier, binocular in-phase signal (Bin), with correlated noise in the two eyes (correlation,  $r = 1$ ). Bottom row: contrast modulation (CM) of a noise carrier, binocular anti-phase signal (Anti), with anti-correlated noise in the two eyes ( $r = -1$ ).

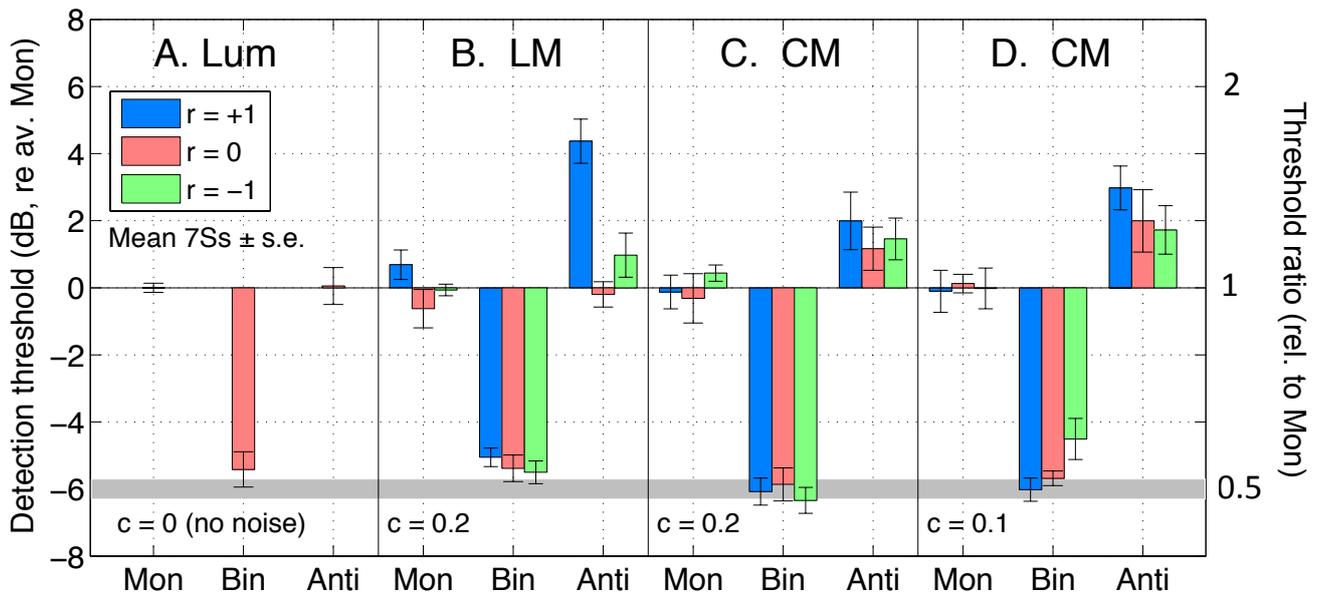
## Functional architecture for binocular CM



