Chapter 4: Additivity Experiments

1 Introduction

In the last two chapters we have shown evidence for subject’s ability to efficiently pool Fourier power across planar regions of frequency space. In this chapter we test one of the prominent predictions of the planar power detector model: additive pooling of energy on the plane. The primary reason for testing additivity is that the existence of an additive law suggests observers are using specialized pooling mechanisms to detect stimuli, since contrast pooling across frequency bands is typically subadditive [15, 14]. Conversely, a non-additive pooling rule would not require the use of specialized mechanisms, and parsimony of strategy would argue that the pooling observed in detecting the stimuli is due to a general rule.

Another reason to test additivity is that the ideal pooling strategy for this task is additive. By measuring the observer’s conformance to an additive law, we can infer what proportion of the subject’s inefficiency is due to an inappropriate pooling rule. Finally, the results of the analysis in the last chapter could be substantially improved by having knowledge of the pooling rule, either in the strength of the conclusions if the rule is additive, or by suggesting an improved psychometric model for a re-analysis.

In the rest of the paper, we explain why the planar power detector model predicts additivity, and describe the stimuli and the experimental procedure. We then explain the data analysis and show the results. We discuss the results in terms of models of motion processing, possible relations to physiology, and the possibility of partial adaptability of orientation pooling.

1.1 Additivity predictions

The planar power detector additively pools power concentrated around a common plane to measure the spectral energy $E$. If we split the signal into several non-overlapping bands, then the detector can be described as summing the energies within each signal band, weighted by the detector’s spectral sensitivity. If $|S(\omega)|^2 = \sum_i |S_{bi}(\omega)|^2$ denotes the signal spectrum and $|H(\omega)|^2$ denotes the power spectrum of the detector, then the output $E$ can be expressed as:

$$ E = \int \sum_i w_i E_{bi} $$

where $w_i$ are the weights for each band.
Where $E_{b_i}$ is the signal energy in band $b_i$, and the $w_i$ are weights which represent the average effect of the detector’s sensitivity within band $b_i$ on the signal energy in the band. We show in the Appendix that the performance of a planar power detector is well described by a psychometric function which only depends on the energies in the signal bands and the background noise power level $N$. If we denote the psychometric function by $\Psi$, then the probability of a correct response $R_i$ for an observer relying on a planar power detector is given by:

$$p(R_i = 1) = \Psi\left(\sum_i w_i E_{b_i}, N\right)$$

(2)

If we fix $N$, then for any fixed probability correct the weighted sum of the signal energies must equal a constant: $\sum_i w_i E_{b_i} = c$. Assuming that the observer’s performance is entirely based on planar power detectors, then we should observe the same performance for any combination of the energies whose weighted sum is equal to the same constant $c$.

*Prediction 1:* In a task in which an observer relies on planar power detectors, the observer should additively combine the energies from bands which intersect a common plane in frequency space.

*Prediction 2:* If planar configurations of power are specially processed by the visual system, we would expect subadditive combination of non-planar configurations of power.

To test these predictions we used an experimental paradigm similar to the one used in the previous experiments.

### 1.2 Experimental Logic

The basic idea is to combine two sets of bandpass signals in different ratios of signal energy. If the signal sets are being additively combined, then performance should be determined by the total signal energy independent of the ratio. We used two different types of bandpass signal sets, *plaid* signals and *bandpass* signals. The *bandpass* signals are produced by passing spatio-temporal white noise through a filter which is bandpass in spatial orientation, spatial frequency magnitude, and temporal frequency. These signals can be described as noisy 'gratings', since the stimuli have dominant spatial and temporal frequencies. The *plaid* signals are formed by summing two equally weighted bandpass components which have different peak spatial orientations but which lie on a common plane. Since each of the components can be described as a noisy 'grating', it is natural to call these combinations 'plaids'.

The plane in frequency space associated with each velocity can be specified by two frequencies on the plane, since the plane is constrained to pass through the origin. Thus each plaid signal determines a unique plane through the component’s peak frequencies. The minimum number of bands which can be used to test for pooling on and off a common plane is three: a plaid which defines the common plane and a third 'test' band which can be added to the plaid in various energy ratios to test additivity. Two different plaids signals and three different bandpass signals were used in three different experimental conditions. The expected signal spectra for the three conditions is shown in figure 1.

Two of the conditions, designated *In-Plane* and *Asymmetric* involve combining plaid and bandpass signals which lie on a common plane in various energy ratios. In the In-Plane condition the plaid signal has bandpass components which are symmetrically oriented around the direction of motion of the translation specified by the common plane. The bandpass signal has an orientation which lies between the plaid orientations, and thus is in direction of the plaid pattern motion. Since the additivity prediction does not depend on which bands are traded off in the plane, we chose a different combination of the same three bands for the second condition. In the Asymmetric condition, the plaid is asymmetric and is formed from the signals in the In Plane condition by combining one of the plaid components with the bandpass signal. The bandpass signal in the Asymmetric condition is just the other plaid component from the In Plane
condition. In this condition the combination signal has the same perceived direction of motion as the In Plane condition, however, the plaid and bandpass signals alone are perceived to move in directions different from the common motion.

The third condition was designated *Off Plane*, and involved combining the symmetric plaid from the In Plane condition with the bandpass signal from this condition rotated out of the common plane to lie near the zero velocity. This off plane bandpass signal was chosen because the spatial frequency spectra of the components in the In Plane and Off Plane conditions is nearly identical, and it allows us to separate orientation specific pooling from motion specific pooling without the problems potentially caused by motion opponency.

Two aspects of the experimental design promoted the reliance of the observer on planar power detectors. The first aspect was that the combination signal had a large range of spatial orientations. Using a large range of orientations encourages the observer to use an internal filter which is broadband in orientation, like a planar power detector. The second aspect was that all the combinations of the plaid and bandpass signals were randomly intermixed. Random intermixing encourages the observer to use a strategy which combines across the ensemble of signals presented. For the experimental setup, additivity is not the optimal combination rule. The ideal observer in this task uses power detectors which are matched to the plaid spectrum and the bandpass spectrum, computes the energies in each signal band for a large set of possible plaid and bandpass power levels, and averages the exponentiated energies across the set of signal power levels. However, if the observer’s most sensitive detectors for the task are planar power detectors, the best the observer can do is to sum together the bands in the plane weighted by the expected power for the band.

