

# Fourth-root summation of contrast over area: No end in sight when spatially inhomogeneous sensitivity is compensated by a witch's hat

Alex S. Baldwin

School of Life and Health Sciences,  
Aston University, Birmingham, UK



Tim S. Meese

School of Life and Health Sciences,  
Aston University, Birmingham, UK



Measurements of area summation for luminance-modulated stimuli are typically confounded by variations in sensitivity across the retina. Recently we conducted a detailed analysis of sensitivity across the visual field (Baldwin, Meese, & Baker, 2012) and found it to be well described by a bilinear “witch’s hat” function: Sensitivity declines rapidly over the first eight cycles or so, but more gently thereafter. Here we multiplied luminance-modulated stimuli (4 cycles/degree gratings and “Swiss cheeses”) by the inverse of the witch’s hat function to compensate for the inhomogeneity. This revealed summation functions that were straight lines (on double log axes) with a slope of  $-1/4$  extending to  $\geq 33$  cycles, demonstrating fourth-root summation of contrast over a wider area than has previously been reported for the central retina. Fourth-root summation is typically attributed to probability summation, but recent studies have rejected that interpretation in favor of a noisy energy model that performs local square-law transduction of the signal, adds noise at each location of the target, and then sums over signal area. Modeling shows our results to be consistent with a wide field application of such a contrast integrator. We reject a probability summation model, a quadratic model, and a matched template model of our results under the assumptions of signal detection theory. We also reject the high threshold theory of contrast detection under the assumption of probability summation over area.

## Introduction

As the area of a sine-wave grating increases, it becomes easier to detect (Hoekstra, van der Goot, van den Brink, & Bilsen, 1974; Savoy & McCann, 1975). For a patch of grating presented in the center of the visual field, the function that plots threshold against area (on log–log axes) is curved, being initially steep and then

shallower, such that there is only marginal benefit from increasing the diameter of the grating beyond eight cycles or so (Robson & Graham, 1981; Tootle & Berkley, 1983; Rovamo, Luntinen, & Näsänen, 1993). There are several processes that contribute to the shape of this function. The steep initial improvement is thought to be due to linear summation within spatial filter elements (Meese, 2010). The further improvement beyond this point has traditionally been attributed to probability summation over local filter elements (e.g., Robson & Graham, 1981). The curvature towards an asymptote is explained by inhomogeneous sensitivity across the visual field, where contrast sensitivity declines with eccentricity (Howell & Hess, 1978; Foley, Varadharajan, Koh, & Farias, 2007). Baldwin, Meese, and Baker (2012) ruled out an explanation of this inhomogeneity as being due to receptor density, but Bradley, Abrams, and Geisler (2014) have demonstrated that an account in terms of retinal ganglion cell density is plausible. In the absence of within-filter summation and visual field inhomogeneity (and under the assumptions described later), probability summation would produce a log–log summation slope of about  $-1/4$  (consistent with the intermediate part of the empirical summation slope; e.g., Meese, Hess & Williams, 2005) and for this reason is sometimes referred to as fourth-root summation.

Recent work has provided a serious challenge to the probability summation interpretation of the fourth-root summation rule. Studies involving classification images (Baker et al., 2014), interdigitated micropattern stimuli known as Battenbergs (Meese, 2010), plaid-modulated grating stimuli known as Swiss cheeses (Meese & Summers, 2007, 2009; Baker & Meese, 2011; Meese & Baker, 2011), and measurement of the psychometric slope under various conditions of extrinsic uncertainty (Meese & Summers, 2012) have all

Citation: Baldwin, A. S., & Meese, T. S. (2015). Fourth-root summation of contrast over area: No end in sight when spatially inhomogeneous sensitivity is compensated by a witch’s hat. *Journal of Vision*, 15(15):4, 1–12, doi:10.1167/15.15.4.

concluded that the probability summation account is wrong (see also Foley et al., 2007; Morgenstern & Elder, 2012). Instead, the preferred model for area summation is one in which local contrasts are extracted by local filter elements (analogous to simple-cell receptive fields), followed by square-law contrast transduction, the addition of independent Gaussian noise to the output of each transducer, and global summation over the stimulus region (Meese & Summers, 2012). In this “noisy energy” model, the cascade of nonlinear transduction and ideal summation of inputs (combining signal and noise) results in a fourth-root summation rule; our view is that this has masqueraded as probability summation (often with a tacit assumption of a linear transducer; see Tyler and Chen, 2000), explaining why the probability summation model has held sway for so long.

One complication with conventional experiments involving centrally placed patches of grating is that retinal inhomogeneity means it can be difficult to assess the spatial extent of the summation process. Recently we made a detailed measurement of contrast sensitivity across the central visual field and found it to be well-described by a witch’s hat function where sensitivity declines rapidly over the first eight cycles or so and more gently thereafter (Baldwin et al., 2012). This function was then built into the noisy energy model, producing very good parameter-free predictions of experimental results when overall sensitivity was normalized (e.g., Meese & Summers, 2012). However, this does not overcome the measurement problem outlined above, and it remains unclear how far the summation process extends. Here we take a different approach. Instead of building the witch’s hat into the model, we constructed a weighting function from its inverse and used this to compensate for the inhomogeneous sensitivity by multiplying it with stimuli (gratings and Swiss cheeses). Essentially, this applied the conventional normalization technique that is commonplace in two-component subthreshold summation experiments. By doing so we found empirical fourth-root summation functions that extended to a stimulus diameter of at least 33 cycles (the largest size we tested). We show that the noisy energy model provides an excellent prediction of these results (with a single free parameter to control overall sensitivity) and, once again, we reject accounts in terms of probability summation.

## Methods

### Equipment

Stimuli were stored in a CRS ViSaGe (Rochester, Kent, UK) and presented on a gamma-corrected CRT

monitor (Eizo Flexscan T68, Bracknell, Berkshire, UK) with a 14-bit gray-level resolution. The monitor had a refresh rate of 120 Hz and a mean luminance of 75 cd/m<sup>2</sup>. It was viewed from a distance of 1.19 m, having a resolution of 48 pixels per degree of visual angle (12 pixels/cycle for the 4 cycles/degree stimuli used here).

### Observers

Data were collected from three observers: ASB, DHB, and TSM. The observers were 22, 28, and 46 years old, respectively. All three were psychophysically experienced (ASB and TSM are authors). Optical correction appropriate for the viewing distances tested was worn if required. All experiments were performed binocularly with natural pupils.