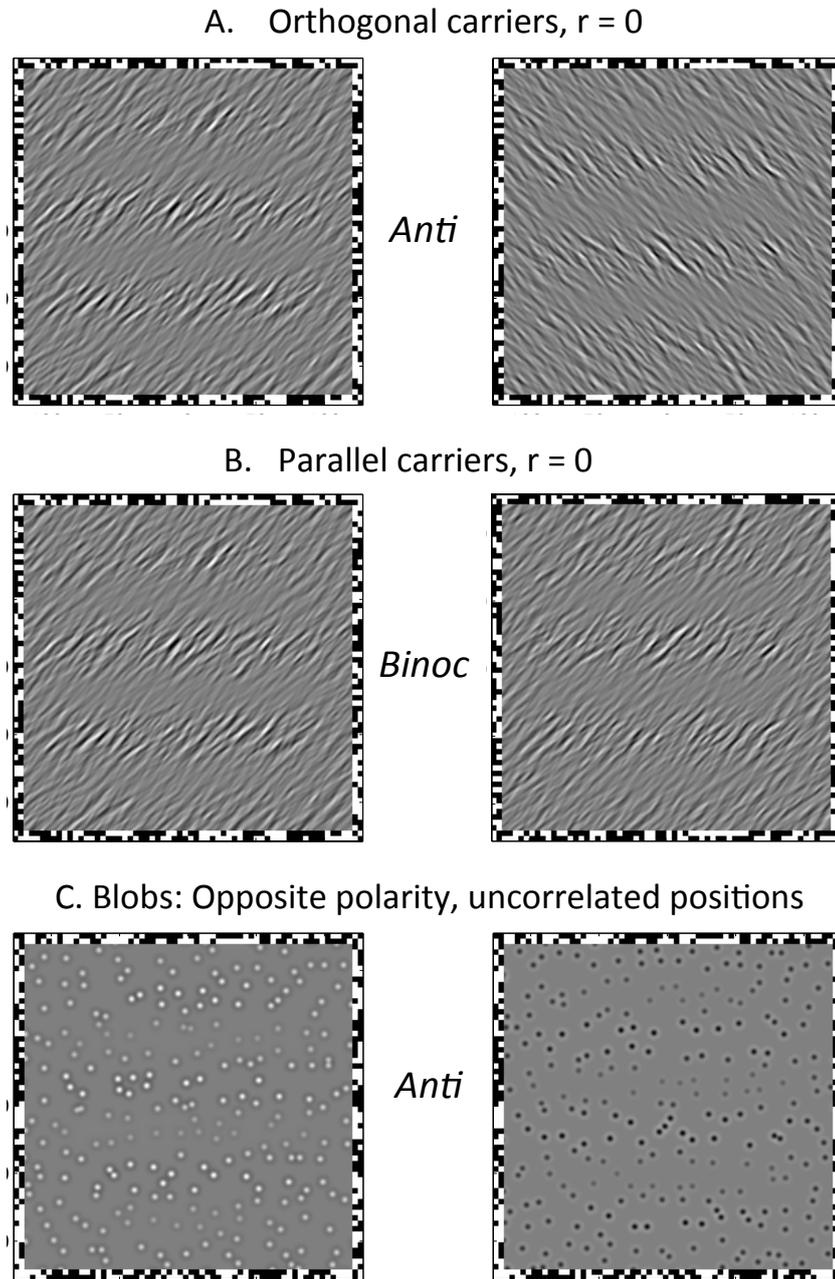
**Figure 2.** Functional architecture of a model that accounts in detail for our results. Monocular recovery of CM envelope signals by a standard filter-rectify-filter process (stage 1) is followed by contrast-weighted binocular summation (stage 2). The binocular weights are controlled by the relative contrast or contrast energy in each eye. Amplitude of the binocular response varies with the relative spatial phase of the target in the two eyes. Performance is limited by independent Gaussian noise in each monocular path. A steady background level of response ( $r_0$ ) is also added to both monocular paths ( $Bgd L$ ,  $Bgd R$ ) before binocular summation. The observation ( $Resp$ ) made in each interval is the largest of the three response amplitudes ( $r_L$ ,  $r_B$ ,  $r_R$ ). In a 2AFC trial, the model observer chooses the interval with the larger  $Resp$ .



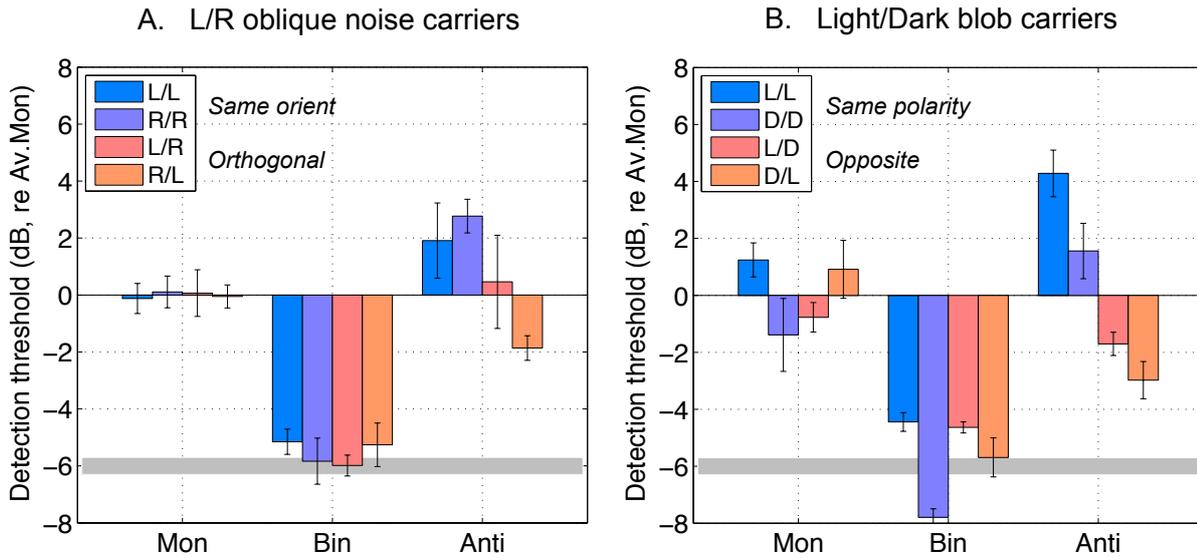
**Figure 3.** Visualizing the behaviour of our 3-channel model for CM detection. Columns (i) and (ii) illustrate the CM inputs to the model in the four key conditions of Experiment 1 (rows A-D). Column (iii) shows how the full-wave rectified output of stage 1 (Fig. 2) renders CM as a spatial variation in output intensity, to which the oriented second-stage filter is sensitive. Column (iv) shows the noisy, weighted outputs of the left and right eye second-stage filters, while column (v) shows their binocular sum. Note the greater binocular response (in A(v) vs B(v)) due to binocular summation; and the lack of a coherent binocular response to antiphase modulation (in C(v)) which, due to cancellation of the two signals, is similar to the noise-only condition (D(v)). Column (vi) shows which of the three response images (L,B,R) had the greatest amplitude at the signal frequency on that trial. With noise sampling variation, this *max* response image can from trial to trial switch between L and B in the monL condition (row B), or between L and R in the Anti condition (row C). The subtle influence of the background response  $r_0$  (eqn. 8,9) is not illustrated here, but is discussed in the text and Appendix.