1.3 Data Presentation

The natural way to present the data from the additivity experiments is to plot the constant % correct thresholds in \((E_b, E_{pl})\) space, illustrated in figure 2. Additivity shows up in this type of plot as a straight line with negative slope which connects the plaid alone and bandpass alone thresholds. To assess deviations from additivity we used a generalized summation equation:

\[
1 = \left(\frac{E_b}{T_{E_b}}\right)^\alpha + \left(\frac{E_{pl}}{T_{E_{pl}}}\right)^\alpha \tag{3}
\]

\[
c = E_b^\alpha + s \cdot E_{pl}^\alpha \tag{4}
\]

Additivity plots \([15, 33]\) and analogues to equation 4 have been previously used to assess additivity. Equation 4 occurs naturally in the context of probability summation \([14]\) and as the solution to a functional equation for pooling \([35]\).

For \(\alpha > 1\), equation 4 leads to subadditive combinations, i.e. to thresholds which are larger than the sum of the component thresholds. Also for integral \(\alpha > 1\) this expression can represent probability summation among \(\alpha\) different independent bands. For \(\alpha < 1\), this equation leads to superadditive combinations, i.e. to thresholds which are smaller than the sum of the component thresholds. The form of the equation was chosen as a simple way to parametrize additive vs. non-additive combinations. By fitting this equation to the resulting thresholds, we can determine the summation rule from the value of \(\alpha\).

2 Methods

2.1 Stimuli

The method for producing the stimuli was the same as chapter 2. Stimuli were produced by passing spatio-temporal gaussian white noise through the set of filters described below.
Figure 1: Filters used in additivity experiments. The illustrations are presented as a table with the rows corresponding to different additivity conditions and the columns to the bandpass filters and their combination. The three different additivity conditions, In Plane, Off Plane, and Asymmetric designate properties of the stimuli.
2.2 Filters

All of the filters used were rotated copies of a single bandpass filter. This filter has the following functional form:

\[ BP(\omega_r, \omega_\theta, \omega_\phi) = W_r(\omega_r)W_\theta(\omega_\theta)W_\phi(\omega_\phi) \]  

(5)

where \( W_x \) is a smooth box function on the variable \( x \) (see methods section Chap 2). Smooth box functions were used because they allow fine control over the placement and smoothness of the spectral boundaries. \( W_r \) had a transition region width of 1.45, and low-high frequency cutoffs of (0.49,7.6), where the frequency radius of the sphere is given by \( \omega_r = \sqrt{\omega_x^2 + \omega_y^2 + (\omega_t/2.1)^2} \). \( W_\theta \) and \( W_\phi \) had a transition widths of 8 degrees, and the high low cutoffs which spanned 36 degrees.

All of the other filters were simply combinations of 3-D rotated copies of this base filter. If we choose the base position of the \( BP \) filter to be centered around the \( \omega_x \) axis, then we can describe the positions of the other filters by the composition of two rotations of \( BP \). Let \( \vec{\omega} \) represent the vector \([\omega_x, \omega_y, \omega_t]\), and \( R_\alpha(\phi_0) \) and \( R_\alpha(\theta_0) \) denote the 3-D rotation matrices which leave the \( \omega_x \) and \( \omega_t \) axes fixed respectively. The composition \( R_x R_t \) rotates the filter away from the \( \omega_x \) axis by \( \theta_0 \) degrees and then away from the spatial frequency plane by \( \phi_0 \) degrees. Most of the filters were rotated up to lie in a common plane which specified a downward motion with a speed of 1.93 deg/sec.

**In Plane Condition** The symmetric plaid filter is formed by summing together two copies of the \( BP \) filter. The rotations angles for \( R_x \) are \( \theta_0 = 28 \) deg and \( \theta_0 = 152 \) deg, which makes the orientations symmetric around the \( \omega_y \) axis. Each of the filters have \( R_t \) rotation angles of \( \phi_0 = 36.9 \) deg so that the bands lie in a common plane.

The bandpass filter for the Symmetric condition has rotations angles of \( \theta_0 = 90 \) deg and \( \phi_0 = 36.9 \) deg so that the plaid and bandpass filters lie on a common plane, with the bandpass spatial orientation orthogonal to the direction of motion specified by the plane.
Figure 3: Diagram illustrating the data collection method. Data was collected for 6 different constant $E_{pl}/E_d$ ratios shown as gray arrows. The ratios are presented to the right of the arrows. The gray circles represent the points the data was collected at. The ratios were chosen so that the distances $d$, between points along the diagonal are equal on normalized energy axes. The data was analyzed by fitting Weibull functions to the data along each constant ratio ray.

Asymmetric Condition The asymmetric plaid filter is also formed by summing together two copies of the $BP$ filter. The rotations angles for $R_x$ are $\theta_0 = 28$ deg and $\theta_0 = 90$ deg. The bandpass filter for this condition has $\theta_0 = 152$ deg. Each of the filters have $R_t$ rotation angles of $\phi_0 = 36.9$ deg so that the bands lie in a common plane.

Off Plane Condition In this condition the plaid filter is the same as in the In Plane condition. The bandpass filter has the same $\theta_0 = 90$ deg, but it does not lie in the plane. $\phi_0 = 3$ deg for this condition, hence the filter is nearly centered around the spatial frequency plane.

2.3 Procedures

Data were collected using a 2IFC task, in which subjects discriminated signal plus noise and noise alone intervals. Signal stimuli in each condition were additive mixtures of one of the plaid stimuli with one of the bandpass stimuli. Data from each condition was collected in separate sessions, but the sessions were intermixed. Subjects were provided with knowledge of which stimuli were to be detected at the beginning of each session. Subjects were also given two hours practice on each condition prior to data collection.