### Stimuli

Two types of stimuli were used: circular 4 cycles/degree sine-phase horizontal gratings (Figure 1a, b), and Swiss cheese modulated versions of those gratings (Figure 1c through f; after Meese & Summers, 2007). Stimuli were windowed by a raised-cosine envelope, which declined from unity to zero over a distance of 12 pixels. The nominal stimulus diameters that we report are the full-widths at half magnitude (i.e., 12 pixels wider than the diameter of the plateau due to the raised-cosine skirt of the envelope). The Swiss cheese modulations had a spatial frequency of 0.8 cycles/degree and were applied in cosine ( $\phi = 90^\circ$ ) and anticostine ( $\phi = 270^\circ$ ) phases (Figure 1c and e). The centers of the stimuli were at the patterns’ maximum and minimum contrasts for cosine and anticostine phase modulations, respectively. Eight sizes were used for the gratings (1.3 to 33.0 cycles in diameter), the larger five of which were also used for the Swiss cheeses. We express stimulus contrast in dB (re 1%), given by  $20 \times \log_{10}(c)$ , where  $c$  is Michelson contrast in percent. For convenience, in the graphical presentation of our results we also express stimulus area as 20 times the  $\log_{10}$  of the nominal stimulus diameter squared, relative to the smallest stimulus area. Note also that our error measures were derived by calculating the root-mean-square (RMS) errors between empirical thresholds and model predictions, where each is expressed in dB.

Stimuli were presented with both (a) flat (uncompensated) contrast profiles, and (b) witch’s hat compensation for the inhomogeneous retinal field. The compensated stimuli were multiplied by the inverse of the attenuation surface measured for each observer, as reported in table 4 of Baldwin et al. (2012; e.g., Figure 1b, d, and f). This was to counteract the effects of inhomogeneous sensitivity with the aim of producing

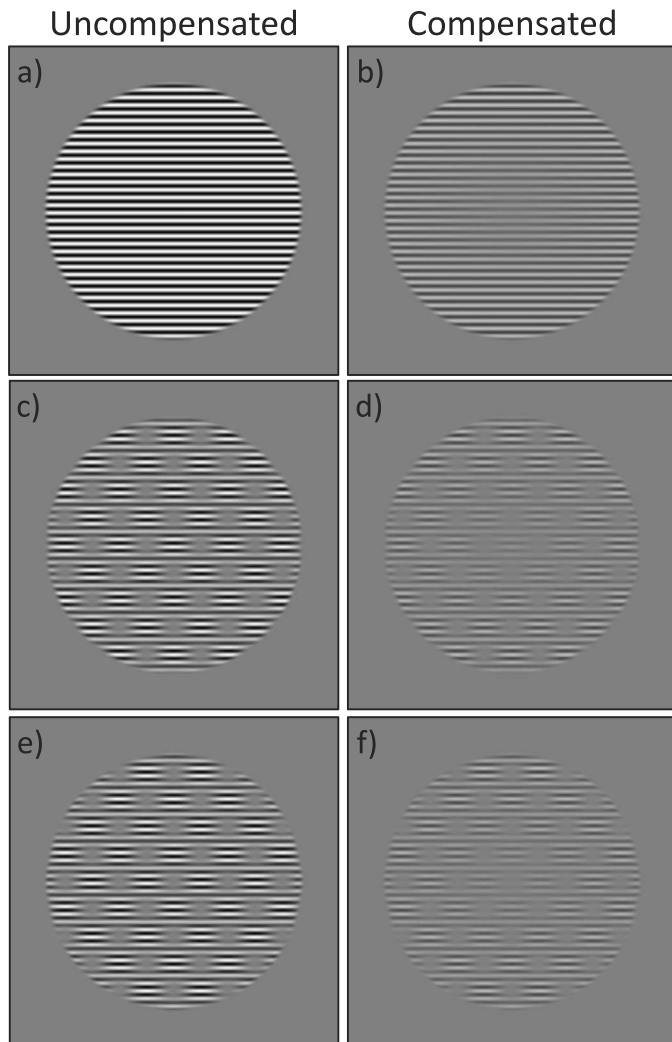


Figure 1. Stimulus examples (the stimuli shown are the largest we used). Stimuli in the left and right columns are uncompensated and compensated, respectively, by multiplication with a witch’s hat. The first row (a–b) shows gratings (sometimes called full stimuli). The second (c–d) and third (e–f) rows show Swiss cheeses in the cosine ( $\phi = 90^\circ$ ) and antic cosine ( $\phi = 270^\circ$ ) phases, respectively.

an effectively flat contrast response profile over area at the summation stage in the visual system. The nominal contrasts of our stimuli are the contrasts of their grating carriers before modulation (by a Swiss cheese) and/or compensation (by a witch’s hat), as appropriate.

A quad of fixation points (black  $2 \times 2$  pixel squares) were placed snugly around each stimulus (Summers & Meese, 2009; i.e., the virtual edge of the quad was matched to the full diameter of the stimulus). Observers were able to use these markers to infer stimulus size and the central location of the display, where they fixated (Meese & Summers, 2012). The stimulus duration was 100 ms.

## Procedures

Thresholds were measured using a two-interval forced-choice technique with auditory feedback on the correctness of the observer’s response (beeps of different tones) provided after each trial. A pair of three-down one-up staircases were used to control the stimulus contrast for each condition using a step size of 3 dB. Each staircase terminated after either 70 recorded trials or 12 staircase reversals, whichever occurred first. For each staircase, recording began after the first two reversals, where the step sizes were 12 and 6 dB, respectively. The observers repeated each condition four times, except for TSM in the Swiss cheese condition where only two replications were performed. Stimuli were blocked by size and by stimulus type (whether they were gratings or Swiss cheeses). Compensated and noncompensated versions of each stimulus were interleaved within a block. For the Swiss cheese stimuli, the two modulator phase conditions were also interleaved within a block. (For empirical and theoretical comparisons between blocked and interleaved designs, see Meese & Summers, 2012.)

Contrast detection thresholds and slopes of the psychometric functions were calculated by fitting a Weibull function to the percent-correct data (collapsed across staircase and repetition) using Palamedes (Prins & Kingdom, 2009). Note that this “pool then fit” approach has been shown to provide slightly more accurate estimates of the slope of the psychometric function than the alternative “fit then pool” approach, in which Weibull functions are fit to the data from each session and then averaged (Wallis, Baker, Meese, & Georgeson, 2013). Parametric bootstrapping was then performed such that the threshold and psychometric slope values reported are the median values from the bootstrap population averaged across observer (1,000 samples per threshold per observer), with 95% confidence intervals (typically smaller than symbol size).

The study was performed under the tenets of the declaration of Helsinki.

