1 Model details and motivation

A schematic explanation of our model is shown in figure 1. Conceptually the model combines three signals to produce decisions about motion direction: M_s is a sum of first order motion energy inside a relatively restricted region (shown by the smaller, green dashed circle), $\bar{\Delta_x}$ is an estimate of the change of position of a tracked target, within the same restricted region (indicated by a black line,) M_I is a repulsion due to first order motion signals integrated across a wider region of space (shown by the larger, red dashed circle).

For M_S, we begin by assuming that the strength of the local motion signal is linear with respect to carrier direction content C. However, as spacing decreases and there are more elements within the summation region, we expect signals from multiple elements to sum, so that the local motion signal is proportionate to the direction content of individual elements and inversely proportional to the spacing (since the number of elements near the integration field is inversely proportional to the spacing.)

$$M_S = \frac{C}{S}$$

For dX, we assume that envelope motion is computed by taking two successive noisy estimates of the position of a given element. The strength of the envelope motion estimate therefore has uncertainty proportional to the uncertainty of the successive position estimates. The sensitivity to displacements, β_{Δ_x} will be inversely proportional to this uncertainty. We assume that as flanking elements come closer to the tracked element, the the visual system will be less able to isolate a single element to determine its position,, so the sensitivity to changes in position will collapse. We model this by making $\beta_{\Delta_x}(S)$ a sigmoid function of the spacing between elements (figure 1): :

$$\beta_{\Delta_x}(S) = \beta_0 \left(2 - \frac{2}{1 + e^{-\frac{S_C}{S}}} \right)$$

This function results in sensitivity that approaches β_0 at large spacings and approaches zero at small spacings, as plotted in the lower graph of figure 1. β_0 is the sensitivity to envelope displacement for uncrowded targets; S_C describes the distance over which spatial interference between targets takes place. S_C has a natural interpretation as the spacing at which the threshold for discriminating a feature (in this case envelope motion) increases by a constant factor, which is often used as an empirical definition of the "critical distance" in crowding. As can be seen from this function, our model assumes that sensitivity to envelope motion approaches an asymptote above S_C, and therefore does not change as a function of the number of targets, once the targets are sufficiently far apart.

In other words, decreasing the spacing between elements has opposite effects on the strengths of the carrier and the envelope motion. For the carrier motion estimate, cramming more elements into the summation region increases the strength of the carrier motion signal, which is a simple summation of motion