Figure 3 illustrates the data collection method. Signal energy was varied using the method of constant
stimuli for six different constant plaid-bandpass energy ratios ($E_{pl}/E_{bp}$), shown in the figure as grey rays. Five different combination energies were used to estimate the psychometric function along each ray, shown as open circles in the figure, for a total of 30 different combinations. At each combination, 100-120 trials were collected. It required 3/4 hours to collect all the trials for a condition. To avoid subject fatigue the data collection for each condition was split into one hour sessions.

To ensure that the constant ratio rays were evenly distributed, we used estimates of the subject’s thresholds for plaid and bandpass stimuli alone to distribute the measurements across the ($E_{pl}$,$E_{bp}$) plane. We measured the subject’s thresholds for each of the plaid and component stimuli alone using the method of constant stimuli. We then determined the ratios which caused the constant ratio rays to divide the line connecting the 80% correct $E_{pl}$ and $E_{bp}$ estimates into equal length segments. Thus the ratios were different for each subject and condition. The energy ratios used are presented in table 1.

Thresholds were determined by fitting Weibull functions to the detection data along each constant ratio ray using a maximum likelihood procedure. Error bars for the thresholds were computed from the inverse numerical Hessian of the likelihood function for threshold, which were cross-validated using a parametric bootstrap procedure. In the bootstrap procedure, 1000 data sets were simulated by sampling from the binomial distribution with the parameter $p$ given by the measured probability correct. Maximum likelihood fits of the parameters were then generated for each data set. The resulting distributions of fitted parameters were used to estimate the standard error on the parameters.

**2.4 Data Analysis**

Additivity was assessed by fitting the following equation to constant %correct threshold points along each of the constant ratio rays:

$$E_{bp}^c + (sE_{pl})^\alpha = c^\alpha$$

This equation represents a Minkowski metric model of pooling [14] which frequently arises in the context of probability summation models. In the present context it provides a simple means for parametrizing additivity through the exponent $\alpha$. $\alpha < 1$: superadditive, $\alpha = 1$: additive, $\alpha > 1$: subadditive. The ‘slope’ $s$, gives us a measure of the relative weights given to the plaid and bandpass energies, when additivity holds.

The model was fit to the data using non-linear least squares minimization. The squared distances along the constant ratio rays between the measured threshold energies and the curve described by the equation were inversely weighted by the variances of the threshold estimates. The sum of these weighted distances were minimized over three parameters, $\alpha$, $s$, and $c$ using the Broyden-Fletcher-Goldfarb-Shanno variable metric multidimensional minimization method [25].

Statistics on the fits were generated using a parametric bootstrap procedure. In the procedure, bootstrap fits of the psychometric functions were used to generate 1000 estimates of each of the energy thresholds. Least squares fits for each of the estimates was performed, generating distributions for $\alpha$, $s$, and $c$.  

<table>
<thead>
<tr>
<th>Subject</th>
<th>Condition</th>
<th>$E_{pl}/E_{bp}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PS</td>
<td>In Plane</td>
<td>1:0 6.6:1 2.5:1 1.1:1 0.4:1 0:1</td>
</tr>
<tr>
<td></td>
<td>Off Plane</td>
<td>1:0 5.8:1 2.2:1 0.96:1 0.36:1 0:1</td>
</tr>
<tr>
<td></td>
<td>Asymmetric</td>
<td>1:0 6.0:1 2.3:1 1.0:1 0.38:1 0:1</td>
</tr>
<tr>
<td>ML</td>
<td>In Plane</td>
<td>1:0 6.4:1 2.4:1 1.1:1 0.4:1 0:1</td>
</tr>
<tr>
<td></td>
<td>Off Plane</td>
<td>1:0 4.5:1 1.7:1 0.7:1 0.3:1 0:1</td>
</tr>
<tr>
<td></td>
<td>Asymmetric</td>
<td>1:0 5.9:1 2.2:1 1.0:1 0.38:1 0:1</td>
</tr>
</tbody>
</table>

**Table 1:** Table of constant ($E_{pl}/E_{bp}$)


Variances for the parameter estimates were computed from these distributions. One way ANOVAs and T-tests were performed on the parameter estimates using these variance estimates as the within-condition variances. Since we could use as many bootstrap samples as desired, the within-condition degrees of freedom were effectively infinite. We used a large positive number $10^5$, instead of infinity for the number of degrees of freedom.

3 Results

3.1 In Plane Results

The results for the In Plane condition are shown in figure 4, plotted in plaid-bandpass threshold energy space. Each data point represents the threshold energy along a constant ratio ray for one of four different % correct values: 60%, 70%, 80% and 90%. The error bars are oriented along the constant ratio rays, and were estimated from the psychometric function fit.

The dashed lines represent approximate constant performance contours, while the solid lines represent the curves generated by the best fitting parameters of the pooling equation. When additivity holds the performance contours should lie along straight lines, or equivalently, the best fitting additivity equation exponents $\alpha$ should be 1. We see that for both subjects the best fitting curves are essentially linear. The best fitting $\alpha$s for each constant performance curve are gathered into a table on the right side of the figures. Inspection shows that all of the $\alpha$s are very close to 1. Subject PS shows a small but consistent trend for $\alpha$ to increase with increasing % correct, however, none of the alpha are significantly different from 1 (T-test, 0.05 level). The alpha for Subject ML do not show an increased trend, and are clustered more tightly around 1. Thus the visual system can be described as additively pooling the bands in this condition.

3.2 Asymmetric Results

The results for the Asymmetric condition are similar to the On Plane condition, as predicted by the planar power detector model. The $\alpha$ estimates for subject PS are 0.2-0.3 higher than in the On Plane condition, however, none of the $\alpha$ are significantly different from 1 (T-test, 0.05 level). Thus the visual system is able to additively pool power as long as the components lie on a common plane. It also shows that subjects are not using a detection approach which requires the phenomenal motion of the plaid and bandpass components to be the same for additive pooling to occur.

3.3 Off Plane Results

The results for the Off Plane condition are very different. Notice that the best fitting curves are curved in the subadditive direction. T-tests show that all of the $\alpha$ for this condition are significantly greater than 1 at the 0.01 level. Thus subjects are not able to additively pool bandpass power off the common plane specified by the symmetric plaid. This result agrees with the results from Chapter 2, which showed that subjects are less efficient at detecting non-planar stimuli.