## Results and discussion

The results for the full grating stimuli are shown in Figure 2 for the stimuli without witch’s hat compensation (Figure 2a) and with compensation (Figure 2b). If summation were linear then thresholds would decline with a slope equal to that of the steepest dashed line (–1). The shallower dashed lines represent square-root (–1/2) and fourth-root (–1/4) summation. The results without compensation have the typical bowed appearance, replicating previous studies (e.g., Rovamo et al., 1993; Meese & Summers, 2007, 2012). However, with

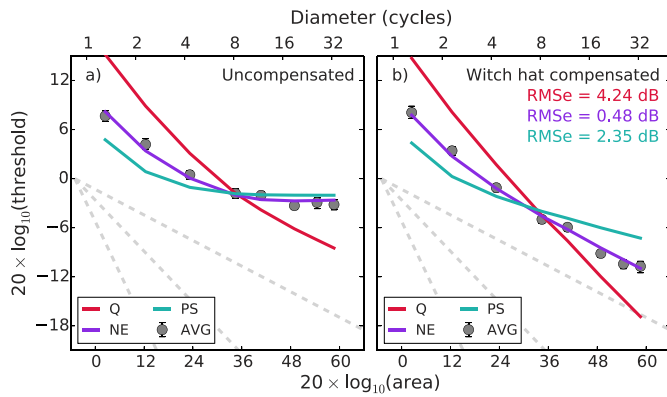


Figure 2. Thresholds for full circular gratings, averaged across three observers and plotted against area (solid symbols). Panel (a) is for uncompensated gratings and panel (b) is for compensated gratings. The dashed gray lines show slopes of  $-1$ ,  $-1/2$ , and  $-1/4$ . Error bars here and in other figures show 95% confidence intervals, often smaller than symbol size. The colored curves are fits for the quadratic (Q), noisy energy (NE), and probability summation (PS) models, each with a single free parameter to control vertical offset in the plot. Note that the RMS errors were calculated across left and right panels. The average standard error across observer thresholds (after normalizing for overall sensitivity) was 0.65 dB.

the compensation in place, the bowing of the function is much less pronounced. Comparing the data with the fiducial contours (gray dashed lines), it starts with a slope of around  $-1/2$  and rapidly diminishes to a slope of  $-1/4$  where it remains. It would seem that our attempt to compensate for the effects of sensitivity loss with eccentricity was successful because performance continues to improve over a much greater range of the central visual field than has been seen previously.

The results for the Swiss cheese stimuli are shown in Figure 3, where the same data are shown in each row (note the reduced range of both axes compared to Figure 2). For these stimuli there was very little benefit from increasing stimulus diameter for the uncompensated stimuli (left column), consistent with previous observations (Meese & Summers, 2007). In our preferred model this is because the benefits of the extra signal are weak (coming from increasingly peripheral retina) and largely offset by the detrimental contribution of further noise. This effect is not specific to the Swiss cheese stimuli but is also seen in the full grating stimuli for the same stimulus diameters, as a comparison with the pale gray symbols (replotted from Figure 2) confirms. However when witch’s hat compensation is introduced (right column) sensitivity once again improves with a summation slope of about  $-1/4$  over the full stimulus range tested (e.g., compare data symbols with the shallowest dashed gray line in Figure 3b).

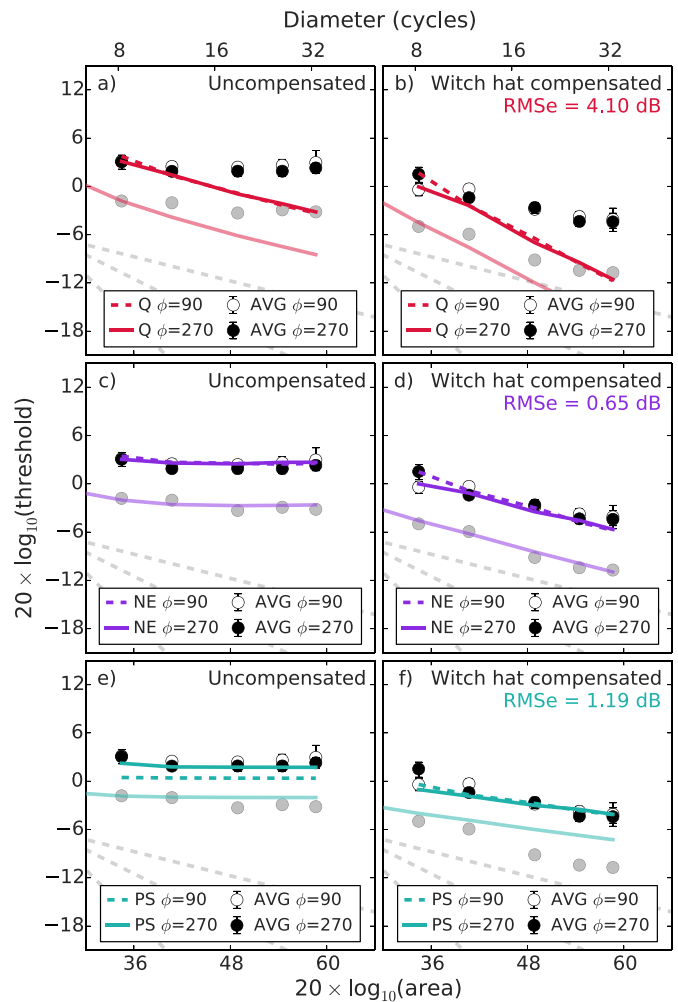


Figure 3. Thresholds for the Swiss cheese stimuli, averaged across observers and plotted against area for two modulation phases (white and black symbols). The left and right columns are for uncompensated and compensated stimuli, respectively. Different rows are for the same human data, but different models. The continuous colored curves are predictions (no free parameters) for the quadratic (Q) model (top row), the noisy energy (NE) model (middle row), and a probability summation (PS) model (bottom row). The dashed gray lines in each panel show slopes of  $-1$ ,  $-1/2$ , and  $-1/4$ . The right-hand part of the results and model curves from Figure 2 are replotted here with reduced opacity. Note that the RMS errors were calculated across left and right panels. The average standard error across observer thresholds (after normalizing for overall sensitivity) was 0.63 dB.

### Slopes of the psychometric functions

To supplement our model analysis we also report the slopes of the psychometric functions (Weibull  $\beta$ ). Consistent with other studies (Mayer & Tyler, 1986; Meese & Summers, 2012; Wallis et al., 2013), psychometric slope did not vary systematically over different stimulus sizes (not shown). Combining data across the

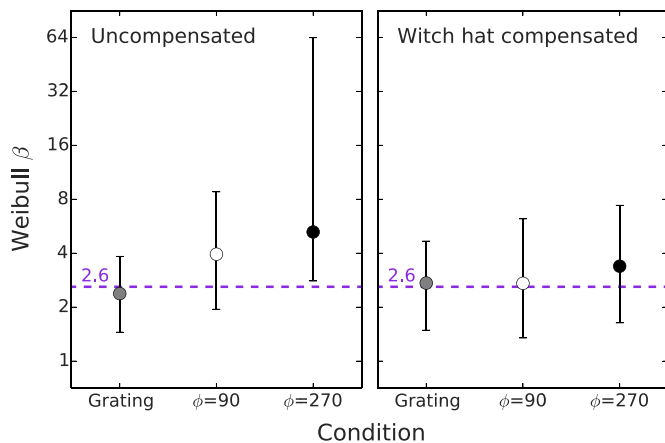


Figure 4. Slopes of the psychometric functions (the Weibull  $\beta$  parameter) for the six stimulus conditions (Figure 1) collapsed across area and observer. The purple dashed line at  $\beta = 2.6$  is the noisy energy model prediction (it is also the quadratic model prediction). The linear transducer in the ideal matched template model predicts  $\beta = 1.3$ . Under the assumptions of HTT, probability summation predicts  $\beta = 4$  (derived from the fourth-root form of the empirical summation slope). One interpretation of fourth-root summation supposes a transducer exponent of 4, predicting  $\beta = 5.2$  (see modeling text for details).

three observers gave the slopes for each condition shown in Figure 4, with an overall median  $\beta$  of 2.8 (the individual observer values were 1.8, 3.2, and 3.0 for ASB, DHB, and TSM, respectively). Previous work by Robson and Graham (1981), Mayer and Tyler (1986), and Meese and Summers (2012) found average slopes of 3.5, 3.5, and 3.6, which are a little steeper than the slopes from the study here. Our values agree with those measured by Wallis et al. (2013), who also report slopes of 2.8. Previous work has shown that the psychometric function is essentially stationary for practiced observers and that slightly better estimates of the slope are achieved when the results are collapsed across multiple sessions before curve fitting, as we did here (Wallis et al., 2013).