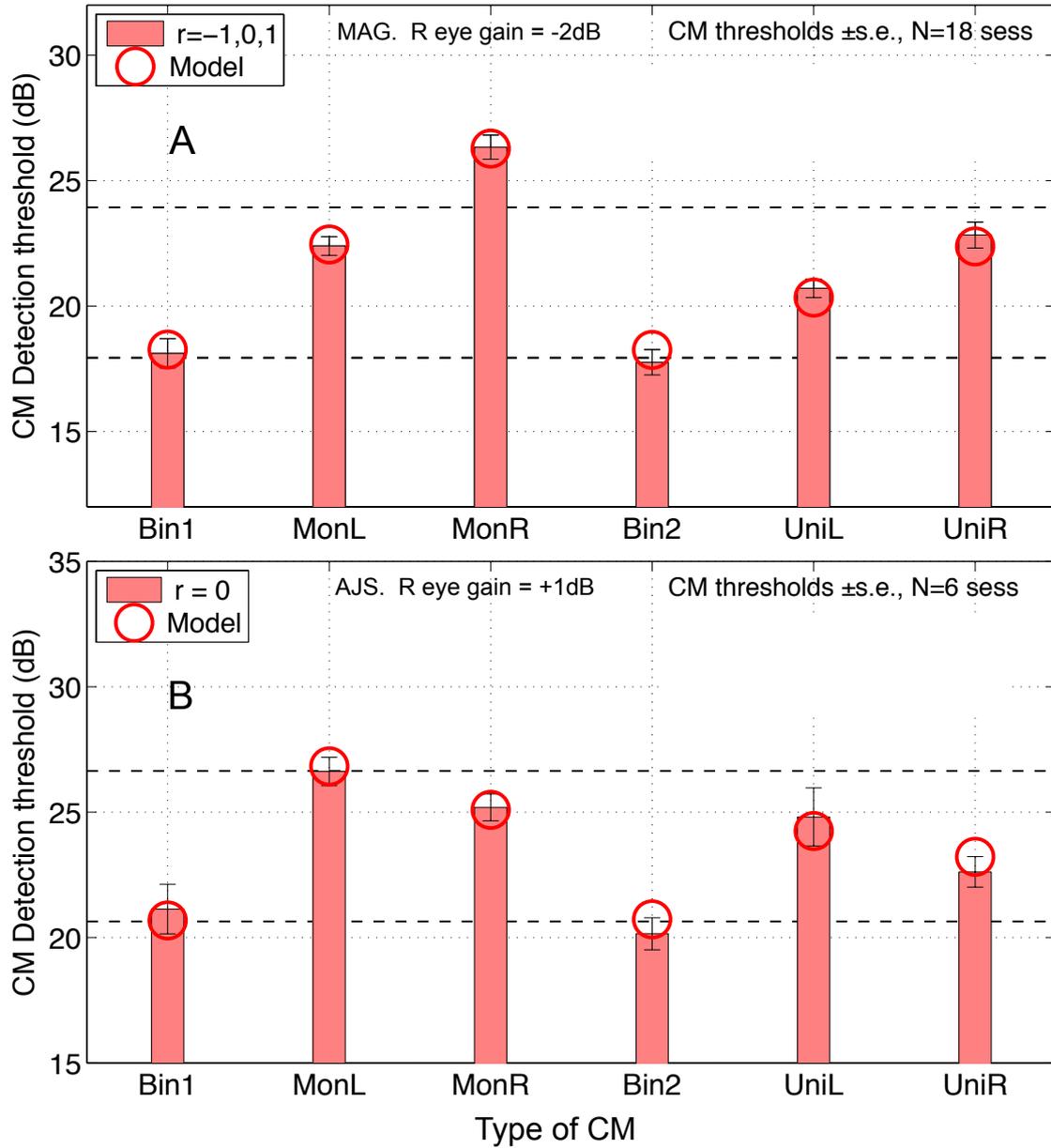
**Figure 4.** Experiment 1. Threshold modulations for monocular, binocular and antiphase signals in dB (mean of 7 Ss) expressed relative to the mean monocular threshold [which was averaged over subjects and over the levels of correlation ( $r$ ), separately for each of the 4 waveform conditions (Lum, LM, CM, CM)]. On the left axis, 0 dB is this monocular baseline; a binocular (Bin) threshold of -6dB (grey band) would be a halving of threshold (doubling of sensitivity) for binocular detection relative to monocular (see right axis). Error bars show  $\pm 1$  standard error of the thresholds, also in dB. To avoid inflating the error variance, these error bars were computed after subtracting absolute differences in observer sensitivity within each panel. As in repeated-measures analysis-of-variance, such between-subject differences are irrelevant to an evaluation of the stimulus-related effects. A: luminance grating (Lum). B: luminance grating in noise (LM). C, D: contrast modulation (CM). In B and C, r.m.s. contrast  $c$  of the 6 c/deg carrier was 0.2; in D it was 0.1. Colours in B-D represent correlation  $r$  between left and right eye carriers ( $r = 1, 0, -1$ ); the Lum grating (A) had no carrier.