There is a highly significant trend for $\alpha$ to decrease as % correct increases for both subjects, which is discussed below.

3.4 Additivity exponents

The fitted additivity exponents $\alpha$ are summarized in figure 7. As noted before, the salient feature of the data is that exponents are clustered around 1 for the In Plane and Asymmetric conditions, but are significantly
On Plane Configuration

Figure 4: In Plane additivity data. None of the estimated $\alpha$ are different from 1 at the 0.05 level using a T-test.
Asymmetric Planar Configuration

Figure 5: Asymmetric additivity data. None of the estimated $\alpha$ are different from 1 at the 0.05 level using a T-test.
Figure 6: Off Plane additivity data. All of the estimated $\alpha$ are significantly different from 1 at the 0.001 level using a T-test.
Figure 7: Additivity exponents for the three conditions.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Comparison</th>
<th>p: 60%</th>
<th>p: 70%</th>
<th>p: 80%</th>
<th>p: 90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>PS</td>
<td>In Plane vs. Asymmetric</td>
<td>0.19</td>
<td>0.29</td>
<td>0.26</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>In Plane vs. Off Plane</td>
<td>&lt; 0.001</td>
<td>0.003</td>
<td>&lt; 0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Asymmetric vs. Off Plane</td>
<td>&lt; 0.001</td>
<td>0.008</td>
<td>&lt; 0.001</td>
<td>0.006</td>
</tr>
<tr>
<td>ML</td>
<td>In Plane vs. Asymmetric</td>
<td>0.33</td>
<td>0.75</td>
<td>0.77</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>In Plane vs. Off Plane</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Asymmetric vs. Off Plane</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Table 2: Table of $\alpha$ comparisons across condition at each probability correct, showing the probability that the compared $\alpha$s are drawn from the same distribution using the T-test statistic. The Off Plane condition is significantly different from the other two conditions at better than the 0.01 level for both subjects.

greater than 1 for the Off Plane condition. The planar power detector model predicts that the pooling rule will be additive and hence equivalent for conditions In Plane and Asymmetric, but will be significantly non-additive for the Off Plane condition. We performed a set of T-tests of the $\alpha$ estimates between the conditions at each % correct level, the results of which are presented in table 2. The analysis shows that the two conditions in which the bands are coplanar, In Plane and Asymmetric, are not significantly different (max $p = 0.19$). In contrast, both of these conditions are significantly different ($p < 0.01$) from the Off Plane condition.

### 3.4.1 Invariance across %correct

In our analysis of the ideal weighting strategy we predicted that the additivity result should hold equally for any % correct slice we choose to analyze. We tested for invariance of the fitted pooling equation parameters across the four % correct levels by performing a one-way ANOVA. The results of the analysis are gathered in table 3. The estimates of $\alpha$ across % correct are not significantly different at the 0.05 level for the In Plane and Asymmetric conditions for either subject. Thus for the conditions in which all of the bands lie in a single plane, the assumption of additivity independent of the level of performance cannot be falsified.

There is a significant trend in the Off Plane data which is nearly identical for both subjects. The trend is toward decreasing $\alpha$ with increasing % correct. The trend could be the result of pooling efficiency
Table 3: Table of ANOVA results for the estimates of $\alpha$ across % correct. $p$ gives the probability that each of the $\alpha$ are drawn from the same distribution, i.e. that there is no significant change in alpha across the % correct slices. The Off Plane condition is significant at the 0.05 level for both subjects, shown by the starred $p$ values.

changing with % correct, such that subjects pool more efficiently at higher signal levels. If this hypothesis were true we would expect the zero mixture Plaid thresholds for the In Plane and Off Plane conditions to be roughly equivalent and for the mixture thresholds in the Off Plane condition to improve with % correct. Expressed in terms of the slopes of psychometric functions, it predicts that the zero mixture plaid psychometric slopes will be equivalent while the non-zero mixture will be shallower in the Off Plane condition.

Instead we find that the Weibull slope parameters $\beta$ along the plaid only axis are steeper for the On Plane than Off Plane conditions (PS: $\beta_{\text{On Plane}} = 1.69$, $\beta_{\text{Off Plane}} = 1.14$, ML: $\beta_{\text{On Plane}} = 1.4$, $\beta_{\text{Off Plane}} = 1.13$), while the average mixture slopes are essentially equivalent ( PS: $\beta_{\text{On Plane}} = 1.57$, $\beta_{\text{Off Plane}} = 1.52$, ML: $\beta_{\text{On Plane}} = 1.4$, $\beta_{\text{Off Plane}} = 1.49$ ). The decreased slope of the plaid energy psychometric function in the Off Plane condition is the result of lower plaid thresholds at the low signal level/low % correct end in the Off Plane condition than the In Plane condition, and slightly higher thresholds at the high signal level/high % correct end. This suggests that the plaid stimuli are being processed differently in the presence of the Off Plane bandpass stimuli than the In Plane bandpass stimuli. The result suggests that the source of the variation across alpha is in the processing of the plaid component stimuli. To test this idea, we refit the Off Plane data using the plaid only mixture data from the On Plane condition. The resulting $\alpha$ estimates showed no trend across % correct, PS: $\alpha_{\text{mod}} = (2.7, 2.6, 2.6, 2.7)$, ML: $\alpha_{\text{mod}} = (2.1, 1.9, 1.9, 2.0)$.

3.5 Weighting across bands and the fitted slopes

Given the result of additivity, the slopes for the In Plane and Asymmetric conditions can tell us something about the relative weighting across the three bandpass components which comprised the mixtures in the two conditions. The slopes are given by the parameter $s$ in the pooling equation 4, and the fitted slopes are shown in figure 8.