## Modeling

In Appendix A we present the mathematical development of four different models of area summation. In all models the stimulus was first multiplied by the witch’s hat to simulate the inhomogeneous contrast sensitivity (for the compensated stimuli this transformed it back to the original stimulus). The stimulus was then filtered with horizontal sine- and cosine-phase Cartesian-separable log-Gabor filters, with a spatial frequency bandwidth of 1.6 octaves and an orientation bandwidth of  $\pm 25^\circ$  (Meese, 2010). A single filtered image was constructed from the sum of the sine- and

cosine-phase filters. This spatial filtering is needed to capture the initial steepness of the summation slope (Meese & Summers, 2012). Note that in our model, the sampling density of the filters was matched to that of the image. Within reason, this simplification has no material impact on our conclusions.

Models were fitted to the full grating results by minimizing the RMS error (in dB) with a single (and uninteresting) free parameter that determined overall sensitivity (i.e., it was an offset parameter that slid the model curves up and down the plots to find the best fit). This overall sensitivity was used to produce the predictions for the Swiss cheese stimuli with no further parameters.

## Noisy energy model

This is our favored model developed in previous work (Meese, 2010; Meese & Summers, 2012). Each pixel in the filtered image is subject to square-law transduction and additive Gaussian noise, followed by summation over the stimulus area. This was implemented by weighting the stimulus with a template derived from the envelope of the full stimulus for that block. A feature of this model is that the summation template (i.e., the weight of the contributions of signal and noise in the summing device) is matched to the stimulus region, over which it is uniform (i.e., it is not matched to the local contrast modulations; Meese & Summers, 2007). For Swiss cheese stimuli the contribution from the “hole” regions in the cheese is dominated by local noise since the signal levels there are so low.

The model derives its name (Meese & Summers, 2012) from the fact that its summation characteristics depend on (a) square-law transduction of local contrast (identical to the energy model), and (b) the dependency of internal noise on stimulus area, as does the ideal summation model. (For this reason, Meese, 2010, referred to the noisy energy model as the “combination model” because it combined the characteristics of two of the competing models outlined in the exposition of that work.) The noisy energy model predicts an asymptotic summation slope of  $-1/4$  for compensated stimuli.

## Quadratic model (sometimes known as the energy model)

This model (e.g., Manahilov & Simpson, 1999, 2001) features square-law contrast transduction at each point on the output of the linear filtering stage. Performance is limited by (notional) additive Gaussian noise before the decision stage, independent of stimulus area. (We have called this model the quadratic model here, to emphasize that it sums the squares of its inputs, and to

avoid confusion with different model implementations that have also been referred to as the “energy model.”) This model predicts an asymptotic summation slope of  $-1/2$  for compensated stimuli.

A similar threshold prediction would be made by a matched template model (e.g., Burgess & Ghandeharian, 1984) where the observer constructs a template matched to the stimulus profile (after witch’s hat attenuation) and cross-correlates this with the stimulus. Because the standard deviation of internal noise grows with the square root of stimulus area, this model also predicts a summation slope of  $-1/2$  for compensated stimuli. In fact for the grating stimuli, the matched template (ideal summation) model makes threshold predictions that are indistinguishable from the quadratic model (for details see Meese & Summers, 2012). Note that for the quadratic model the  $-1/2$  summation slope derives from square-law transduction, whereas for the template model it derives from ideal summation of signal and early noise.

#### **Probability summation model under signal detection theory**

Models of probability summation under signal detection theory (SDT) involve a max operation across noisy mechanisms. This is mathematically complicated (e.g., Kingdom, Baldwin, & Schmidtmann, 2015), but under several conditions can be approximated by a fourth-root summation rule using Minkowski summation with an exponent of 4 (e.g., Tyler & Chen, 2000). The unusual conditions in which the max operator implementation of probability summation can produce more summation (equivalent to an exponent of 2) have been shown to be inconsistent with experiments of the type here (Meese & Summers, 2012). There is good evidence that the contrast transducer in human vision is nonlinear, being well approximated by a squaring exponent of 2 around threshold (Meese, 2010; Meese & Summers, 2009, 2012). The combined effects of these two nonlinearities means that probability summation can be approximated here using Minkowski summation with an exponent of 8 ( $= 4 \times 2$ ). This model predicts an asymptotic summation slope of  $-1/8$  for compensated stimuli.

#### **Fourth-root summation and probability summation under high threshold theory**

The fourth-root summation model has a long history in vision science. At one level it can be treated as a descriptive model; our results do have a  $-1/4$  slope over much of their range after all. It is also a good approximation to probability summation using a max operator when the contrast transducer is linear (Tyler & Chen, 2000). Finally, it is also the prediction for

probability summation under high threshold theory (HTT) when the slope of the psychometric function (Weibull  $\beta$ ) equals 4 (Robson & Graham, 1981), though more generally, probability summation under HTT gives a summation slope of  $1/\beta$ . Note that HTT underlies the common conception of probability summation, where it is understood that one calculates overall sensitivity by combining the probabilities of detecting the individual components using the standard statistical procedure for combining probabilities. This model predicts an asymptotic summation slope of  $-1/\beta$  for compensated stimuli.

### **Comparing the model predictions with our data**

The fits of the first three of our models are shown by the colored curves for the full grating stimuli in Figure 2 and for Swiss cheese stimuli in Figure 3. In Figure 3 the predictions for the cosine and antic cosine phase modulators ( $\phi = 90^\circ$  and  $270^\circ$ ) are shown with dashed and solid curves, respectively. The single parameter fit of the noisy energy model (purple curve in Figure 2a, b) is very good. Note how well it captures the initial steepness of the data (owing to within-filter summation) and then levels off to horizontal when there was no compensation (Figure 2a) and to a slope of  $-1/4$  when the witch’s hat stimulus compensation was in place (Figure 2b). As mentioned in the Introduction, this fourth-root behavior in the model is due to the cascading quadratic effects of square-law transduction and integration of internal noise with the signal (i.e., the internal noise at the decision variable increases with stimulus diameter). This is not seen in the uncompensated case (Figure 2a) because of the loss of sensitivity to the signal with eccentricity. The predictions (no free parameters) for the Swiss cheese stimuli (purple curves in Figure 3c, d) are also very good. Of the three models plotted in Figures 2 and 3, the RMS errors from the noisy energy model are by far the best.

