Optimal Attentional Modulation of a Neural Population

Ali Borji[a,*], Laurent Itti[a,b,c]

[a]*Department of Computer Science, University of Southern California, 3641 Watt Way, Los Angeles, CA 90089*
[b]*Department of Neuroscience, University of Southern California, 3641 Watt Way, Los Angeles, CA 90089*
[c]*Department of Psychology, University of Southern California, 3641 Watt Way, Los Angeles, CA 90089*

Abstract

Top-down attention has often been separately studied in the contexts of either optimal population coding or biasing of visual search. Yet, both are intimately linked, as they entail optimally modulating sensory variables in neural populations according to top-down goals. Designing experiments to probe top-down attentional modulation is difficult because non-linear population dynamics are hard to predict in the absence of a concise theoretical framework. Here, we describe a unified framework that encompasses both contexts. Our work sheds light onto the ongoing debate on whether attention modulates neural response gain, tuning width, and/or preferred feature. We evaluate the framework by conducting simulations for two tasks: 1) classification (discrimination) of two stimuli s_a and s_b and 2) searching for a target T among distractors D. Results demonstrate that all of gain, tuning, and preferred feature modulation happen to different extents, depending on stimulus conditions and task demands. The theoretical analysis shows that task difficulty (linked to difference Δ between s_a and s_b, or T and D) is a crucial factor in optimal modulation, with different effects in discrimination vs. search. Further, our framework allows us to quantify the relative utility of neural parameters. In easy tasks (when Δ is large compared to the density of the neural population), modulating gains and preferred features is sufficient to yield nearly optimal performance; however, in difficult tasks (smaller Δ), modulating tuning width becomes necessary to improve performance. This suggests that the conflicting reports from different experimental studies may be due to differences in tasks and in their difficulties. We further propose future electrophysiology experiments to observe different types of attentional modulation in a same neuron.

Keywords: top-down attention, neural modulation, neural coding, gain, tuning width, feature selectivity

1. Introduction

Optimal neural coding, or efficient coding, suggests that sensory systems have evolved to optimize the representation of the world around us. Two seemingly different fields of study, neural coding and visual search, have addressed neural modulation. The former has mainly investigated the optimal tuning width for a population of neurons (often one value for all neurons) in stimulus reconstruction and discrimination tasks (e.g., Berens et al. (2011); Zhang & Sejnowski (1999); Wang et al. (2012); Jazayeri & Movshon (2006)). For example the question of whether sharpening or broadening a neuron's tuning might improve performance has attracted significant interest (e.g., Zhang & Sejnowski (1999); Pouget et al. (1999)). Computational studies of top-down biasing of visual search, on the other hand, have primarily addressed optimal gain modulation (e.g., Navalpakkam & Itti (2007); Scolari & Serences (2009, 2010); Scolari et al. (2012)). Optimal neural modulation, in general, is a complex optimization problem since several variables such as statistics of stimuli, task variability, limitations of neural systems (e.g., number of neurons and parameters, metabolic cost, noise), and coupled nonlinear dynamics are involved. Here, we present a reconciled and abstract account of optimal neural modulation by solving for the best set of gain, tuning width and preferred feature of

*Corresponding author
Email addresses: `borji@usc.edu` (Ali Borji), `itti@pollux.usc.edu` (Laurent Itti)

individual neurons to maximize classification and visual search performance. We use terms *attention* and *optimal neural modulation* interchangeably since the term "attention", as currently used in the literature, refers to a highly heterogeneous class of phenomena. Characteristics of these phenomena vary significantly depending on the specific context in which the nervous system is operating (e.g., different time scales, tasks, environments, etc.).

1.1. Overview of attentional modulation

Finding a friend amidst several hundred passengers at an airport can be a nightmare. Yet, our brain handles the explosion of information efficiently by filtering out irrelevant or distracting stimuli, and by drawing our gaze to salient and relevant visual stimuli, through a process known as visual attention (James, 1890; Treisman & Gelade, 1980a; Tsotsos, 1992; Desimone & Duncan, 1995). Specifically, visual attention is believed to help in at least two ways: goal-driven or top-down attention (Corbetta & Shulman, 2002; Yarbus, 1967; Borji & Itti, 2013) might help in focusing on relevant image regions that resemble our friend's appearance, thereby accelerating our search, and stimulus-driven or bottom-up attention (Koch & Ullman, 1985) might alert us to salient image regions like moving cars, pedestrians or dollies in our way, thereby avoiding accidents (Itti & Koch, 2001). Together, top-down and bottom-up attention help us select a few relevant and salient image regions for further processing, including recognition, representation, awareness and action (Desimone & Duncan, 1995; Crick & Koch, 1998). Please see Itti & Koch (2001), Hayhoe & Ballard (2005), Macknik et al. (2008), Eckstein et al. (2009), Nakayama & Martini (2011), Tatler et al. (2011), Schütz et al. (2011), Nakayama & Martini (2011), Kowler (2011), Eckstein (2011), Carrasco (2011), Baluch & Itti (2011), and Borji & Itti (2012) for recent reviews of attentional mechanisms at behavioral and neural levels.

There exists at least three types of attention – *spatial* (Posner et al., 1980; Moran & Desimone, 1985; Kastner et al., 1999; Womelsdorf et al., 2006; Talsma et al., 2007), *feature-based* (Treue & Trujillo, 1999; Saenz et al., 2003; Sohn et al., 2005; Maunsell & Treue, 2006; Serences & Boynton, 2007; J. F. Jehee, 2011) and *object-based attention* (Duncan, 1984, 1996; Roelfsema et al., 1998; Kanwisher & Wojciulik, 2000; Reynolds et al., 2003; Chen, 2012; Cohen & Tong, 2013), depending on whether the basic unit of attentional deployment is a spatial location/region (e.g., the attentional "spotlight" (Treisman & Gelade, 1980b; Crick, 1984; Brefczynski & DeYoe, 1999)), visual feature (e.g., color, orientation), or an object.

Attention offers several behavioral advantages. It is it is known to:

- Improve processing of stimuli at the attended location (Posner et al., 1980),

- Improve detection of faint stimuli and to lower contrast thresholds (Carrasco et al., 2000; Baldassi & Verghese, 2005),

- Improve feature discrimination (Lee et al., 1999),

- Increase spatial resolution (He et al., 1996; Yeshurun & Carrasco, 1998),

- Reject unwanted stimulus noise (Lu & Dosher, 1998; Ling et al., 2009),

- Increase the rate of visual processing (Carrasco & McElree, 2001),

- Affect appearance (Liu et al., 2006).

In effect, attention filters out irrelevant stimuli from the visual input and enables neural resources to be focused on the relevant locations, features and objects (Zhang et al., 2011).