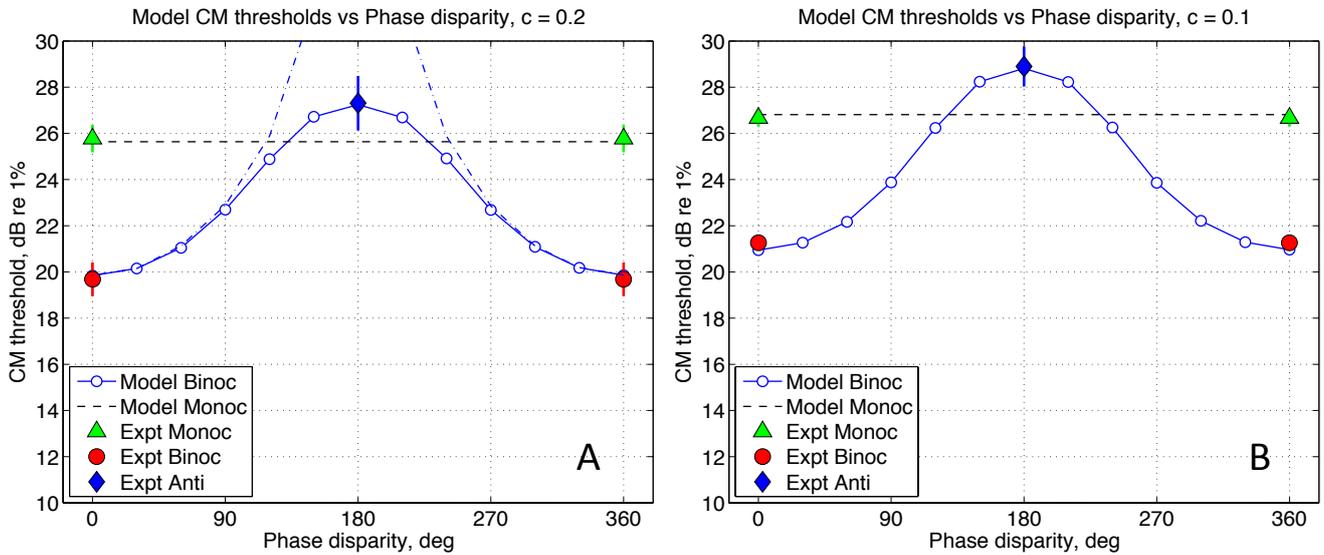
**Figure 5.** Example image pairs from Experiments 2 & 3. The carriers were dynamic as in experiment 1: a sequence of 4 different noise samples per 216 ms presentation. A, B: Experiment 2. Carriers were dynamic, oriented bandpass noise that had (A) *orthogonal* orientations ( $\pm 45^\circ$  from vertical) in the two eyes, or (B) the *same* orientations. CM signals could be binocular out-of-phase (A) or in-phase (B), or monocular. C: Experiment 3. Similar to experiment 2, but the carriers were dynamic, random arrays of circular blobs with light or dark centres (and dark or light surrounds) that had the *same* or *opposite* polarity across the two eyes. The blobs were luminance-balanced, so that adding light or dark blobs to the image did not change mean luminance. Each blob was defined as the circular second-derivative of a Gaussian ( $\partial^2 G / \partial x^2 + \partial^2 G / \partial y^2$ ), where  $G(x, y; s)$  is an isotropic Gaussian blob with standard deviation  $s = 2.23$  min arc (1.9 pixels), giving a peak SF of 6 c/deg. Blob positions were initially on a 16x16 square grid, but were jittered at random in  $x$  and  $y$ , with the constraint that blob centres were never closer than  $2s$  apart. Every image was jittered independently. Contrast modulation was Mon, Bin or Antiphase, as previously.