We can use the slopes in the In Plane and Asymmetric conditions to derive estimates of the weights across the three stimulus bands. Let $w_a$ denote the weight for the Asymmetric bandpass component, $w_b$ the weight for the In Plane bandpass component, and $w_c$ the weight for the remaining band. An equivalent rule for naming the bands is to label them $\{c, b, a\}$ going counterclockwise in spatial orientation from the $\omega_x$ axis. Then in the In Plane condition, the slope $s_{\text{in}}$ in $s_{\text{in}}E_{\text{bp}} + E_{\text{pl}} = e$ is equal to

$$s_{\text{in}} = \frac{w_b}{(w_a + w_c)/2}$$  \hspace{1cm} (7)

Thus $(w_a + w_c) = s_{\text{in}}/2w_b$. In the case $w_a = w_c$, then these weights are 0.74 for subject PS and 0.48 for
subject ML, relative to \( w_b = 1 \). In the Asymmetric condition the slope is given by:

\[
s_{\text{asym}} = \frac{w_a}{w_b + w_c} / 2
\]  

(8)

If we assume that the weighting is identical in the two experiments and set \( w_b = 1 \), then we can use equations (7) & (8) to solve for the weights for \( w_a \) & \( w_c \) relative to \( w_b \). The resulting weights for both subjects are gathered in table 4.

Instead of an equal split between the weighting as expected by symmetry, assuming the invariance of weights across conditions leads to a large bias in the weighting of bands corresponding to leftward and downward moving gratings over those moving rightward and downward. However, to derive the weights, we only considered the slopes and not the magnitude of the thresholds. If both In Plane and Asymmetric weights are equal, then the threshold energies at a given % correct should also be equal, since the conditions obey equation 2. Thus, we can use the common weights to compute a predicted Asymmetric bandpass threshold from the In Plane threshold energy at a given % correct by multiplying \( w_a \) by the In Plane plaid threshold. We find that the actual thresholds in the Asymmetric condition are 27% & 34% higher on average than the predictions from the In Plane condition for subjects PS and ML respectively. Using the variance of the Asymmetric bandpass thresholds to construct a T-test, we found that these thresholds are significantly different from the predictions at the \( p < 0.02 \) level. Thus, it seems likely that the assumption that the weights are constant across the two conditions is false.

The other possibility is that subjects are able to adjust their weighting across bands, using different sets of weights for the In Plane and Asymmetric conditions. Subjects will perform better in the experimental conditions if they weight the three frequency bands by scalars proportional to the expected energy in each band. Because the additivity experiments involved intermixing different sets of stimuli in the In Plane and Asymmetric conditions, the expected energies, and hence the optimal weights are different for the two conditions. We computed the optimal weights for each subject (see Appendix B) in each condition, shown in table 4. The weights show that subjects can do better on average if they weight the bandpass band more heavily than the plaid bands. If we assume that subjects equally weighted the plaid bands in each condition (i.e. In Plane: \( w_a = w_c \), Asymmetric: \( w_b = w_c \)), we can compare the optimal weights with

\[\text{Figure 8:}\] Slopes of the fits the three conditions. A one-way ANOVA of the slopes across % correct shows no significant effect for either subject in any condition (largest probability is 0.43).
Subject | Condition | Hypothesis | $w_a$ | $w_b$ | $w_c$
--- | --- | --- | --- | --- | ---
PS | In Plane | Optimal weights | 0.65 | 1 | 0.65
| | Adaptable weights | 0.74 | 1 | 0.74
| | Fixed weights | 1.15 | 1 | 0.33
| Off Plane | Optimal weights | 1 | 0.61 | 0.61
| | Adaptable weights | 1 | 0.58 | 0.58
| | Fixed weights | 1.15 | 1 | 0.33
ML | In Plane | Optimal weights | 0.63 | 1 | 0.63
| | Adaptable weights | 0.52 | 1 | 0.52
| | Fixed weights | 0.86 | 1 | 0.11
| Off Plane | Optimal weights | 1 | 0.61 | 0.61
| | Adaptable weights | 1 | 0.67 | 0.67
| | Fixed weights | 0.86 | 1 | 0.11

Table 4: Estimates of the relative weights across the bands in the plane under two different hypotheses, the weights are fixed across the In Plane and Asymmetric conditions or the weights are adaptable to the demands of each condition. The weight estimates are presented below the optimal weights for the condition (see Appendix B). $w_a$ and $w_c$ correspond to the bands in the plane oriented rightward and leftward of the direction of motion (downward) respectively, while $w_b$ corresponds to the band whose orientation coincides with the movement direction. Assuming that both the weights are adaptable and equal weighting of bands within the plaids leads to weight estimates which are not significantly different from the optimal weights. Assuming the weights are fixed across the conditions leads to a strong weight bias toward leftward moving bands, which we conclude is less likely. See text for details.

4 Discussion and Conclusions

We have shown that planar configurations of power are additively pooled. This result constitutes the strongest evidence we are aware of for the existence of planar power detectors. The three conditions together verify that planar configurations are special, while the additive law suggests that the detectors are indeed power detectors. In terms of contrast, additivity in power (or energy) is a quadratic summation rule. It is possible that the results are actually due to a nonadditive summation rule acting on contrast which has been subjected to a non-linearity other than squaring, such that the conjunction of the summation rule and the non-linearity conspire to produce the apparent additivity. However, the exact cancellation that this would require seems implausible. Regardless, the net result is that configurations of power lying on a common plane obey a summation rule which is equivalent to the rule used by planar power detectors, while configurations of power which do not lie on a common plane are detected using a suboptimal rule.

For the Off Plane condition, the exponents indicated subadditivity. A simple hypothesis is that the plaid and bandpass stimuli are independently processed and pooled by probability summation.

4.1 Context sensitive weighting across the plane

The results suggested that observers are adapting their weighting strategy across bands which lie in a common plane to the demands of a task. These adaptations are modest, involving only 25-30% changes
in the values of the weights, and hence do not contradict the conclusion that the visual system has a limited capacity to adapt the weights across bands. One way to account for the adaptability of weighting is to postulate the existence of a large set of fixed planar power detectors, each with different orientation sensitivities. The adaptability of performance could then be attributed to the observer relying on detectors with different spatial sensitivities in the two conditions. This possibility leads to the confounding of spatial structure and velocity within a detector, which must then be disambiguated by making separate estimates of the spatial structure to avoid image structure dependent biases in the encoded velocities.