Attentional modulation is widespread in the visual brain and has been observed in multiple areas along the cortical hierarchy including:

- V1 (Motter, 1993; Watanabe et al., 1998; Martinez et al., 1999; Huk & Heeger, 2000; Saenz et al., 2002; Verghese et al., 2012),

- V2 (Motter, 1993; Luck et al., 1997),

- V4 (Haenny & Schiller, 1988; Spitzer et al., 1988; Motter, 1993; Connor et al., 1997; Luck et al., 1997; McAdams & Maunsell, 1999; Williford & Maunsell, 2006; David et al., 2008; Ipata et al., 2012),

- MT (Treue & Maunsell, 1996; Treue & Trujillo, 1999; O'Craven et al., 1997; Saenz et al., 2002; Sohn et al., 2005),

- Lateral Intra-Parietal cortex (LIP) (Bushnell et al., 1981; Colby et al., 1996; Gottlieb et al., 1998; Bisley & Goldberg, 2003),

- Frontal Eye Fields (FEF) (Bichot & Schall, 2002; Bichot et al., 2005; Moore & Fallah, 2004),

- Subcortical structures like Lateral Geniculate Nucleus (LGN) (O'Connor et al., 2002) and Superior Colliculus (SC) (Fecteau & Munoz, 2006; Munoz et al., 1991).

Attentional effects are task-dependent. In separate studies, attention to color/shape has been shown to enhance BOLD activity in V4, while attention in a speed discrimination task increases activity in MT, and attention in a contrast discrimination task increases activity in V1 (Corbetta et al., 1990; Beauchamp et al., 1997; O'Craven et al., 1997; Huk & Heeger, 2000; Verghese et al., 2012). In fact, simply instructing observers to pay attention to different aspects of a same stimulus on different blocks of trials triggers different observable attentional modulation effects, in distinct anatomical and functional cortical areas. For example, Watanabe et al. (1998) showed, using one stimulus with superimposed translating and expanding fields of dots, differential attentional modulation of BOLD activation, depending on whether the task was to attend to the translating or the expanding feature of the stimulus.

Although different neural mechanisms for attention have been reported, the physiology literature presently appears to be divided. Attention to a neuron's preferred location or feature could:

- Cause a leftward shift in the neuron's contrast response function thus increasing the effective contrast of the stimulus (Reynolds et al., 2000; Martinez-Trujillo & Treue, 2002),

- Increase the response gain of the neuron a.k.a multiplicative scaling (Treue & Trujillo, 1999; McAdams & Maunsell, 1999; Womelsdorf et al., 2008; Boynton, 2009; Reynolds & Stoner, 2009; Scolari & Serences, 2010; Scolari et al., 2012; Saproo & Serences, 2010),

- Decrease the neuron's tuning width a.k.a bandwidth scaling (Moran & Desimone, 1985; Spitzer et al., 1988; Haenny & Schiller, 1988),

- Increase neuron's baseline or spontaneous activity a.k.a additive scaling (Luck et al., 1997; Chelazzi et al., 1998; Chawla et al., 1999; Kastner et al., 1999),

- Shift neurons tuned to nearby locations towards the attended location (Connor et al., 1996; Womelsdorf et al., 2006; David et al., 2008; Ipata et al., 2012),

- Modulate neuronal interactions through neuronal synchronization (Fries et al., 2001; Womelsdorf et al., 2007; Womelsdorf & Fries, 2007).

Note that the underlying mechanisms responsible for these observed effects at the single-unit level may be more complex, for example involving biasing or winner-take-all (WTA) competitions among neurons in a local population (Desimone & Duncan, 1995; Lee et al., 1999), or through gain modulation of upstream neurons (McAdams & Maunsell, 1999). Figure 1 illustrates four possible types of attentional modulation of a neural population. Here, we discard the additive scaling since it has been argued that uniform translation of a tuning function does not affect the coding precision of that tuning function (Cover & Thomas, 1991) (but see Saproo & Serences (2010), Paragraph 4 in the Discussion section) and hence information content of a neural population. Further, this simplification makes our analysis easier and tractable.

1.2. Optimal attentional modulation

To gain better insight into above-mentioned discrepancies, we propose a unified account for optimal modulation of neural activity over two tasks: 1) *stimulus classification* (which of two stimuli was presented on the basis of the neural response pattern) and 2) *visual search* (i.e., enhancing the representation of the target stimulus, thus making search easier). Target selection often comes up in the context of a real world task such as visual search where the observer may be looking for a particular target, or for an unknown target that is the odd-ball. Our proposed framework can extend to additional tasks, including match-to-sample (as a neuron's response to the matching stimulus is enhanced while response to any non-matching stimulus is suppressed), discrimination, and stimulus reconstruction.
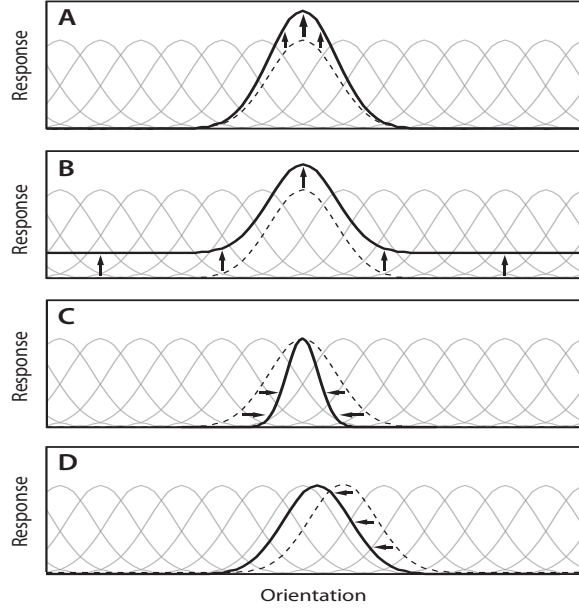

Figure 1: This illustration depicts four possible attention-induced modulations of a neural population to a given visual task (here classification and visual search). Attention theoretically can: **A:** Increase the gain of some important neurons a.k.a multiplicative scaling. This modulation selectivity increases the gain of the neurons that are more useful to find the target in visual search (or two classes in the classification and discrimination tasks). **B:** Enhance response amplitudes in a feature-nonspecific manner a.k.a additive scaling. **C:** Increase the selectivity of a neuron by modulating its tuning width (here sharpening) a.k.a bandwidth scaling, and **D:** Shift tuning curves of neurons around to concentrate on important regions of the feature space (or shifting physical spatial receptive field of a neuron). Faint tuning curves correspond to the neural population before modulation, dotted black curve is the neuron under investigation, and the solid black curve is the modulated tuning curve. Here, we discard case B to make our simulations easier and tractable. Further, it has been argued that this case does not affect information decoding much.