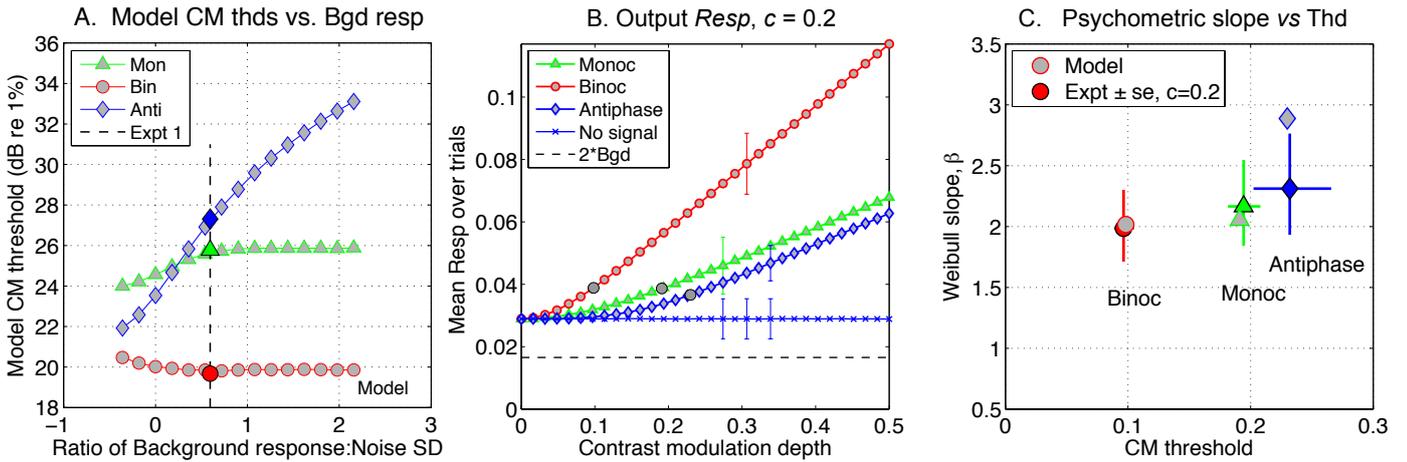
**Figure 6.** Experiments 2 & 3. Detection thresholds for Monocular, Binocular and Antiphase CM signals (mean of 3 Ss,  $\pm 1$  s.e.) expressed in dB relative to the average monocular baseline, as in Fig. 4. A: Experiment 2. Carriers were oriented bandpass noise (Fig. 5A,B) with the *same* orientation (both left oblique, L/L, or both right oblique, R/R) or *orthogonal* (L/R, or R/L) orientations in the two eyes. Binocular advantage for CM was unaltered by the potentially rivalrous nature of the orthogonal carriers. B: Experiment 3. Similar to experiment 2, but carriers were random arrays of light blobs (L) or dark blobs (D) (Fig. 5C) that had the *same* (L/L, D/D) or *opposite* polarity (L/D, D/L) across the two eyes.



**Figure 7.** Experiment 4. CM detection thresholds are plotted in dB for monocular modulation in the left or right eye (MonL, MonR; unmodulated carrier in the other eye), for unocular modulation (UniL, UniR; no carrier in the other eye), and for binocular in-phase modulation (Bin1, Bin2; tested twice in each session). Lower dashed line shows the mean Binoc threshold; the upper dashed line is 6dB higher. Circles show the fit of the 3-channel model (Fig. 2) - see Discussion/*Contrast weighted summation*. (A) subject MAG;  $\sigma = 0.0104$ ,  $r_0 = 0.0076$ ,  $r_0/\sigma = 0.73$ , rms error = 0.30dB. (B) subject AJS;  $\sigma = 0.0163$ ,  $r_0 = 0.0039$ ,  $r_0/\sigma = 0.24$ , rms error = 0.46dB.