The quadratic model (red curves in Figures 2 and 3a, b) fares much less well. The predicted benefit of area is too great, with the square-law transduction producing a slope of  $-1/2$  (compare with the intermediate dashed gray lines). Because the predictions for the matched-template model would be very similar to the quadratic model (see Quadratic model section above), then these results also lead us to reject the matched-template model. Similar conclusions have been drawn in a previous study (Meese & Summers, 2012) where the matched-template model was considered in more detail.

The probability summation following square-law transduction model (cyan curves in Figures 2 and 3e, f) also fails badly. It predicts far too little summation, both with increasing stimulus diameter (the summation slopes are too shallow) and also between the full

and Swiss cheese stimuli (the average difference in the human data is 5.5 dB, whereas the model predicts 3.1 dB). Indeed, the main reason that the probability summation model fares so well for the compensated Swiss cheese stimuli (Figure 3f) is because it underestimated the comparison sensitivities for the full gratings in this region of the summation curve (Figures 2b and 3f).

Perhaps not surprisingly, the fourth-root model (not shown) fared nearly as well as the noisy energy model (bearing in mind that the inclusion of spatial filtering means that the initial part of the predicted summation slope is steeper than  $-1/4$ ). It did not fare quite so well in predicting the summation between the full and Swiss cheese stimuli: An average of 3.9 dB for the fourth-root model, by comparison to 5.5 dB in the human results (the noisy energy model predicts 5.2 dB). The fourth-root summation model is largely descriptive in origin (excepting the spatial filtering, retinal inhomogeneity, and template, common to all of our models), and so we must ask what processes it is intended to summarize. As mentioned above, one interpretation is in terms of probability summation under the assumptions of HTT and linear contrast transduction (Robson & Graham, 1981; Meese & Williams, 2000). However, HTT has been discredited (e.g., Nachmias, 1981). We will return to this point when we consider the slope of the psychometric function. Another possible interpretation of a fourth-root summation slope is in terms of probability summation arising from a max operator under SDT (Tyler & Chen, 2000). However, the version of that model that predicts this slope also involves a linear transducer, which is almost certainly wrong (Meese, 2010; Meese & Summers, 2009, 2012). A third interpretation would be linear summation following a transducer with an exponent of 4 (Graham, 1989). We will return to this possibility when we consider the slope of the psychometric function below. Finally, the combined effects of a square-law transducer and the integration of signal and noise within a matched template also predict a fourth-root rule. Thus, the success of the fourth-root model can be seen as deriving from its similarity to the noisy energy model. Indeed, for the full grating stimuli, the threshold predictions by the two models are indistinguishable. However, the models can be differentiated when summation is assessed by filling in the holes of Swiss cheeses (as shown by the vertical offset between grating and Swiss cheese thresholds in Figure 3). In the noisy energy model, templates are not matched to the plaid modulations in the stimulus, just to the overall stimulus size. Thus, noise is constant for a fixed stimulus diameter so the summation effects derive from a signal exponent of 2 (the square-law transducer) and are greater than the fourth-root prediction.

## Model predictions for the slopes of the psychometric functions

The analysis above is sufficient to demonstrate that the noisy energy model provides a better account of our data than the other models we have considered. However, we can also provide a brief analysis based on the slopes of the psychometric functions (Weibull  $\beta$ ). The noisy energy model and the quadratic model both involve square-law ( $p = 2$ ) transduction of signal contrast ( $c^p$ ). In the absence of uncertainty (Pelli, 1985), this predicts  $\beta = 1.3 \times 2 = 2.6$  (Pelli, 1987; Tyler & Chen, 2000; May & Solomon, 2013), very close to the average of  $\beta = 2.8$  here, and in agreement with the slopes from some of our individual conditions (see Figure 4). It seems likely that the small deviation from the  $\beta = 2.6$  prediction arises from uncertainty (see Meese & Summers, 2009), which appears to be greatest in the uncompensated case, particularly when  $\phi = 270^\circ$  (i.e., when there was no signal contrast in the center of the visual field). Thus, it is clear that our preferred model from above (the noisy energy model) is consistent with the slopes of empirical psychometric functions.

The template-matching model has a linear transducer and so predicts a psychometric slope of  $\beta = 1.3$  (Pelli, 1987), at odds with our estimates in this study (Figure 4). In principle, this shortcoming might be overcome by supposing a fairly high level of intrinsic uncertainty across all stimulus conditions. However, other experiments in which we have assessed intrinsic uncertainty by manipulating extrinsic uncertainty suggest that high levels of uncertainty are unlikely for grating stimuli (Meese & Summers, 2012). As mentioned earlier, one interpretation of the fourth-root summation model involves a contrast transducer of  $p = 4$ ; however, this predicts  $\beta = 1.3 \times 4 = 5.2$ , which is much higher than what we found empirically (see Figure 4).

For the probability summation model under the assumptions of HTT, a psychometric slope of  $\beta = 4$  is implied by the fourth-root summation curve for detection thresholds. However, our psychometric slopes were not consistent with this prediction (Figure 4). Alternatively one might take the slope of the psychometric function to predict the slope of the summation function, but that predicts a summation function that is too steep (a slope of  $-1/2.8$ , on average) compared to the human results (Figure 2b; a slope of  $-1/2.8$  lies between the upper two pale dashed lines which have slopes of  $-1/4$  and  $-1/2$ ). This mismatch between two empirical measures (slopes of psychometric functions and summation functions) serves as further evidence to reject HTT under the widely held assumption that area summation is achieved by probability summation under that model.

## General discussion and conclusions

Using witch's hat compensation for the loss of contrast sensitivity with retinal eccentricity (Baldwin et al., 2012), we measured spatially extensive fourth-root summation of contrast (at 4 cycles/degree) across the central visual field for a stimulus diameter range of 1.3 to 33 cycles (an area factor of 644). This result was lawful over the full range tested when the minor deviations from the fourth-root rule (at small stimulus sizes) were accommodated by spatial filtering, typical of that known in central human vision (e.g., Meese, 2010). These results (along with the slopes of the psychometric functions) were not consistent with probability summation under SDT or HTT, a template-matching model or a simple quadratic model, but they were consistent with a model involving square-law contrast transduction and the integration of signal and internal noise over area (the noisy energy model; Meese, 2010).