Let $p(\mathbf{r}|s_a)$ and $p(\mathbf{r}|s_b)$ be probability distributions of population activity $\mathbf{r}$ to two stimuli s_a and s_b. The goal of optimal population modulation is to find the best set of parameters for each of n sensory neurons (i.e., $\theta_i = [g_i, \sigma_i, \mu_i]$ including gain, tuning width, and feature selectivity) such that:

$$\phi^* = \arg\max_{\phi} f(p(\mathbf{r}(\phi)|s_a), p(\mathbf{r}(\phi)|s_b)), \quad \phi = [\boldsymbol{\theta}_{i=1...n}] \tag{1}$$

where f denotes the task objective function. For *classification* and *discrimination* tasks, f can be the mutual information between neural activity and behavioral response, or classification accuracy (e.g., linear discrimination error). Here we choose to maximize the inverse of minimum discrimination error (MDE) as the optimality criterion for the classification task. It has been shown that MDE has several advantages over other criteria such as Fisher Information (Berens et al., 2011). For *visual search* tasks, we choose to maximize signal to noise ratio (SNR). The concept of SNR has been suggested by psychophysicists as measured by

the amount of overlap between target (="signal") and distractor (="noise") response distributions. If the purpose is *reconstruction* (i.e., estimate the true value of the presented stimulus on the basis of the noisy neural response $\mathbf{r}$: $\hat{s} = \arg\max_s p(s|\mathbf{r}) \propto \arg\max_s p(\mathbf{r}|s)p(s)$), then f can be the inverse of the mean squared error (MSE) between estimated stimulus (by means of a decoding method such as maximum-likelihood or population vector) and the actual input stimulus.

Optimizing above objective functions is a complex and time consuming process. For the brain this would be an optimization across many (usually thousands of) neurons, involving many different parameters which seems to be very daunting. Note that this does not happen instantly, rather it is a slow process of an organism learning to perform a task. Further, the stimulus distribution is also not available at once and demands the organism to interact with the environment and observe sensory data over time. Indeed, previous work by Baluch & Itti (2010) has shown that human observers become increasingly more efficient at biasing their visual system toward search targets in a triple conjunction search task. This suggests that humans learn over time how to bias the setting of their neural parameters so as to maximize task performance. Navalpakkam & Itti (2007) proposed a three-phase mechanism for learning top-down attentional modulation. In the first phase, bottom-up and top-down cues (learned previously) are applied to render some visual items salient. In the second phase, distributions of target and distractor features are learned through past trials, preview of picture cues, verbal instructions, etc. and in the third phase, optimal top-down gains (as well as other parameters) are computed (See Navalpakkam & Itti (2007), Fig. 2). These gains will be later recalled and applied during future search trials.

2. Theoretical Perspective

We formalize, in the Bayesian sense, how attention may modulate neural activity to optimize task performance. In classification tasks, the goal is to distinguish between a stimulus from class $C = 1$ (defined by a distribution of features $P(s|C = 1)$ in some dimension such as orientation) from a stimulus from class $C = -1$ (defined by a distribution of features $P(s|C = -1)$). In visual search, class $C = 1$ is considered the target T that is to be found among distractors D ($C = -1$).

We assume that the incoming visual display is processed by a population of n neurons tuned to different features. We further assume that all neurons have idealized and homogeneous tuning functions. Let $\mathbf{r}(s) = [r_1(s), r_2(s), \ldots, r_n(s)]$ denote the population vector of responses to input stimulus s. Assuming independent neurons, the probability distribution of response to a single stimulus s is:

$$L_r(s) = p(\mathbf{r}|s) = \prod_{j=1}^{n} p(r_j|s) \tag{2}$$

Classification. In classification tasks, a Bayesian ideal observer needs to estimate $\hat{C} = \arg\max_C P(C|\mathbf{r}) = \arg\max_C P(\mathbf{r}|C)P(C)/P(\mathbf{r})$ where $\hat{C}$ represents the estimated class (out of m classes). This equation means that the classifier chooses the class that was most likely to have caused the observed response pattern $\mathbf{r}$ on the basis of the stimulus conditional response distributions. For a two-class problem, the optimal neural decision variable depends on distributions of neural response to classes $P(\mathbf{r}|C = 1)$ and $P(\mathbf{r}|C = -1)$, each defined as:

$$p(\mathbf{r}|C) = \int p(\mathbf{r}|s)p(s|C)ds = \int L_r(s)p(s|C)ds \tag{3}$$

Thus, to maximize classification performance, the MDE objective function (the error of the ideal observer model) tries to minimize the overlap between neural response distributions to the two classes:

$$MDE(C = 1, C = -1) = \frac{1}{2} \int \min\big(p(\mathbf{r}|C = 1), p(\mathbf{r}|C = -1)\big)d\mathbf{r} \tag{4}$$

Discrimination is a special case of classification, with $p(s|C = 1) = d(s - s_a)$ and $p(s|C = -1) = d(s - s_b)$, where d denotes the Dirac delta function. In (Berens et al., 2011), authors have used MDE to solve for the optimal tuning width of a neural population in reconstruction and discrimination tasks.

Visual Search. Assuming that attention during visual search is guided to locations of high neural activity, search performance can be optimized by maximizing the strength of the signal (expected total neural response to the target $C = 1$) relative to the noise (expected total neural response to the distractors $C = -1$). Thus, using the above formulas, SNR can be written as:

$$SNR(C = 1, C = -1) = \frac{\sum_i E(r_i|C = 1)}{\sum_i E(r_i|C = -1)} = \frac{\sum_i \int r_i p(r_i|C = 1) dr_i}{\sum_i \int r_i p(r_i|C = -1) dr_i} = \frac{\sum_i \int \int r_i p(r_i|s) p(s|C = 1) ds dr_i}{\sum_i \int \int r_i p(r_i|s) p(s|C = -1) ds dr_i}$$
(5)

A closed-form solution for optimal gain modulation using SNR has been previously proposed in (Naval-pakkam & Itti, 2007). Please note that here we attempt to solve visual search in feature space, irrespective of spatial organization of items in the search array. The SNR formulation has been shown to be capable of explaining a large number of psychophysics findings in the visual search literature (Verghese, 2001; Naval-pakkam & Itti, 2007; Scolari & Serences, 2009, 2010; Scolari et al., 2012; J. F. Jehee, 2011). In addition, it has been shown that feature-based attention occurs independently of spatial attention (David et al., 2008), and feature-based attention changes activity globally throughout the visual-field representation (Maunsell & Treue, 2006; Treue & Trujillo, 1999; McAdams & Maunsell, 1999; Saenz et al., 2002; Serences & Boynton, 2007). In other words, attentding to a spatial location all features in that location are enhanced (McAdams & Maunsell, 1999; Boynton, 2009; Ling et al., 2009; Reynolds & Stoner, 2009). Conversely, attention to a specific feature results in global biases to that feature across the entire visual field (Treue & Maunsell, 1996; Treue & Trujillo, 1999; Saenz et al., 2002; Serences & Boynton, 2007).

3. Simulation Results

We run two numerical simulations to investigate the optimal coding quality of a population of neurons under a range of stimulus conditions. The goal of this analysis is to reveal patterns or profiles of modulations depending on tasks and stimuli. Understanding how different patterns arise in different conditions can help design future experiments to pinpoint the neural basis of attentional modulation. In the first simulation, for simplicity and tractability, we choose a neural population of size 6 and we exhaustively search the parameter space. We then run a second, larger simulation with 60 neurons on the most interesting cases. To illustrate our simulations, we consider the feature dimension of stimulus orientation, although our results apply interchangeably to other features such as color, spatial location, or direction of motion.