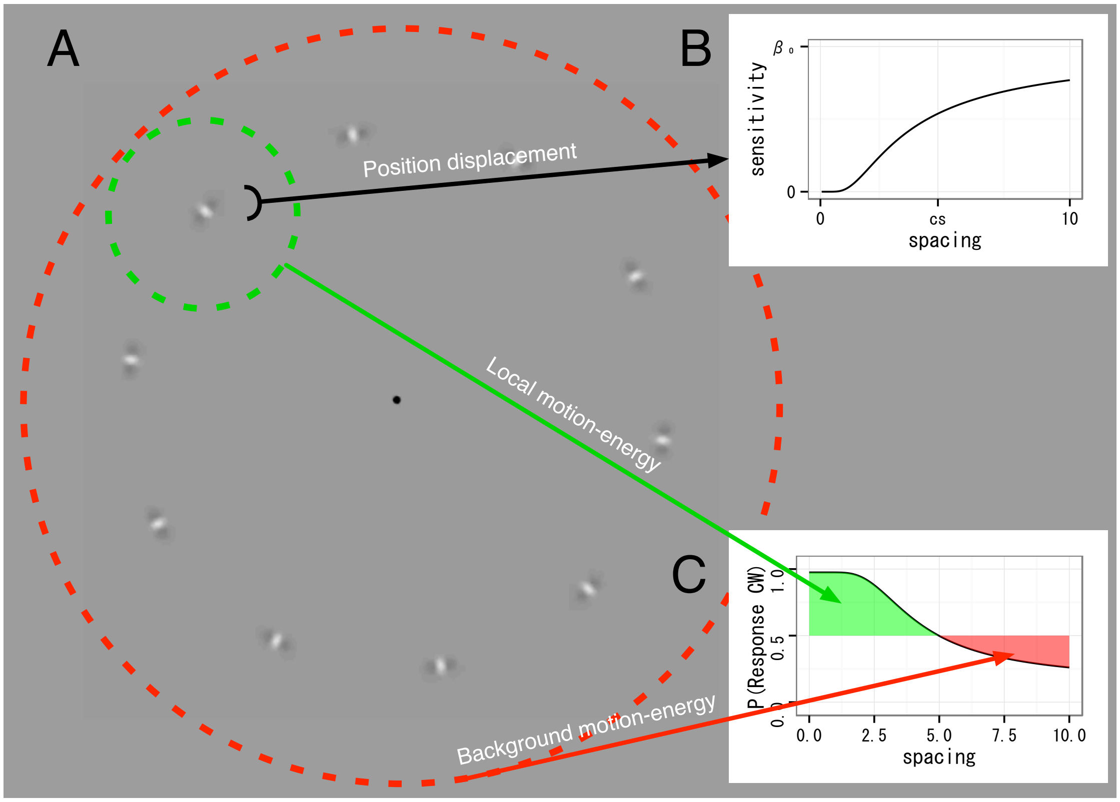

Figure 1: Schematic of motion model. (A) A typical motion stimulus is shown. The motion stimulus contains three sources of information: the positional displacement of an individual, attended element, a summation of first order motion energy within a local region, illustrated by the green dashed circle, and a summation of motion energy over most or all of the visual field, illustrated by the red dashed line. (B) Envelope motion is estimated by taking two successive noisy estimates of the position of a given element. Due to "crowding" the uncertainty in position $1/\beta(S)$ increases as a function of the flanker spacing S . The form of $\beta(S)$ is a sigmoid, as in (B). Its parameters are β_0, reflecting the sensitivity in position of an uncrowded target, and S_C, which is the distance at which discrimination thresholds for envelope motion double. (C) Observers attempt to determine the direction of envelope motion but their perceptions are biased by the first-order motion content of the stimulus. This has two components, a local summation, and a background average. The local component is a summation of all motion energy within a local region of fixed size around the attended object, as illustrated by the green dashed circle. As the spacing between targets decreases, more targets are brought into the summation region, causing the bias to move toward the direction of the carrier motion (shaded green region of panel C). The second contribution of carrier motion is to an average over much of the visual field. When local summation is weak (target spacing is wide) we find that subjects are biased against the carrier motion. Therefore there is a change between summation and repulsion as a function of spacing (green and red shaded region of the graph in panel C.) The strength of repulsion is a function of the average motion energy content of elements in the scene. Observers' responses are modeled as a thresholded sum of the noisy position-displacement measurement and the biases due to short-range and long range motion energy.

energy within that region. For the envelope motion estimate, cramming more elements into the summation region makes it difficult to estimate the position of any individual element, and thereby weakens the signal. Thus, decreasing the spacing results in carrier motion having a stronger influence on the final percept, and envelope motion having a weaker influence.

One addition to this basic model is required to account for the effect of first order motion signals outside our small summation region. The third component M_I is an "induced motion" effect, whereby the perceived direction of motion of an attended target is also influenced by the motion-energy of the full display. The influence of this process is generally repulsive; carrier motion outside the integration field biases observers towards perceiving the individual elements as moving in the opposite direction. However, this effect can have a nonlinearity, which (see Results). We modeled this using a function defined as:

$$M_I = \left(\beta_{I_a} C + \beta_{I_b} C|C|\right),$$

that is, a linear component controlled by the coefficient β_{I_a} and a second-order, odd-symmetric component controlled by the coefficient β_{I_b}. This second component was necessary to account for the nonmonotonic effect of direction content on subjects' responses (see section **??**.)

These three components contribute to the modeled subjects' responses according to a probabilistic rule;

$$\Pr\left(\text{clockwise} \mid \Delta x, C, s, n\right) = \text{logit}(\beta(S)\Delta_x + \beta_S M_S + \beta_I M_I + k)$$

where $\text{logit}(r) = \left(1 + e^{-r}\right)^{-1}$ is the standard logistic cumulative distribution function. The free parameters of the model are β_0, S_C, β_S, β_{Ia}, and β_{Ib}. We only use k as necessary account for an overall clockwise or counterclockwise bias; for subjects with no such bias we use $k = 0$.