Another way to account for adaptability is to postulate planar power detectors in which the weighting across spatial orientation can be modified by perceptual learning. This possibility would be a way of implementing optimal velocity estimators for the stochastic stimuli. Optimal velocity estimation can be implemented using a population of detectors tuned to different planes in frequency space. The spectrum around which the detector pools will be given by the expected spatial frequency spectrum projected unto the plane specified by the velocity. If the visual system uses takes this expectation over a window of time on the order of the experiments, then the adaptability could be accounted for by an attempt at optimality (i.e. the development of perceptual expertise).

Neurally, this possibility could involve the feedback from some site which modifies the weighting based on the pattern of correct and incorrect decisions. The idea that the visual mechanisms are flexible and can be modified by perceptual learning has been recently suggested by several researchers as a means of accounting for the stimulus and task specificity of several kinds of perceptual expertise (e.g. spatial hyperacuity)[7, 1, 8, 10, 12]. Physiological evidence for receptive field structure being modified by feedback also exists. Recently, McLean & Palmer[18] have shown that the phase selectivity of neurons in primary visual cortex can be modified by associative learning, and were able to change the direction selectivity in one cell in ten. In audition, Weinberger and colleagues [9, 36, 11] have presented evidence that the peak frequency of auditory neurons in the guinea pig and cat could be modified by classical conditioning. In visual area MT in monkeys, Treue (1996)[34] showed that neuronal firing rates could be modulated by attention, suggesting that MT neurons have the potential to be modified by feedback from higher stages of visual processing.

In addition to the difference in On Plane and Asymmetric weighting, we found evidence that the same plaid stimuli are processed differently in the In Plane and Off Plane conditions. This difference cannot be explained by an instantaneous interaction between the plaid and the bandpass stimuli, since the trials we are considering did not include the bandpass stimuli. Due to the subadditive combination in the Off Plane condition, we might expect some elevation of the plaid thresholds over the In Plane condition. For instance, the bandpass stimuli in the Off Plane condition may increase the number of irrelevant detectors the subject attends to in detecting the plaid stimuli (increased signal uncertainty)[24], or it might cause the visual system to use a detector which is more poorly matched to the plaid signal than in the In Plane condition. The latter possibility is less plausible given the fact that plaid thresholds are actually lower for smaller %correct values in the Off Plane condition. Instead, the lower thresholds suggest that subjects may actually be using detectors which are better matched to the plaid stimuli in the Off Plane condition, but are attending to many irrelevant detectors. This possibility could produce the pattern of results, because at low signal values the irrelevant detector responses will increase the decision variable variance less than the positive weighting of the middle band would in the In Plane condition. We plan to investigate the possibility that the frequency weighting for plaid stimuli differs in the two conditions using the perturbation analysis from last chapter.
4.2 Relations to physiology

There are a number of studies which suggest that the most probable location of planar power detectors in the brain is in the human analogue to simian visual cortical area MT. Simoncelli and Heeger (1998) [30] have recently modeled a great deal of electrophysiological data recorded from this area with a modified planar power detector model. A subset of the neurons in MT, designated either 'pattern motion selective'[20] or 'Type II'[2, 28] have several of the properties expected for planar power detectors. This set of MT neurons are tuned to speed and direction [17, 27, 16] and are relatively insensitive to the spatial pattern characteristics[20, 3]. Movshon et al. has also shown that the spatial and temporal frequency tuning in some of these cells covary, which is required for pattern invariant speed tuning[21]. Lesion studies show that MT neurons are critically involved in both the computation of direction of motion[22] and speed [23], while electrical stimulation can produce directional biases in a perceptual task[29].

In addition, motion opponency for MT neurons similar to that found in the perturbation analysis has been reported by several investigators[19, 27, 31, 26, 38, 39]. The opponency is manifested as a suppression of neuronal firing rates for motions in the direction opposite to the cell’s preferred direction. This inhibition has a component coextensive with the classical receptive, best demonstrated by experiments of Qian and Anderson [26], who showed that superimposing random dot patterns suppressed firing rates, especially when the opposite moving dots were locally paired in space. Snowden (1991) showed that this inhibition is well modeled by a divisive interaction and Mikami et al.[19] showed that the inhibition can be maximal for opponent motions which have speeds different than the cells preferred speed. Finally, several researchers have shown that there is a strong inhibitory interaction for motions outside of the cells classical receptive field[19], and that this inhibition is structured but spatially inhomogeneous[38, 39, 37].

The existence of spatially inhomogeneous interactions outside the classical receptive has been reported by many researchers[4, 32, 5, 38, 39]. On the basis of these findings all of these researchers have suggested that area MT may be involved in computations more complicated than simple image velocity estimation, such as figure/ground segmentation, motion in depth, and computing surface curvature and orientation in depth[37]. In addition, it has been shown that MT neurons are also tuned to binocular disparity[17], which has led to the speculation of a role for MT in motion in depth computations [4]. Finally, there are suggestions that MT neurons may play a role in determining targets for tracking eye movements [13]. These additional properties of MT neurons could be used to better assess experimentally whether the psychophysically defined planar power detectors have their basis in visual area MT.

4.3 Model for velocity estimation

The preceding experiments have suggested the outlines of the properties of velocity detectors. However, the visual system is more often faced with the problem of estimating image velocities than detecting image movement. In this section we discuss how the results fit into a velocity estimation scheme, and outline some possibilities for further testing.

Velocity estimation is a simple extension of translation detection. We propose a model based on the experimental evidence and a set of assumptions motivated by optimal velocity estimation in conditions of uncertainty. We assume: 1) The visual system makes local estimates of velocity. 2) The visual system makes use of the expected spatial structure of the signal in making velocity estimates.