3.1. Small-scale simulation

We assume a conventional model of neural response, where the i-th neuron ($i \in [1 \ n]$, in a population of $n = 6$ equi-spaced uncorrelated neurons in [0 180]) has a bell-shaped tuning function:

$$f_i(s) = g_i \times \left(\lambda_1 + \lambda_2(\frac{1}{2} + \frac{1}{2}cos(s - \mu_i))^{20\sigma_i}\right); \quad p(r|s) = \frac{1}{\sqrt{2\pi v_i^2}}e^{-\frac{(r - f_i(s))^2}{2v_i^2}}$$
(6)

where s is the scalar stimulus feature (here orientation) and μ_i is the preferred feature of neuron i. The parameter g_i is the multiplicative gain. The parameter σ_i controls the width of the tuning curve. Large σ corresponds to steep tuning curves with small width. The parameters λ_1 and λ_2 set the baseline rate to 5Hz and the maximal rate (amplitude) to 50Hz. The firing activity of each neuron is assumed to follow a Gaussian distribution with Poisson-like noise, where variance is identical to mean spike count (i.e, $v_i^2 = \bar{r}_i(s) = 10f_i(s)$). We estimate MDE and SNR (Eqns. 4 and 5) using Monte Carlo techniques, by iteratively sampling from $p(s|C)$, and, for each s, many times from $p(r|s)$ to finally estimate $p(r|C)$ (similar approach as in Berens et al. (2011)).
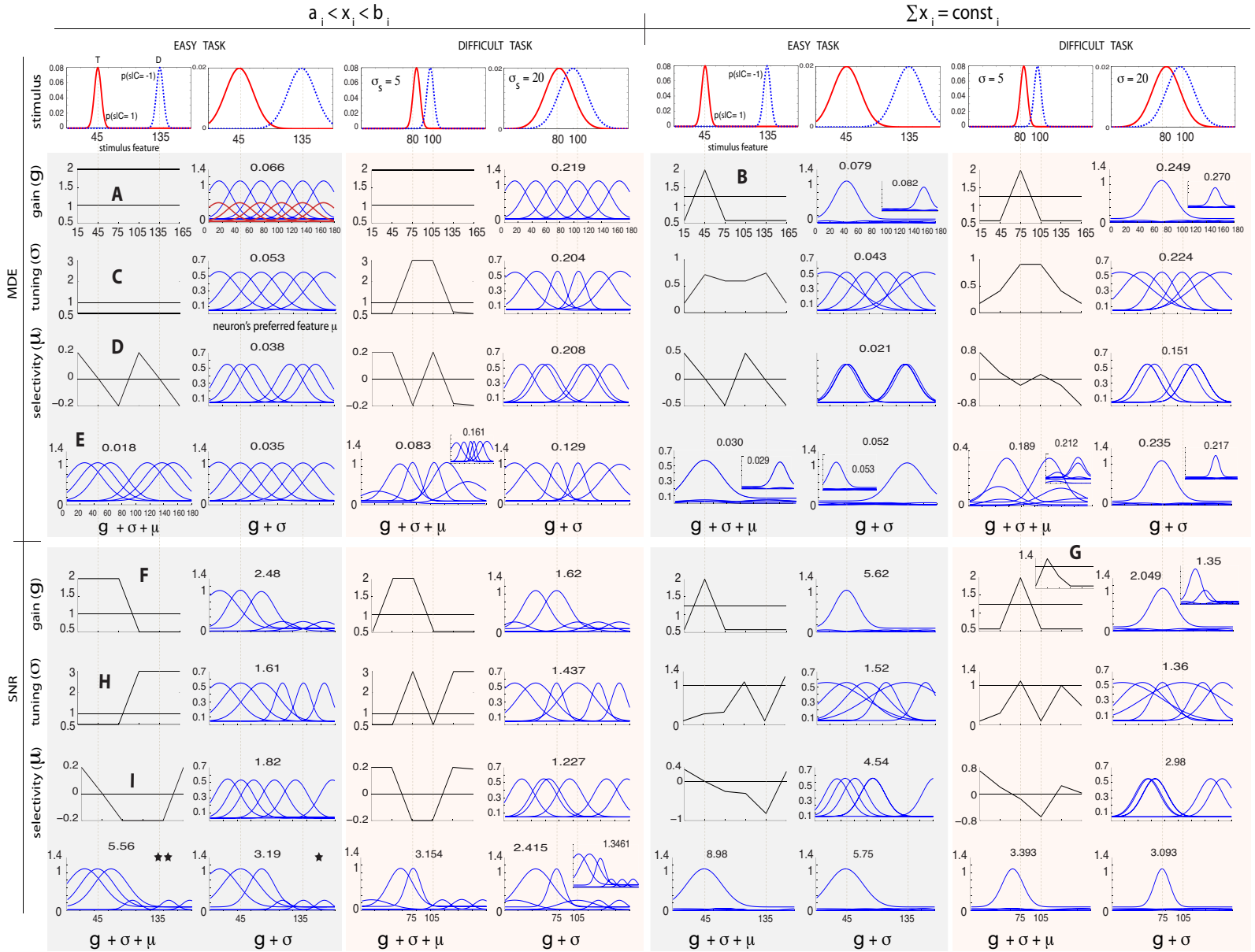
Figure 2: Optimal attentional modulation for classification (MDE; top row) and visual search tasks (SNR; bottom row). The lower the MDE, the better (opposite is true for SNR). Left and right columns correspond to two parameter regimens for 2 classes of stimuli: 1) coarse classification (or easy search) C=1 (target; solid line) at 45° and C=-1 (distractor; dashed line) at 135°, 2) fine classification (or hard search) C=1 at 80° and C=-1 at 100°, each at two uncertainty levels ($\sigma_s = 5°$ and $\sigma_s = 20°$). **A**) MDE, g, regimen 1: gains should be maximized for all neurons in both tasks. The red tuning curves represent the default tunings ($\sigma = 1$). **B**) MDE, g, regimen 2: All gain is allocated to one of the two classes. **C**) MDE, σ, regimen 1: all neurons in easy task should be widened. In other cases, neurons at 2 classes should be sharpened while the rest should be widened. **D**) MDE, μ: neurons should be moved to locations of classes in all cases. **E**) MDE, $g + \sigma$ and $g + \sigma + \mu$ has the superposition of individual effects. **F**) SNR, g, regimen 1: gains of neurons nearby target should be enhanced. In regimen 2, gain at the target should be amplified in easy task. **G**) In difficult search ($\sigma_s = 20°$), the gain of the exaggerated neuron should be amplified more than the neuron at the target. **H**) SNR, σ, easy task: neurons nearby target should be widened while neurons near distractor should be sharpened (see text). In difficult search task, neurons near target should be sharpened while neurons near distractor should be widened. **I**) SNR, μ: neurons should be moved toward the target and away from the distractor.

We consider two types of constraint regimens on neural parameters. The *first regimen* constrains each free parameter to change only within a restricted window, to adhere to biophysical constraints. Note that, otherwise, in visual search, a trivial solution to optimize SNR would be for every neuron to shift its preference to the target feature, change its tuning to infinitely narrow, and enhance its gain infinitely. However, such unbounded changes would likely consume enormous energy (every spike is costly), would prevent neurons from adapting to dynamically changing environments, and are implausible given the electrophysiological observations described in the Introduction. Thus, to prevent indiscriminate changes leading to this mathematical singularity, we constrain each free parameter to change only within a restricted window. We set bounds for g_i to [0.5 2], for σ_i to [0.5 3], and for μ_i to [−0.2 0.2] (in radian, ∼ 11.46°). A default value of 1 for g_i and σ_i, and 0 for μ_i means no change.