Our results show that the summation process extends up to at least 33 stimulus cycles, possibly more. Indeed, if the process were part of a visual hierarchy involved in assessing the size and/or area of objects and textures (Meese & Baker, 2011) then we should expect signal integration to extend across the entire retinal field, since the dimensions of real-world objects and textures are not constrained by retinal image size. However, other work we have done has fallen short of this conclusion. Analysis and results of contrast detection of various Swiss cheese stimuli (see Baker & Meese, 2011), and reverse correlation analysis of suprathreshold Battenberg stimuli (Baker & Meese, 2014), suggest that contrast integration over area operates up to only about 12 cycles. We see two possible explanations for the inconsistency between our previous work and that here:

(1) The correct conclusions about the extent of contrast integration are drawn in our current work, with previous work being compromised by the loss of sensitivity with retinal eccentricity. For example, Baker and Meese (2011) built witch's hat compensation into their modeling, but not their stimuli (in which they manipulated carrier and modulator spatial frequencies, not diameter). A loss of experimental effect in the results (such as that in Figures 2a and 3a here) limits what the analysis can be expected to reveal. Indeed, Baker and Meese (2011) found it difficult to put a precise figure on the range of contrast integration, and aspects of their analysis hinted at a range of >20 cycles for two of their three observers. Baker and Meese (2014) made no allowance for eccentricity effects in their reverse correlation study. The contrast jitter applied to their target elements ensured they were above threshold, and so the effects of contrast constancy should come into play (Georgeson,

1991); however, we cannot rule out the possibilities that either (a) the contrast constancy process was incomplete or (b) internal noise effects not evident at detection threshold (e.g., signal dependent noise) compromised the conclusions.

(2) The correct conclusions about the extent of contrast integration come from our previous work. Our current work points to lawful fourth-root summation, but not necessarily signal integration across the full range. On this account, signal integration takes place up to a diameter of about 12 cycles and a different fourth-root summation processes take place beyond that point. For example, from our results here we cannot rule out the following possibility: Beyond an eccentricity of  $\sim 1.5^\circ$  the transducer becomes linear and overall sensitivity improves by probability summation (Tyler & Chen, 2000), but uncertainty (Pelli, 1985; Meese & Summers, 2012) for more peripheral targets causes the slope of the psychometric function to remain steeper than  $\beta = 1.3$  (May & Solomon, 2013).

We think Occam's razor would favor the first account over the second.

*Keywords:* contrast detection, psychophysics, noisy energy model, Swiss cheese stimuli, area summation, spatial summation, witch's hat, retinal inhomogeneity.

## Acknowledgments

This work was supported by an Engineering and Physical Sciences Research Council (EPSRC, UK) grant to TSM and Mark Georgeson (EP/H000038/1). The raw data for the experiments reported here can be found at the following doi: <http://dx.doi.org/10.17036/8733940b-83e9-4cbc-a4b6-9ba0f80e90d8>

Commercial relationships: none.

Corresponding author: Tim S. Meese.

Email: [t.s.meese@aston.ac.uk](mailto:t.s.meese@aston.ac.uk).

Address: School of Life and Health Sciences, Aston University, Birmingham, UK.

## References

- Baker, D. H., & Meese, T. S. (2011). Contrast integration over area is extensive: A three-stage model of spatial summation. *Journal of Vision*, *11*(14):14, 1–16, doi:10.1167/11.14.14. [PubMed] [Article]
- Baker, D. H., & Meese, T. S. (2014). Measuring the



- spatial extent of texture pooling using reverse correlation. *Vision Research*, 97, 52–58.
- Baldwin, A. S., Meese, T. S., & Baker, D. H. (2012). The attenuation surface for contrast sensitivity has the form of a witch's hat within the central visual field. *Journal of Vision*, 12(11):23, 1–17, doi:10.1167/12.11.23. [PubMed] [Article]
- Bradley, C., Abrams, J., & Geisler, W. S. (2014). Retina-V1 model of detectability across the visual field. *Journal of Vision*, 14(12):22, 1–22, doi:10.1167/14.12.22. [PubMed] [Article]
- Burgess, A., & Ghandeharian, H. (1984). Visual detection. I. Ability to use phase information. *Journal of the Optical Society of America, A*, 1(8), 900–905.
- Foley, J. M., Varadharajan, S., Koh, C. C., & Farias, M. C. Q. (2007). Detection of Gabor patterns of different sizes, shapes, phases and eccentricities. *Vision Research*, 47, 85–107.
- Georgeson, M. A. (1991). Contrast overconstancy. *Journal of the Optical Society of America, A*, 8(3), 579–586.
- Graham, N. V. S. (1989). *Visual pattern analyzers*. New York: Oxford University Press.
- Hoekstra, J., van der Goot, D. P. J., van den Brink, G., & Bilsen, F. A. (1974). The influence of the number of cycles upon the visual contrast threshold for spatial sine wave patterns. *Vision Research*, 14, 365–368.
- Howell, E. R., & Hess, R. F. (1978). The functional area for summation to threshold for sinusoidal gratings. *Vision Research*, 18, 369–374.
- Kingdom, F. A. A., Baldwin, A. S., & Schmidtman, G. (2015). Modeling probability and additive summation for detection across multiple mechanisms under the assumptions of signal detection theory. *Journal of Vision*, 15(5):1, 1–16, doi:10.1167/15.5.1. [PubMed] [Article]
- Manahilov, V., & Simpson, W. (1999). Energy model for contrast detection: Spatiotemporal characteristics of threshold vision. *Biological Cybernetics*, 81, 61–71.
- Manahilov, V., & Simpson, W. A. (2001). Energy model for contrast detection: Spatial-frequency and orientation selectivity in grating summation. *Vision Research*, 41, 1547–1560.
- May, K. A., & Solomon, J. A. (2013). Four theorems on the psychometric function. *PLoS One*, 8(10), 1–34.
- Mayer, M. J., & Tyler, C. W. (1986). Invariance of the slope of the psychometric function with spatial summation. *Journal of the Optical Society of America, A*, 3(8), 1166–1172.
- Meese, T. S. (2010). Spatially extensive summation of contrast energy is revealed by contrast detection of micro-pattern textures. *Journal of Vision*, 10(8):14, 1–21, doi:10.1167/10.8.14. [PubMed] [Article]
- 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, doi:10.1167/11.1.23. [PubMed] [Article]
- Meese, T. S., Hess, R. F., & Williams, C. B. (2005). Size matters, but not for everyone: Individual differences for contrast discrimination. *Journal of Vision*, 5(11):2, 928–947, doi:10.1167/5.11.2. [PubMed] [Article]
- Meese, T. S., & Summers, R. J. (2007). Area summation in human vision at and above detection threshold. *Proceedings of the Royal Society B*, 274, 2891–2900.
- 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(4):7, 1–16, doi:10.1167/9.4.7. [PubMed] [Article]
- Meese, T. S., & Summers, R. J. (2012). Theory and data for area summation of contrast with and without uncertainty: Evidence for a noisy energy model. *Journal of Vision*, 12(11):9, 1–28, doi:10.1167/12.11.9. [PubMed] [Article]
- Meese, T. S., & Williams, C. B. (2000). Probability summation for multiple patches of luminance modulation. *Vision Research*, 40, 2101–2113.
- Morgenstern, Y., & Elder, J. H. (2012). Local visual energy mechanisms revealed by detection of global patterns. *Journal of Neuroscience*, 32(11), 3679–3696.
- Nachmias, J. (1981). On the psychometric function for contrast detection. *Vision Research*, 21, 215–223.
- Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America, A*, 2(9), 1508–1532.
- Pelli, D. G. (1987). On the relation between summation and facilitation. *Vision Research*, 27(1), 119–123.
- Prins, N., & Kingdom, F. A. A. (2009). Palamedes: Matlab routines for analyzing psychophysical data. Retrieved July 1, 2013 from [www.palamedestoolbox.org](http://www.palamedestoolbox.org)
- Quick, R. F. (1974). A vector-magnitude model of contrast detection. *Kybernetik*, 16, 6567.
- Robson, J. G., & Graham, N. (1981). Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Research*, 21, 409–418.