An interesting property of the translation detectors we have been discussing is that a population of these detectors can be used to construct a likelihood function for image velocity. Thus the detectors could be part of a general system for estimating local image velocity. The visual system can compute optimal local estimates of velocity by using the expected spatial structure of the image, and the prior distribution of image velocities.
Let \( \hat{S}_e(\omega_x, \omega_y) \) denote the expected spatial spectrum of the signal, and \( W(\omega_x, \omega_y, \omega_t) \) represent the spectrum of the localization window. Then if the texture is moving with velocity \( \vec{v} \), the expected signal spectrum is given by:

\[
S(\vec{\omega}, \vec{v}) = \left( \hat{S}_e(\omega_x, \omega_y) \delta(\omega_t + \vec{v} \cdot \vec{\omega}_{sp}) \right) \otimes W(\omega_x, \omega_y, \omega_t)
\] (9)

This equation represents the expected signal spectrum projected onto the plane given by the velocity and blurred (convolved) by the localization function. Given \( S(\vec{\omega}, \vec{v}) \) we can construct a family of detectors tuned to different values of \( \vec{v} \). The likelihood function over velocity can be approximated by computing the inner product of the input signal power spectrum \( X(\vec{\omega}) \) with a set of detectors each with a different fixed values of \( \vec{v} \).

\[
\log L(X|\vec{v}_i) \simeq \int |S(\vec{\omega}, \vec{v}_i)|^2 |X(\vec{\omega})|^2 d\vec{\omega}
\] (10)

where we have dropped out the scaling factors for simplicity, and used the small signal approximation.

A maximum (or mean) a posteriori estimate of the velocity can be made from this sampled likelihood function by introducing a prior distribution on velocity.

\[
p(\vec{v}|X) = \frac{L(X|\vec{v})p(\vec{v})}{\int L(X|\vec{v})p(\vec{v}) d\vec{v}}
\] (11)

A likely prior on velocity would be a bias for slower speeds. As pointed out by Simoncelli (1993), a bias for slower speeds could explain several perceptual phenomena, including the wagon wheel effect and the fact that one-dimensional signals are typically seen to move in the direction orthogonal to their spatial orientations.

One of the interesting predictions of this sort of model is that speed and direction discrimination based on these filters should obey a Weber’s law in speed, which has been previously shown[6].

5 Appendix

The planar power detector additively pools power lying on a common plane to measure the spectral energy \( E \) around the plane. If we split the plane into several bands, then the planar power detector can be described as adding the output energies within each of the bands.

\[
E = \int |P(\vec{\omega})|^2 |S(\vec{\omega})|^2 d\vec{\omega}
\] (12)

\[
E = \int |P(\vec{\omega})|^2 \sum_i |S_{b_i}(\vec{\omega})|^2 d\vec{\omega}
\]

\[
E = \sum_i \int |P(\vec{\omega})|^2 |S_{b_i}(\vec{\omega})|^2 d\vec{\omega}
\]

\[
E = \sum_i w_i E_{b_i}
\]

Where \( E \) is the output energy, \( |P(\vec{\omega})|^2 \) is the planar power detector spectrum, \( |S(\vec{\omega})|^2 \) is the signal spectrum, and \( |S_{b_i}(\vec{\omega})|^2 \) are the band pass components of the signal spectrum. The weights \( w_i \) represent the effect the Planar filter has on the energy within each signal band \( i \).

Next we show that the performance of the ideal observer has the form

\[
p(R_i = 1) = \Psi(\sum_i w_i E_{b_i})
\] (13)
If we fix \( p(R_i = 1) = p_0 \), then \( \sum_i w_i E_{b_i} \) must also be a constant \( c \). Thus the equal performance contours will lie on hyperplanes in the space of energies within the signal bands, \( E_{b_i} \).

In Chapter 1 we gave an expression for the ideal performance,

\[
p(X_i = 1) = \Phi(0, \mu_{H1} - \mu_{H0}, \sigma_{H1}^2 + \sigma_{H0}^2)
\]

where the mean and variance depend on the signal and receiver filter spectra, and the signal and background noise power levels:

\[
\mu = \mu_{H1} - \mu_{H0} = 2s \int_{\omega} |P(\omega)|^2 |S_n(\omega)|^2 d\omega
\]

\[
\sigma^2 = \sigma_{H1}^2 + \sigma_{H0}^2 = 8(s^2 \int_{\omega} |P(\omega)|^4 |S_n(\omega)|^4 d\omega + 2sN \int_{\omega} |P(\omega)|^4 |S_n(\omega)|^2 d\omega + 2N^2 \int_{\omega} |P(\omega)|^4 d\omega)
\]

where \( s|S_n(\omega)|^2 = |S(\omega)|^2 \), i.e. \( |S_n(\omega)|^2 \) is the normalized signal spectrum and \( s \) is the signal power level. The mean is simply proportional to the energy, and hence is linear in the energies in the bands. The variance can be shown to be linear to an extremely good approximation. In the experiments, \( s \ll N \) and hence the first term in the variance can safely be dropped. \( \int_{\omega} |P(\omega)|^4 d\omega = k_p \) evaluates to a constant \( k_p \).

To evaluate the middle term we notice that when \( |S_n(\omega)|^2 = \sum_i |S_{b_i}(\omega)|^2 \), the term \( \int_{\omega} |P(\omega)|^4 |S_n(\omega)|^2 d\omega \) evaluates to \( \sum i a_i w_i E_{b_i} \), i.e. weighting the signal spectrum by the square of the planar filter spectrum simply scales the energy by a fixed amount \( a_i \).

Thus the mean and variance can be written:

\[
\mu = 2 \sum_i w_i E_{b_i}
\]

\[
\sigma^2 = 16(\sum_i a_i w_i E_{b_i} N + k_p N^2)
\]

When the filter overlap on each of the signal bands is identical, the inner products \( \int_{\omega} |P(\omega)|^4 |S_{b_i}(\omega)|^2 d\omega \) are the same for all the bands, and hence all the \( a_i = a \) are the same. This is true for two of the conditions in the experiment, the In Plane condition and the Asymmetric condition.