Constraint regimen one imposes constraints at the single cell level. Another possibility is to consider constraints at the population level as suggested by Navalpakkam & Itti (2007) where the sum of each parameter over the neural population is constrained (Our *second regimen*, $\sum g_i = 2$, $\sum \sigma_i = 3$, and $\sum \mu_i = 2$). This type of constraint needs more complex mechanisms to impose than constraint type one, for example by means of another neural network or a low-level molecular process. Similar to regimen one, regimen two leads to efficient spending of resources and energy but has more selective pressure as several solutions in regimen one may have equal objective function but in regime two optimization favors most informative neurons. Eventually, our treatment here is theoretical and further biological research is needed to discover which constraint is really implemented in the brain.

We also set the minimum value of g_i and σ_i to be 0.1 to preserve baseline activity. We employ real-valued Genetic Algorithms to exhaustively search the parameter space, in each individual dimension (i.e., g alone), for $g + \sigma$, as well as all three 3 parameters, to maximize SNR and MDE^{-1}. It is worth noting that the qualitative conclusions derived from our simulations do not depend on the exact values of bounds.

Figure 2 shows simulation results obtained by modulating g_i, σ_i, and μ_i in the above manner for two arrangements of stimulus classes: 1) an easy task where two classes are far apart ($C = 1$ at 45° and $C = -1$ at 135°), and 2) a difficult task where two classes are close to each other and thus more similar ($C = 1$ at 80° and $C = -1$ at 100°). We investigate two levels of uncertainty (low $\sigma_s = 5$° and high $\sigma_s = 20$°) on stimulus distributions. For some cases in which solutions are not unique, we also show other good answers in insets. To further study the influence of stimulus distributions and initial parameterization, in Figure 3 we illustrate solutions to some additional cases: 1) when only knowledge about one class is known, 2) three classes of stimuli (two targets and one distractor; See appendix A for heterogeneous search, i.e., one target among two distractors), and 3) narrow default tuning curves ($\sigma_i = 5$). In each test case, we first describe results for classification, then search.

Response gain. In *classification*, under constraint regimen one, all neurons attain the maximum allowed gain, in both easy and difficult tasks. In regimen 2, all gains are concentrated around one of two classes, since both classes are equally important. Interestingly, and possibly counter-intuitively, if we were to distribute the gains equally around both stimulus classes, or equally among all neurons, the MDE would rise (i.e., worse classification). In *visual search*, SNR optimization shows that neurons tuned near the target feature undergo gain enhancement, while neurons tuned near the distractor feature undergo gain suppression (aligned with Navalpakkam & Itti (2007) and Treue & Trujillo (1999)). While in regimen 2, only neurons at the target feature show gain enhancement, in regimen 1 neurons around the target are also enhanced. Interestingly in regimen 2, when target and distractor are very close and overlap is high (Figure 2.F, $T = 80$°, $D = 100$°, $\sigma_s = 20$), in accordance with Navalpakkam & Itti (2007) and Scolari & Serences (2009), we also observe higher gain for the exaggerated neuron (at 45°) than for the neuron best tuned to the target (at 75°). However, unlike Navalpakkam & Itti (2007), baseline activity is sustained in our simulation, which agrees with electrophysiology findings (Chelazzi et al., 1998; Chawla et al., 1999; Kastner et al., 1999; David et al., 2008). Supporting single-unit evidence comes from feature-based attention tasks (Martinez-Trujillo & Treue, 2004; Treue & Trujillo, 1999; McAdams & Maunsell, 1999; David et al., 2008; J. F. Jehee, 2011).

Tuning width. Maximum *classification* accuracy, in the easy task and in regimen 1, is obtained when all
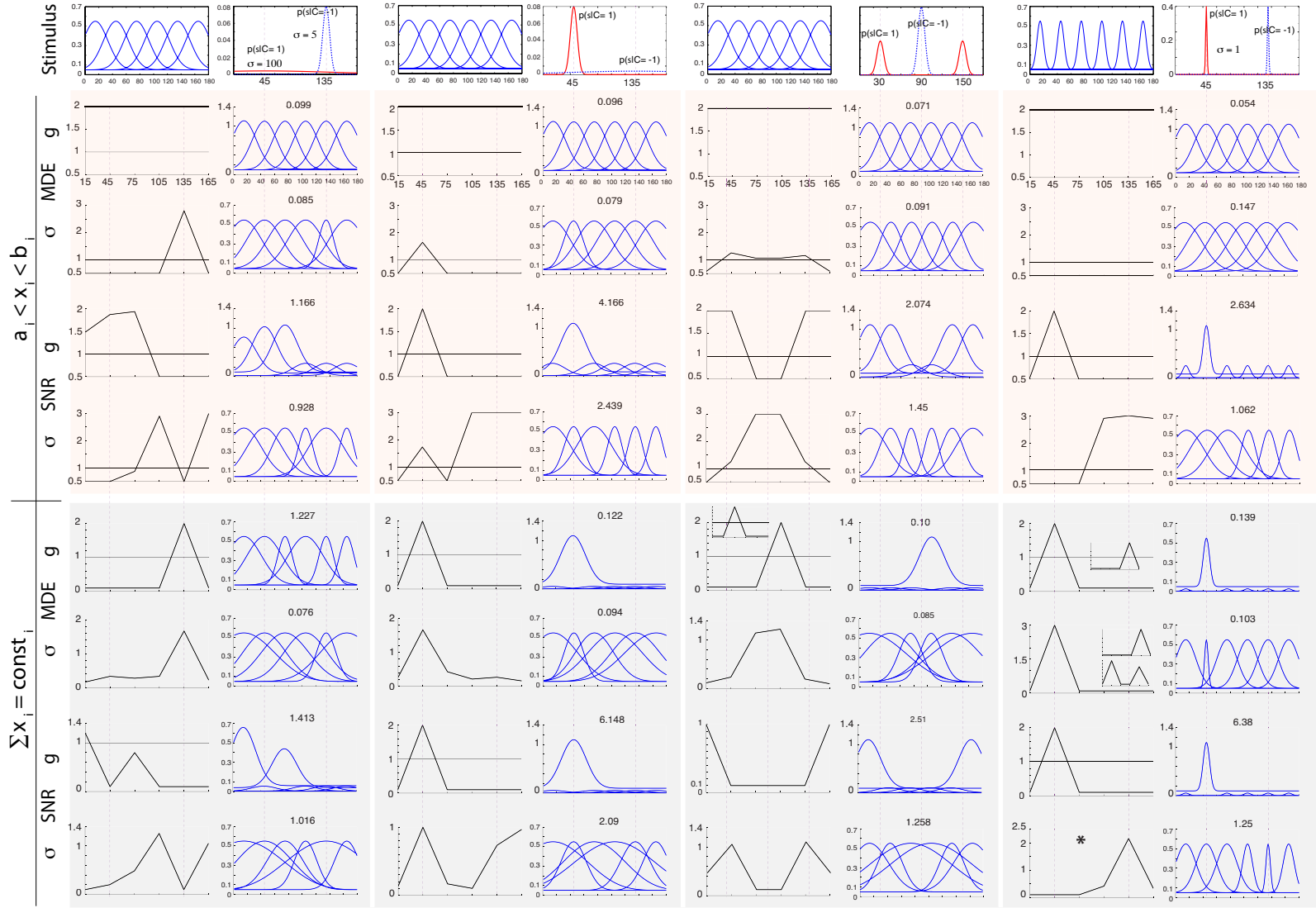
Figure 3: Optimal neural modulation of g, σ and μ for additional cases mentioned in the text (small-scale simulation). Columns from left to right: unknown target (here modeled as a very wide distribution with $\sigma_s = 100$ shown with the dotted blue curve) and known distractor at $135°$ with $\sigma_s = 5$ (solid red curve), known target at $45°$ and unknown distractor, visual search for two targets at $30°$ and $150°$ and a single distractor at $90°$ with $\sigma_s = 5$, easy search for a target at $45°$ with a narrow distribution ($\sigma_s = 1$) and a distractor at $135°$. In each column/setting, the left side shows the original neural population with 6 neurons and below that are the optimal parameters. The right side in each column shows the neural population after modulation. The top rows shows results for regimen 1 while the bottom one corresponds to regimen 2, for both classification (MDE) and visual search tasks (SNR). The panel with * shows the optimal σ in an easy visual search task.