- Rovamo, J., Luntinen, O., & Näsänen, R. (1993). Modelling the dependence of contrast sensitivity on grating area and spatial frequency. *Vision Research*, 33(18), 2773–2788.
- Savoy, R. L., & McCann, J. J. (1975). Visibility of low-spatial-frequency sine-wave targets: Dependence on number of cycles. *Journal of the Optical Society of America*, 65(3), 343–350.
- Summers, R. J., & Meese, T. S. (2009). The influence of fixation points on contrast detection and discrimination of patches of grating: Masking and facilitation. *Vision Research*, 49, 1894–1900.
- Tootle, J. S., & Berkley, M. A. (1983). Contrast sensitivity for vertically and obliquely oriented gratings as a function of grating area. *Vision Research*, 23(9), 907–910.
- Tyler, C. W., & Chen, C.-C. (2000). Signal detection theory in the 2AFC paradigm: Attention, channel uncertainty and probability summation. *Vision Research*, 40, 3121–3144.
- Wallis, S. A., Baker, D. H., Meese, T. S., & Georgeson, M. A. (2013). The slope of the psychometric function and non-stationarity of thresholds in spatiotemporal contrast vision. *Vision Research*, 76, 1–10.

## Appendix A

Here we present the mathematical forms of the models used in the main body of the paper. We assume that the limiting internal noise is additive and Gaussian, meaning that increasing or decreasing its standard deviation will translate model predictions vertically (on a log axis). This was the single offset parameter used in fitting the models to the grating results in Figure 2. These values were then used in making the zero-free parameter predictions in Figure 3. Note that if we are interested only in summation *slope*, then we can set the internal noise  $\sigma$  to a fixed arbitrary value ( $\sigma = 1$ ).

For each model, we first derive the exact model equation used in fitting the behavioral data in the main body of the paper. We then simplify matters by disregarding the effects of spatial filtering and the very minor effects caused by the skirt of the overall stimulus window. We do this to predict asymptotic summation slopes, where the number of signal pixels ( $n$ ) is directly proportional to the square of the nominal stimulus diameter (the full width at half window height).

### Noisy energy model

The signal to noise ratio ( $d'$ ) for the general form of the noisy energy model is

$$d' = \frac{\sum_{i=1}^p (c \cdot s_i)^2 \cdot t_i}{\sqrt{\sum_{i=1}^p (\sigma_i \cdot t_i)^2}}, \quad (\text{A1})$$

where  $c$  is the stimulus contrast,  $s_i$  is the amplitude of the  $i^{\text{th}}$  of  $p$  pixels in the filtered “witch’s hat attenuated” image (note that  $p$  is constant for all images; it is the number of pixels in the display and does not depend on stimulus diameter),  $t_i$  is the amplitude of the template at that pixel (in the noisy energy model presented here the template is matched to the outer boundary of the stimulus and does not match the Swiss cheese modulations), and  $\sigma_i$  is its internal noise standard deviation (Figure A1). Although each pixel has two coordinates ( $x$  and  $y$ ), we collapse these to a single dimension ( $i$ ) for ease of presentation. Solving Equation A1 for  $d' = 1$  gives

$$c_{\text{threshold}} = \frac{\sqrt[4]{\sum_{i=1}^p (\sigma_i^2 \cdot t_i^2)}}{\sqrt{\sum_{i=1}^p s_i^2 \cdot t_i}}. \quad (\text{A2})$$

This is the model equation used to generate the noisy energy model predictions. We can also demonstrate why a fourth-root slope is predicted for compensated gratings where the template is matched to the stimulus ( $t = s$ ), meaning that the template for the nonsignal pixels is zero. In this case we consider the summation that occurs over the  $n$  signal pixels determined by the

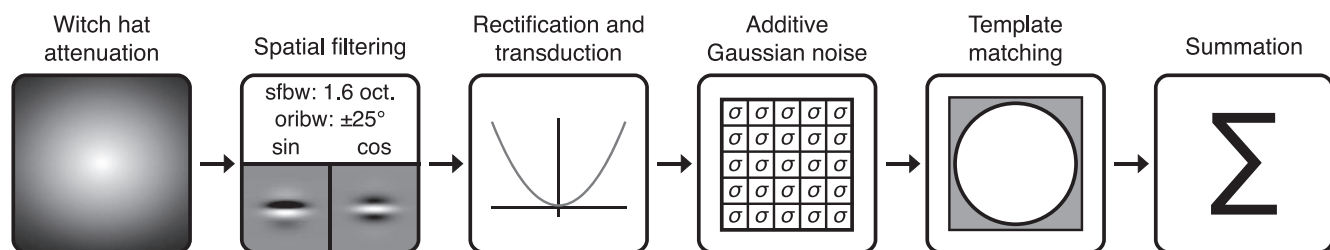


Figure A1. A schematic illustration of the noisy energy model. In this model, the template is matched to stimulus size and the blurred boundary of the envelope. It is not matched to other stimulus modulations in either luminance or contrast.

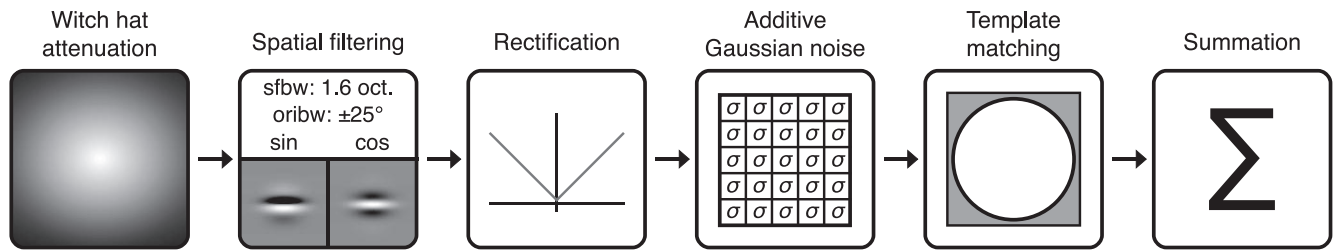


Figure A2. A schematic illustration of the quadratic model. Note that since summation is mandatory across the entire image, this scheme is equivalent to having a single source of late noise after summation.

area derived from nominal stimulus diameter, where the output of the filtering (sum of the sine and cosine phase filters) is uniform (thereby ignoring the minor effects of the blurred boundary to the stimuli).

$$c_{\text{threshold}} = \frac{\sqrt[4]{\sum_{i=1}^n (\sigma_i^2 \cdot s_i^2)}}{\sqrt{\sum_{i=1}^n s_i^3}} \tag{A3}$$

For a constant  $s$  (obtained from quadrature filtering) and  $\sigma = I$

$$c_{\text{threshold}} = \frac{1}{\sqrt[4]{n \cdot s}} \tag{A4}$$

so for any  $s$ , the summation slope asymptotes to

$$c_{\text{threshold}} \propto \frac{1}{\sqrt[4]{n}} \tag{A5}$$

### Quadratic model

The signal to noise ratio for the quadratic model (Figure A2) is

$$d' = \frac{\sum_{i=1}^p (c \cdot s_i)^2}{\sqrt{\sum_{i=1}^p \sigma_i^2}} \tag{A6}$$

and is solved for contrast detection threshold in a

similar manner as for A. The effective internal noise is constant, giving an asymptotic summation slope of

$$c_{\text{threshold}} \propto \frac{1}{\sqrt{n}} \tag{A7}$$

for compensated gratings.