**Figure 8.** Experiment 1. Model fit to CM thresholds at the two carrier contrasts: (A)  $c=0.2$ , and (B)  $c=0.1$ . Filled symbols show mean experimental thresholds  $\pm 1$  s.e. ( $n=7$ ), averaged over the 7 subjects and 3 levels of carrier correlation. Horizontal axis plots phase disparity ( $0$  or  $360^\circ$  for the Binoc condition, red circle;  $180^\circ$  for Antiphase, blue diamond). Monoc thresholds (green triangle) are arbitrarily located at  $0^\circ$  and repeated at  $360^\circ$ . Dashed horizontal line marks the model's Monoc threshold; open circles show how the model's threshold varied with phase disparity. Best-fitting parameters were (A)  $\sigma=0.0139$ ,  $r_0 = 0.0083$ ,  $r_0/\sigma = 0.59$ , rms error =  $0.10\text{dB}$ ; (B)  $\sigma=0.0079$ ,  $r_0 = 0.0054$ ,  $r_0/\sigma = 0.68$ , rms error =  $0.16\text{dB}$ . Antiphase signals cancel each other in the binocular channel, and so a model that uses only the binocular channel output (dash-dot curve in A) fails badly at disparities around  $180^\circ$ .

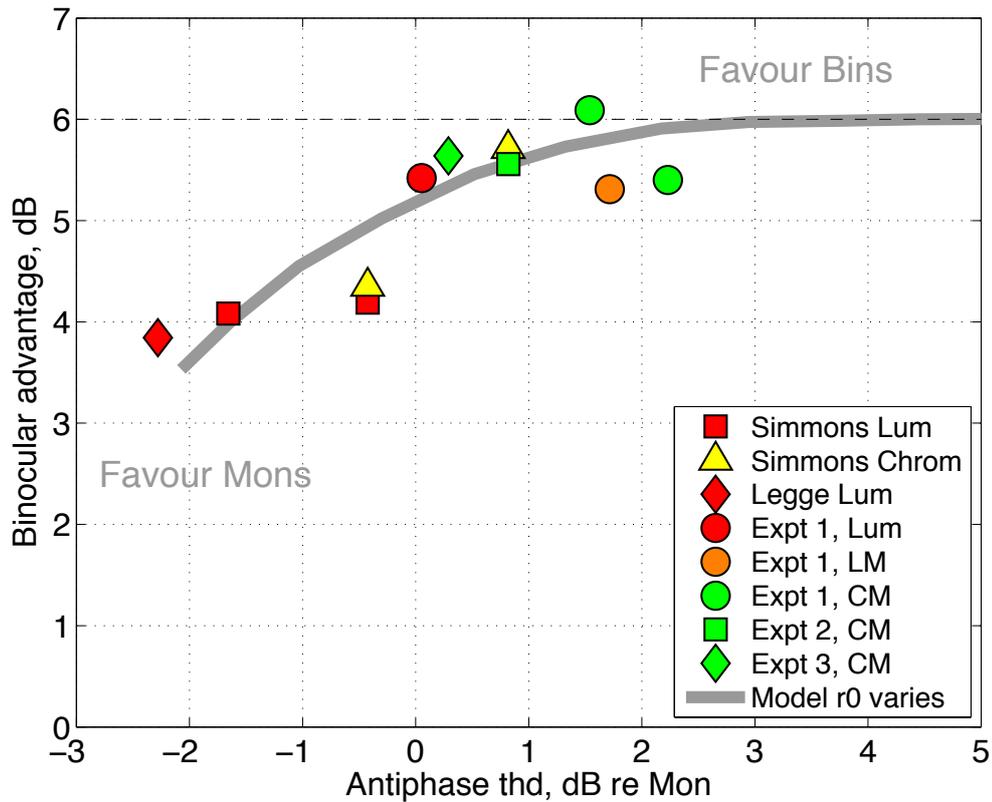


**Figure 9.** Probing the model's responses.

A. Model CM thresholds were computed for different values of the background response  $r_0$ , from -0.36 to 2.16 (expressed in units of the noise SD,  $\sigma$ ), with  $\sigma$  fixed at the best-fitting value (0.0139). While binocular thresholds (circles) reduced slightly with increasing  $r_0$ , monocular thresholds (triangles) increased by about 2dB, and antiphase thresholds (diamonds) became substantially worse by about 11 dB. Colour-filled symbols show how the group mean experimental thresholds (expt 1,  $c=0.2$ ) fell onto the pattern of model thresholds when  $r_0=0.6\sigma$  (vertical dashed line). For the data at  $c=0.1$ , the best-fitting estimate was similar,  $r_0=0.68\sigma$  (not shown).