neurons widen their tuning as much as possible. In other cases (difficult task, regimen 1, and both tasks in regimen 2), optimization leads to sharpening near both stimuli and widening elsewhere (see also Figure 3). In *visual search*, our results suggest that attention causes both narrowing and widening of tuning width, and the choice depends on the difficulty of the task. In regimen 1, in the easy task, neurons at and near the target feature are maximally widened while neurons near the distractor feature are maximally sharpened. In regimen 2, in the easy task, we observe widening of neurons both at target and distractor, which was unexpected. Since neurons tuned near the distractor feature already respond strongly to the distractor (due to our bounds), sharpening would indeed only boost the distractor and lower SNR; however, widening for these neurons represents a "better worst-case scenario", as it will make them respond to both distractor and target, resulting in slightly higher SNR compared to sharpening. When we made the task even easier (Figure 3.A), we then observed that neurons at distractor sharpened. Over the difficult task in both regimens, we observe a sharpening at the target and widening near the distractor, which is the opposite of the easy task in regimen 1. When $p(s|T)$ and $p(s|D)$ do not overlap much (i.e., low uncertainty), and/or tuning curves are narrow and far apart, neural tuning widens near the target and sharpens near the distractor. The opposite happens when $p(s|T)$ and $p(s|D)$ highly overlap or the population is very dense. Note that parameter setting is important in the optimal answers. While exact values might differ for different parameter settings, we believe that patterns will stay the same (e.g., dependency of results to task difficulty). For experimental works, when biophysical properties of a neural population are known, it is easy to run a simulation (with our shared code) and verify a hypothesis. Supporting evidence for sharpening at the target comes from single-unit studies of orientation Spitzer et al. (1988) and spatial tuning Moran & Desimone (1985)).

Preferred feature. In *classification*, optimization moves neurons towards either of the two classes as much as possible, in both regimens over both tasks. The optimal answer in *visual search* is to move neurons toward the target and away from the distractor. Supporting evidence for tuning shifts comes from single-unit studies in feature-based (Ipata et al., 2012; David et al., 2008) and spatial attention (Connor et al., 1996; Womelsdorf et al., 2006).

All parameters. Comparing results obtained for the joint optimization of all parameters and the separate optimization of g, σ, and μ, we empirically find that the superposition of optimal answers to each individual parameter is always a good answer (although we do not have a theoretical guarantee on the optimality or uniqueness of such answer). For example, optimizing gain and tuning width jointly in easy visual search, regimen 1 (See Figure 2*), leads to maximal gain amplification and widening of neurons around the target, while minimizing gains of neurons selective to the distractor. Note that tuning width modulation near the distractor is not important here since their gain has already been minimized. When optimizing all three parameters, in addition to the joint answer of gain and tuning width, neurons are also shifted toward the target and away from the distractor (See Figure 2**). Our results also show that modulation of multiple parameters always yields better performance than optimizing only one or two parameters. This suggests that biological top-down attention may also affect multiple parameters, although most previous reports have focused on one parameter at a time.

Optimal neural modulation in heterogeneous visual search (i.e., one target among two distractors and vice versa) and optimizing g, σ, and μ with 12 neurons shows the same patterns as in Figure 2. These results are shown in Appendix B.

Figure 4 shows the optimal MDE and SNR values (in regimen 1) as a function of target-distractor dissimilarity for g, σ, and $g + \sigma$ (averaged over $T \in \{30°, 40°, 50°, 60°\}$ and $D = T + \{10°, 20°, 30°, 40°, 50°, 60°\}$). Increasing the distance between the two classes leads to decrease in MDE and a ramp up in SNR. This qualitatively matches with human performance as a function of task difficulty (Duncan & Humphreys, 1989). Over both MDE and SNR, modulating both g and σ wins over single parameters. The tuning width is more effective than gain in classification, as seen by lower MDE values of σ than MDE values using g. The opposite occurs in visual search using SNR. One reason why SNR values for σ are small might be because neurons in this simulation are not allowed to sharpen beyond a certain limit.
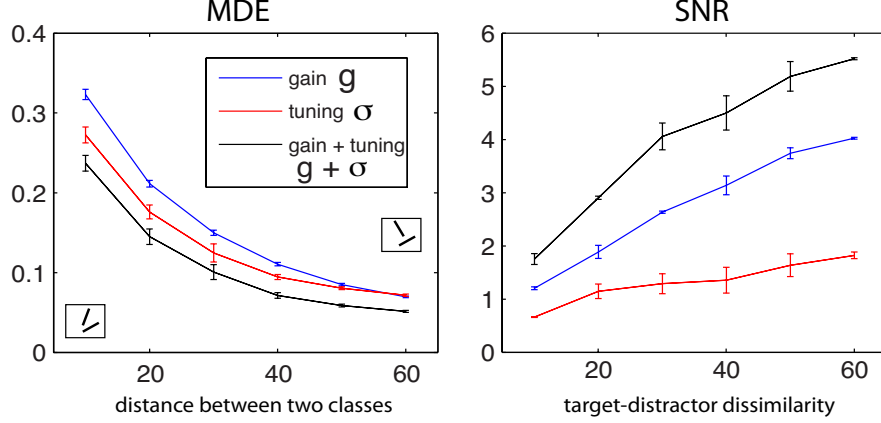
Figure 4: Dependency of objective functions to dissimilarity between two classes for the small-scale simulation with 6 neurons for g, σ, and $g + \sigma$ (averaged over $T \in \{30°, 40°, 50°, 60°\}$ and $D = T + \{10°, 20°, 30°, 40°, 50°, 60°\}$). Left: MDE for classification and Right: SNR for visual search. MDE decreases as two classes become more separate from each other while SNR raises which means that in both cases task becomes progressively easier.