Thus for the conditions in this experiment, the ideal performance can be written:

\[
p(X_i = 1) = \Phi \left( 0, 2 \sum_i w_i E_{b_i}, 16(a \sum_i w_i E_{b_i} N + k_p N^2) \right)
\]

For fixed \( N \) and any fixed probability correct \( p(X_i = 1) \), the ideal observer’s performance is linear in the energies.

It is important to point out that the condition of identical overlap between signal bands is not very important, since the third term dominates the variance expression.

The previous discussion can be easily adapted to the human observer. We assume that the visual system uses an unknown internal filter which is roughly selective for planar regions of spectral power. In addition, the visual system is subject to additional sensory and internal noise. If we assume that the additional noise is equally distributed across all the bands, then the subject’s decision variable variance should be dominated by the effects of the background noise and the sensory and internal noises. In this case both of the signal dependent terms in the variance will be nearly insignificant, and the subject’s performance should be approximately linear in the energies in the signal bands.
6 Appendix B: Ideal observers for the task

The stimuli in the task can be described by the equations:

\[ H_1 = \text{signal present:} \quad r(x, y, t) = a_1 \cdot s_{pl}(x, y, t) + a_2 \cdot s_b(x, y, t) + n(x, y, t) \]

\[ H_0 = \text{noise alone:} \quad r(x, y, t) = n(x, y, t) \]

where \( s_{pl} \) denotes the plaid signal waveform and \( s_b \) denotes the bandpass signal waveform. The constants \( a_1 \) and \( a_2 \) determine the contrast of the signal noises \( s \), hence \( a_1^2 \) and \( a_2^2 \) are proportional to the signal energies. We will compute the ideal for the case in which the signal energies are much lower than the background noise energy. In the equations, \( a_1 \) and \( a_2 \) form a random vector since the data are collected by randomly selecting from a set of \([a_1, a_2]\) pairs.

The Bayes decision for the 2AFC task is to choose the interval \( i \) with the larger likelihood ratio \( L(r) \), averaged over the \( \bar{a} = [a_1, a_2] \) pairs:

\[
E[L(r|\bar{a})_1] \geq E[L(r|\bar{a})_2]
\]

(18)

where the likelihood ratio is the ratio of the conditional probabilities of the waveform \( r \) given signal present and noise alone conditions:

\[
L(r|\bar{a}) = \frac{p(r|H_1, \bar{a})}{p(r|H_0, \bar{a})}
\]

(19)

Let \(|S_{pl}(\omega_i)|^2\) and \(|S_{pl}(\omega_i)|^2\) denote the plaid signal and bandpass signal normalized power spectra respectively, and let \( \bar{S}(\omega_i) = \langle |S_{pl}(\omega_i)|^2, |S_{pl}(\omega_i)|^2 \rangle \).

The likelihood ratio is given by:

\[
\Lambda(R|\bar{a}) = \frac{\prod_{i=1}^{M} \left[ \frac{1}{2\pi a^2 \cdot \bar{S}(\omega_i)} + N \right]^{0.5} \exp \left( -0.5 \sum_{i=1}^{M} \frac{R_i R_i^*}{a^2 \cdot \bar{S}(\omega_i) + N} \right)}{\prod_{i=1}^{M} \left[ \frac{1}{2\pi N} \right]^{0.5} \exp \left( -0.5 \sum_{i=1}^{M} \frac{R_i R_i^*}{N} \right)}
\]

(20)

\[
\Lambda(R|\bar{a}) = \left( \prod_{i=1}^{M} \left( \frac{N^{0.5}}{a^2 \cdot \bar{S}(\omega_i) + N} \right) \right) \exp \left( -0.5 \sum_{i=1}^{M} \frac{\left( \bar{a}^2 \cdot \bar{S}(\omega_i) \right) R_i R_i^*}{a^2 \cdot \bar{S}(\omega_i) + N} \right)
\]

When the background noise power is much greater than the total signal power the likelihood ratio reduces to:

\[
\Lambda(R|\bar{a}) = \exp \left( -\frac{0.5}{N^2} \sum_{i=1}^{M} \left( \bar{a}^2 \cdot \bar{S}(\omega_i) \right) R_i R_i^* \right)
\]

(21)

The likelihood averaged across \( \bar{a}^2 \) is given by:

\[
E[L(r|\bar{a})] = \sum_{\bar{a}} p(\bar{a}) \Lambda(R|\bar{a})
\]

(22)

which follows because \( p(\bar{a}) = p(\bar{a}^2) \), due to the fact that \( \bar{a} \) is constrained to be positive. This expression does not simplify but can be simulated.
If the visual system computes the decision based on the sum of energies in the bands, but can vary the weighting within the bands, then we can compute the optimal weights based on the set of stimulus mixtures used. We can derive the optimal additive rule by using the log likelihood functions.

\[
\log \Lambda(R|\vec{a}) = -\frac{0.5}{N^2} \sum_{i=1}^{M} \left( \vec{a}^2 \cdot \vec{S}(\vec{\omega}_i) \right) R_i R_i^* \tag{23}
\]

The expectation of the log likelihood function over $\vec{a}^2$ is given by:

\[
E[\log \Lambda(R|\vec{a})] = -\frac{0.5}{N^2} \sum_{\vec{a}} p(\vec{a}) \sum_{i=1}^{M} \left( \vec{a}^2 \cdot \vec{S}(\vec{\omega}_i) \right) R_i R_i^* d\vec{a} \tag{24}
\]

where $\vec{a}^*2$ is the mean stimulus power vector. Thus in this case the decision is to compare the energies in the filter $\vec{a}^2 \cdot \vec{S}(\vec{\omega}_i)$ on both intervals and choose the interval with the larger energy.

From the set of $(E_{pl}, E_{bp})$ energy vectors we used we can compute the expected energy and hence the expected weights on the energy bands. These in turn produce expected slopes which we can compare against the data. For subject PS the expected weights for the In Plane condition are $(a, b, c) = (1, 1.55, 1)$, while for the Asymmetric condition is $(a, b, c) = (1.63, 1, 1)$. For subject ML, the expected weights for the In Plane condition are $(a, b, c) = (1, 1.58, 1)$, and for the Asymmetric condition are $(a, b, c) = (1.65, 1.1)$.

References