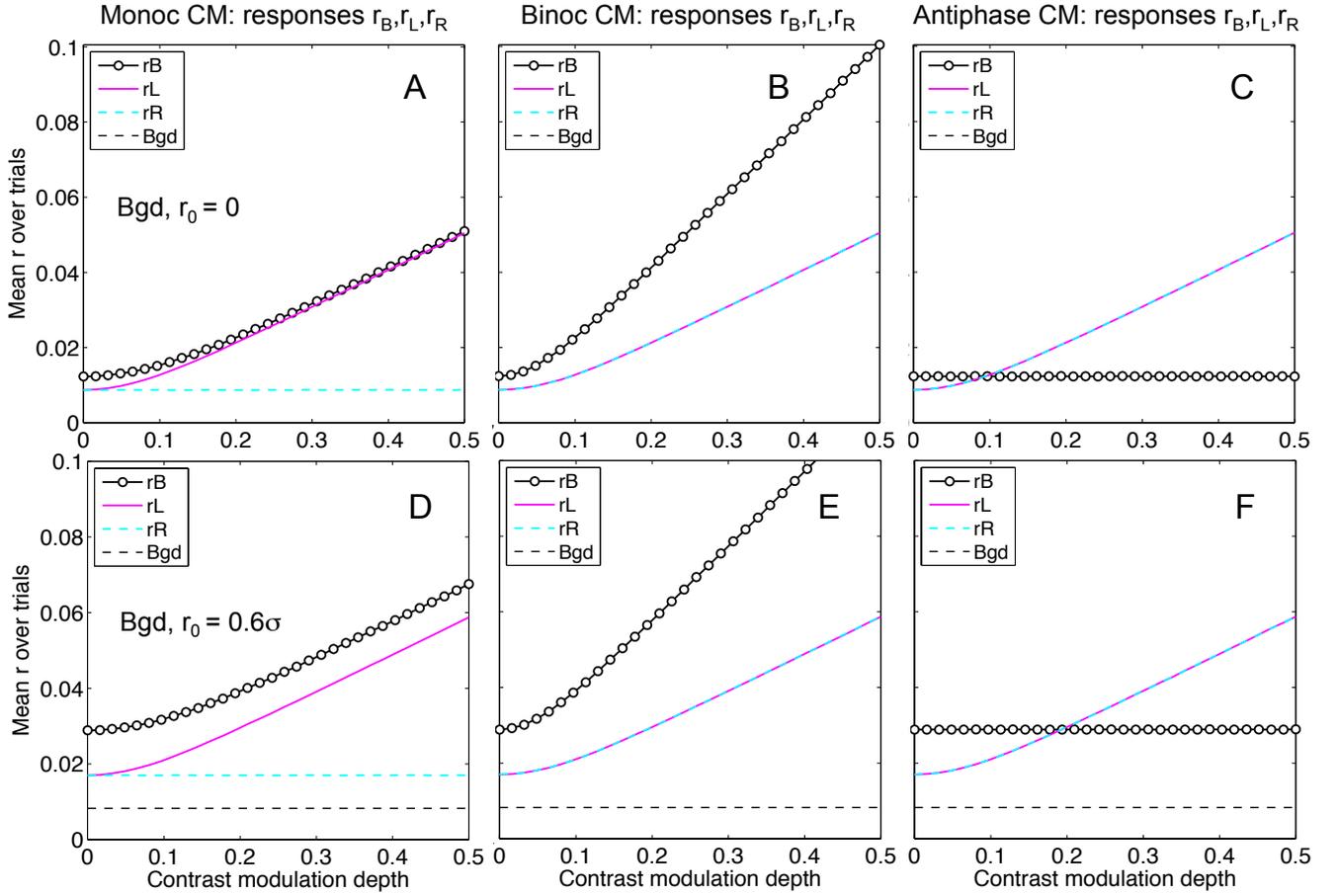
B. Mean *Resp* of the *max* operator for the 3 types of dichoptic CM used in experiment 1. Error bars show  $\pm 1$  s.d. of the response noise, as seen at this output stage. Grey circles mark the model's detection threshold (where  $d'=1.27$ , 81.6% correct) for each type of CM input.

C. Model thresholds and slopes ( $\beta$ ) of the Weibull psychometric functions fitted to model behaviour (grey-filled symbols) agreed well with the group means (coloured symbols  $\pm 1$  s.e.) observed in Experiment 1,  $c=0.2$ , averaged over the 7 observers and 3 levels of carrier correlation ( $r=1,0,-1$ ).