Note on noise correlation. In our simulations so far, we considered optimal modulation of an uncorrelated neural population for the sake of simplicity (i.e., uncorrelated noise). But, noise in the brain is correlated and this might influence the amount of information a neural population conveys (Averbeck et al., 2006) (See also Bejjanki et al. (2011) and Seriès et al. (2004)). Here, we analyze the role of correlations (correlated noise) in optimal modulation of parameters for visual search (i.e., maximizing SNR) on our small scale neural population with 6 neurons.

Following Berens et al. (2011), we model the stimulus-conditional response distribution as a multivariate Gaussian:

$$p(\mathbf{r}|s) = \mathcal{N}(\bar{\mathbf{r}}(s), \Sigma(s)) \tag{7}$$

In above equation, $\bar{\mathbf{r}}(s) = (\bar{r}_1(s), \bar{r}_2(s), \ldots, \bar{r}_6(s))$ and $\Sigma(s)$ represent average spike counts and covariance matrix, respectively. This allows us to inject Poisson-like noise correlations into our simulation (See Berens et al. (2011) and their supplement for more details on adding correlated noise). Results are shown in Fig. 10 (Appendix C) for optimal answers of searching a target at $80°$ and distractor at $100°$ with $\sigma_s = 5°$ (see Fig. 2). We consider 10% noise correlation in our simulations. As it can be seen patterns of results are similar to those shown in Fig. 2 for both constraint regimes and all three neural parameters. This could be because the effect of noise is vanished when averaging the neural activity to targets and to distractors. For future research we encourage a more detailed look at noise correlations (e.g., non-uniform correlations) and how they may affect optimal solutions on larger neural populations.

3.2. Large-scale simulation

The previous analysis revealed different patterns of modulation depending on task and stimulus conditions. Importantly, it revealed that joint optimization of all parameters always yields better performance than optimizing only one parameter. This prompts us to study the relative utility or contribution of modulating each parameter as part of a joint optimization. To further investigate this, we focus on visual search in a larger-scale, more detailed simulation. We simulated a population of $n = 60$ equi-spaced, broad, overlapping Gaussian neurons with preferred stimulus feature μ_i, tuning width σ_i, amplitude λ_2, gain factor g_i, and baseline firing rate λ_1:

$$f_i(s) = g_i \times \left(\lambda_1 + \lambda_2 e^{-(s-\mu_i)^2/2\sigma_i^2}\right), \; i = 1, \ldots, n; \quad p(r|s) = \frac{e^{-f_i(s)} f_i(s)^r}{r!} \tag{8}$$

11

with default tuning width of $10°$, default gains at unity, spacing between preferred orientations of adjacent neurons $3°$) spanning 0-180° in orientation space (Figure 4). In addition, we consider the noise in neural response (to repeated presentations of a same stimulus) to have Poisson variability (used to numerically compute the expectations in the Eq. 5). Here, we set $\lambda_1 = 0$, for simplicity.

We jointly maximized SNR wrt. g_i, μ_i, and σ_i using a multi-start Nelder-Mead simplex algorithm (Nelder & Mead, 1965) (genetic algorithm was too slow in this larger-scale test). We used multiple initial conditions to avoid converging into local optima (20 different initial conditions, each with a random jitter in g_i, μ_i, and σ_i of up to 50% from default values), and considered the solution with maximum SNR. Here, attention can modulate g_i by up to $\pm 50\%$ of its default unity value, and σ_i and μ_i by up to $\pm 50\%$ of the default tuning width (corresponding to **regimen 1** and to avoid numerical instability).

Figure 4 shows how neural parameters may be optimally modulated in an easy search (with an orientation difference between target and distractors of $5\sigma_0 = 50°$), and a difficult search task (smaller orientation difference of $\sigma_0 = 10°$). After modulation, the expected neural response to the target is much higher than the distractor (Figure 4.C) compared to before modulation (Figure 4.B). This effect is more clearly seen in the difficult task, where the initial population response to the target and distractor are similar (Figure 4.B, 2nd column, hence a low SNR), but different after modulation (Figure 4.C, 2nd column), leading to an improvement in SNR. Optimization results here are aligned with our smaller-scale simulation (Figure 2). Interestingly, since here target and distractor are well separated in the easy task, neurons around the target widen while those tuned near the distractor sharpen. In contrast, neurons sharpen near the target and widen near the distractor in the difficult task.

Analysis of tuning curve overlap. How much is SNR dependent on the degree of neural overlap? Over our population of 60 neurons, we change σ from $6°$ to $35°$ and task difficulty from $10°$ to $100°$ and then find the optimal solutions for g, σ, and μ. Figure 5 shows that increasing the overlap between neurons reduces SNR for all parameters regardless of task difficulty. This impairment is more profound in difficult tasks than in easy tasks. In easy tasks, irrespective of the degree of overlap, SNR values using gain are higher than SNR due to σ and μ. SNR using gain increases as the difference between target and distractor increases. Interestingly, there is an interaction between overlap and task difficulty when optimizing for σ and μ (non-monotonic curve shapes in Figure 5).

The analysis of SNR changes as a function of tuning overlap suggests explicit qualitative predictions that could be made when looking across cortical areas (given that orientation tuning inherently broadens as one ascends the visuocortical hierarchy). Moving along the hierarchy, neurons become broader (thus higher overlap among neurons) which eventually causes lower SNR. Also note that the peak of the curves in Figure 5 shifts to the right suggesting that maximum separability happens for more dissimilar stimuli.

Behavioral utility of neural modulation. How useful is the modulation of each neural parameter? To answer this question, we computed a utility statistic $u(p)$ for a parameter $p \in \{g, \sigma, \mu\}$ as the ratio of benefit to SNR obtained by modulating p alone vs. modulating everything. Higher utility values indicate that more performance is achieved by modulating p compared to other parameters, i.e., p is a high-yield parameter to modulate in the particular task and stimulus studied. As seen in Figure 6, $u(g)$ and $u(\mu)$ both decrease with increasing task difficulty, but $u(\sigma)$ does not. Thus, in easy tasks (where the target and distractor differ by $\Delta \geq 40°$) modulating g or μ is more useful, but becomes less useful in difficult tasks. On the other hand, while modulating σ is not very beneficial in easy tasks, it becomes necessary in difficult tasks ($\Delta \leq 25°$). Furthermore, in easy tasks, simulation predicts that the combined modulation of μ and g is sufficient to yield close to best behavioral performance, but their combined utility decreases with increasing task difficulty.