### Matched template model

The signal-to-noise ratio for a model where there is a linear transducer followed by a template matched exactly to the stimulus (Figure A3) is

$$d' = \frac{\sum_{i=1}^p c \cdot s_i \cdot t_i}{\sqrt{\sum_{i=1}^p (\sigma_i \cdot t_i)^2}} \tag{A8}$$

Solving for the contrast threshold prediction in a similar manner as for Equation A5 gives the same result as the quadratic model (Equation A7) for compensated gratings.

### Probability summation model under HTT

For the probability summation model under the assumptions of HTT, the response at each location in

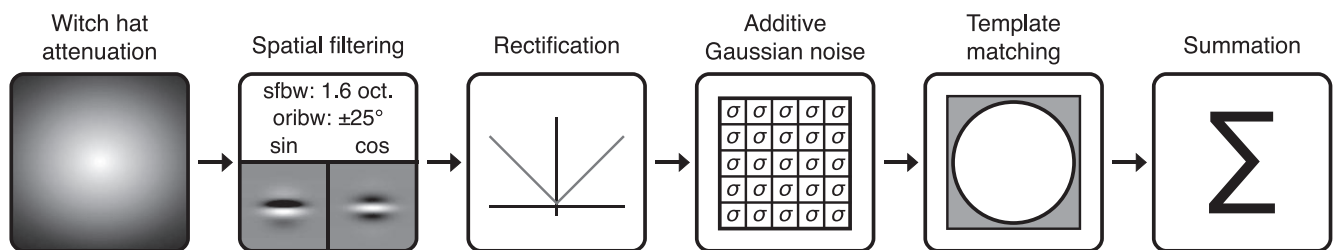


Figure A3. A schematic illustration of the matched template model. In this model, the template is an exact template of the filtered image after attenuation by the simulated spatial inhomogeneity. (However, note that for our purposes here, we consider this model only for the case of compensated gratings. In that case, the situation is equivalent to there being no spatial inhomogeneity [to the extent that our witch’s hat matches that exactly], and no provision for inhomogeneity in the template.)

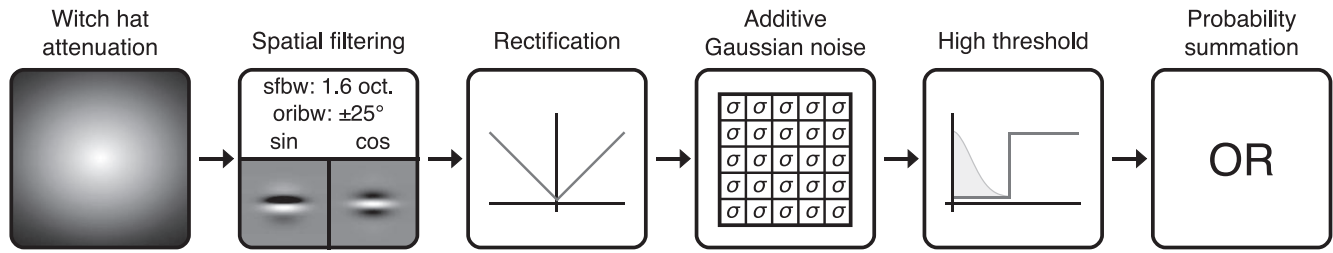


Figure A4. A schematic illustration of the HTT of probability summation.

the image is perturbed by independent noise. There is then a threshold set sufficiently high that it is only rarely exceeded by the noise alone (Figure A4). The probability of detecting the stimulus can be derived by combining the probabilities of detection for each individual pixel

$$P(\text{detect}) = 1 - \prod_{i=1}^n 1 - P(s_i). \tag{A9}$$

The thresholds predicted from such a system are given by Minkowski summation over the detector outputs (assuming that the psychometric function is a Weibull function), with an exponent equal to the slope parameter of the psychometric function (Weibull  $\beta$ ; Quick, 1974):

$$c_{\text{threshold}} = \left( \sum_{i=1}^n s_i^\beta \right)^{-\frac{1}{\beta}}, \tag{A10}$$

and for a constant value of  $s$  we have:

$$c_{\text{threshold}} \propto \frac{1}{\sqrt[n]{n}}, \tag{A11}$$

for compensated gratings.

### Probability summation model under signal detection theory

Although the original probability summation model was based on HTT, which has been discredited (Nachmias, 1981), it has been reformulated under signal detection theory (Pelli, 1985; Tyler & Chen, 2000). This version replaces the high threshold with a max operation over noisy detector outputs (Figure A5). For reasonable assumptions about uncertainty (Tyler & Chen, 2000; see also Meese & Summers, 2012), threshold predictions from this model can be approximated by Equation A10, with  $\beta = 4$ . Lower exponents are justified under particular conditions of uncertainty ( $\beta < 4$ ; see Tyler & Chen, 2000 for details), and higher exponents when the model includes an accelerating contrast transducer (equal to 4 times the transducer exponent; Meese & Summers, 2012).

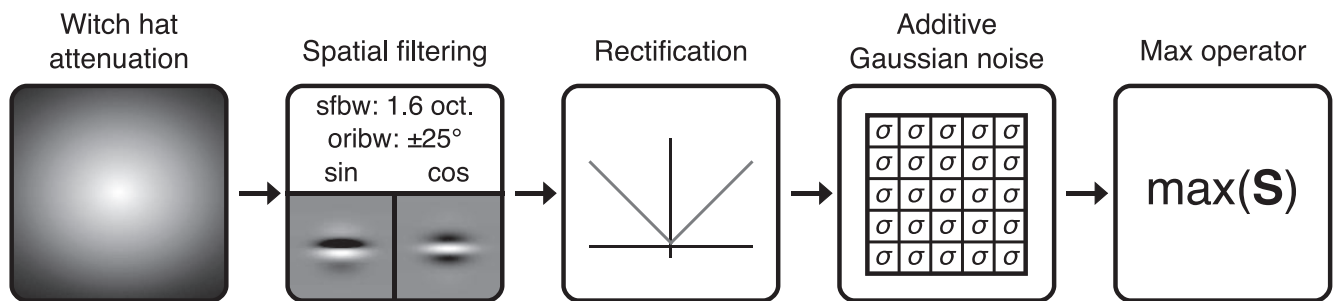


Figure A5. A schematic illustration of a signal detection theory of probability summation.