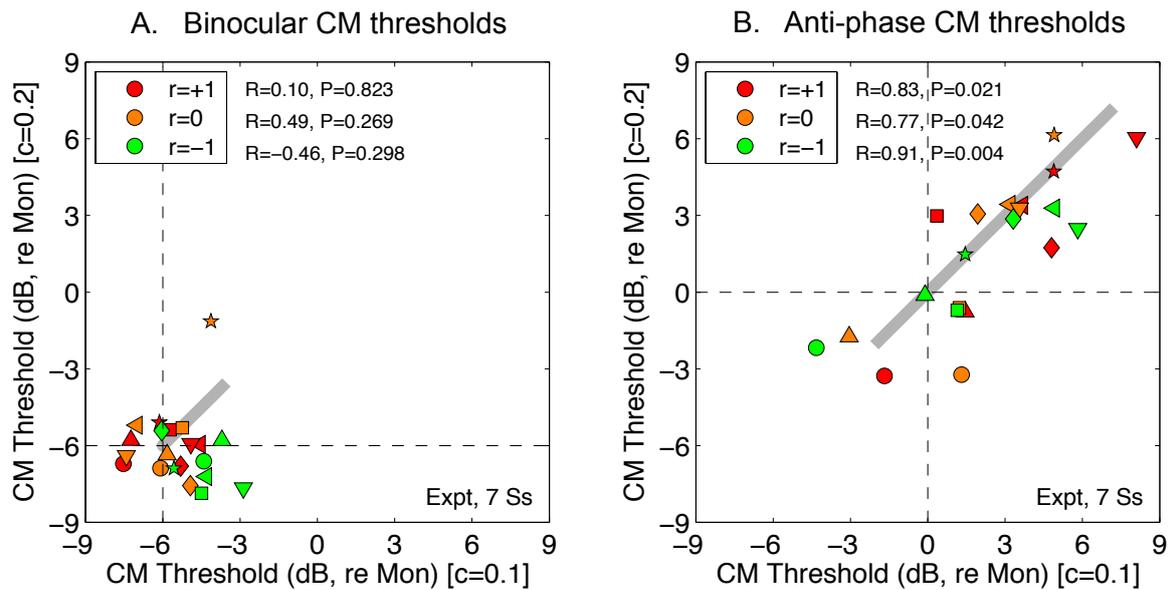


**Figure 10.** Empirical and theoretical relation between binocular in-phase and antiphase detection. Vertical axis is the binocular advantage (Monoc minus Binoc threshold, in dB); horizontal axis is the antiphase threshold, in dB, relative to the Monoc threshold. Upward trend represents increasing binocular superiority at the cost of reducing antiphase sensitivity.

- Red squares & yellow triangles: Simmons (2005), luminance & chromatic grating patches respectively, 0.5 c/deg; orientation horizontal or vertical (rightmost member of each pair of symbols was horizontal); mean of 4 Ss.
- Red diamond: Legge (1984), luminance grating, 0.5 c/deg; mean of 3 Ss.
- Red, orange, green circles: present study expt 1, Lum, LM, CM gratings respectively, 0.75 c/deg; mean of 7 Ss.
- Green square & diamond: present study expts 2,3; CM grating, mean of 3 Ss.
- Grey curve replots model thresholds from Fig. 9A in this new format, where  $r_0$  is now a parameter (it varies along the curve). Moving up the curve represents increasing background response  $r_0$  which increasingly favours the binocular channel.



**Figure A1.** Influence of background response  $r_0$  on model responses to CM, *before* the *max* operator. Each panel shows mean response amplitude of the binocular channel ( $r_B$ , open circles) and monocular ( $r_L$ , magenta solid curve;  $r_R$ , cyan dashed curve) channels at stage 2 (Fig. 2). Top row:  $r_0=0$ ; bottom row:  $r_0=0.6\sigma$ , the fitted value in experiment 1,  $c=0.2$  (Fig. 9A). (A,D): Monocular CM; (B,E): Binocular CM; (C,F) Antiphase CM. Although there is no explicit nonlinear transducer, mean responses rise nonlinearly with modulation depth because of noise and phase uncertainty (see *Model*, sec. 2.2). Black dashed horizontal line is the monocular background response ( $r_0$ ), in the absence of modulation. The summed binocular background response is  $2r_0$ , and this pushes up  $r_B$  relative to  $r_L, r_R$  (compare bottom and top rows), and makes  $r_B$  dominate the output *max Resp*.



**Figure A2.** Experiment 1: Individual differences and  $r_0$ . Symbols show CM thresholds for (A) binocular and (B) antiphase modulation, in dB relative to each subject's monocular threshold. Vertical axis plots thresholds for the higher carrier contrast ( $c=0.2$ ); horizontal axis for the lower contrast ( $c=0.1$ ). Each subject has a different symbol shape; red, orange, green represent the three carrier correlations  $r = 1, 0, -1$ . Significant correlations  $R$  listed in (B) imply consistent individual differences in antiphase thresholds, but for binocular thresholds (A) individual differences were much smaller and apparently random (no significant correlation). Thick grey line shows the range of variation in model thresholds to be expected as  $r_0$  varies over the range shown in Fig. 9A. Model antiphase thresholds (grey line, B) increased markedly with increases in  $r_0$ , but model binocular advantage (grey line, A) changed only a little. This modelling shows that individual differences in background response level  $r_0$  could account fairly well both for the individual differences in antiphase thresholds (B), and the similarity of binocular advantage across observers (A).