4. Discussion and Conclusion

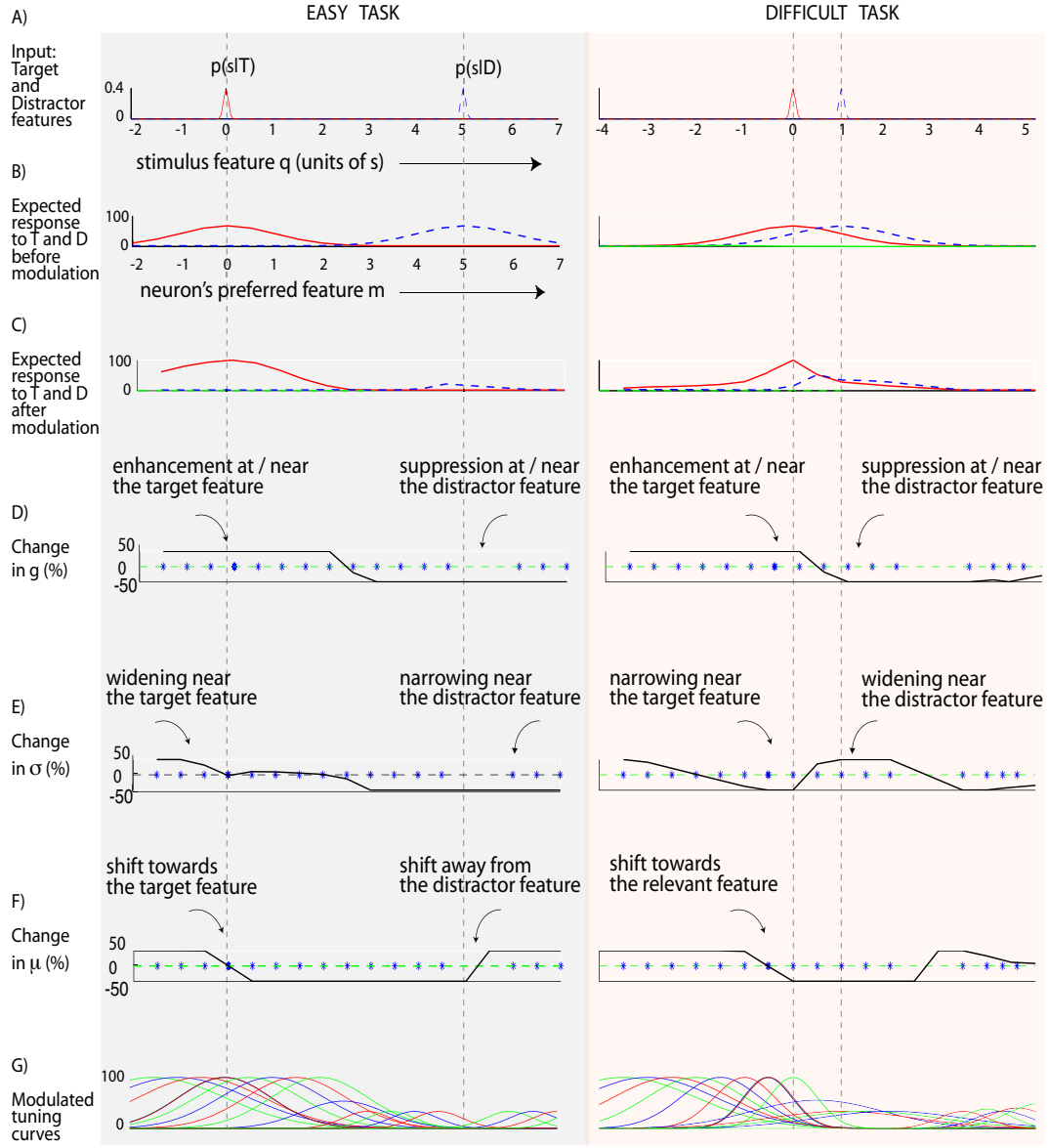Results of two consistent simulations reveal that:

Figure 5: Attentional modulation in easy and difficult visual search. A) The input stimuli. Rows B and C show the expected response of neurons (tuned to different features) before and after modulation. The solid red line is the expected response to the target, while the dotted blue line represents the expected response to the distractor. D) The optimal shift in response gain is shown by the solid black line. Neurons tuned near the target increase their gain, while others tuned near the distractor undergo suppression. E) The optimal shift in neuron's tuning width (σ) is shown here in the solid black line. In the difficult task, neurons tuned to the target feature decrease their tuning width, while nearby neurons widen their tuning width. F) The optimal shift in preferred features μ is shown by the solid black line. A positive shift ($\Delta\mu_i > 0$) indicates neurons shifting to the right, and vice versa. The blue star shows the neuron's preferred feature after the modulation. Neurons shift towards the target feature and away from the distractor feature (as seen by the lack of blue stars near the distractor). G) The optimal tuning curves.
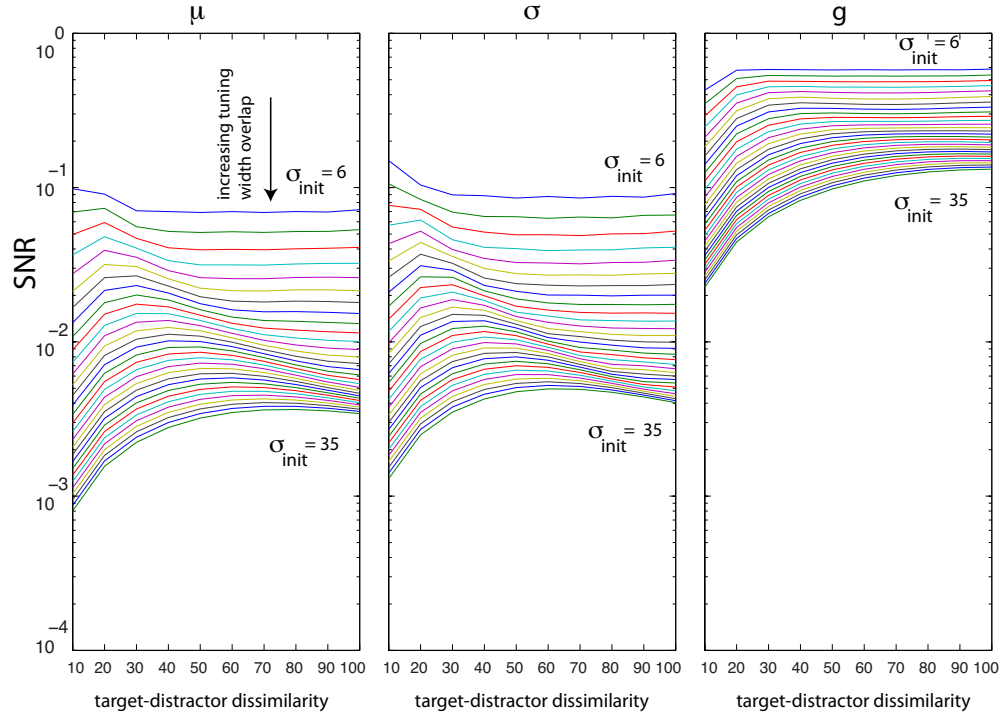
Figure 6: Analysis of tuning curve overlap (σ from $6°$ to $35°$; spacing between neurons is $3°$). The x axis shows task difficulty due to target-distractor dissimilarity (measured by increasing orientation difference between the target and distractor: for $j = 1 : 10$, $T = 60° - j \times 5$, $D = 60° + j \times 5$). The y axis shows the best SNR achieved by optimizing each parameter. Curves from top to bottom indicate higher overlap between neurons. Increasing the neural overlap impairs the SNR due to optimal σ and μ more than SNR by g.
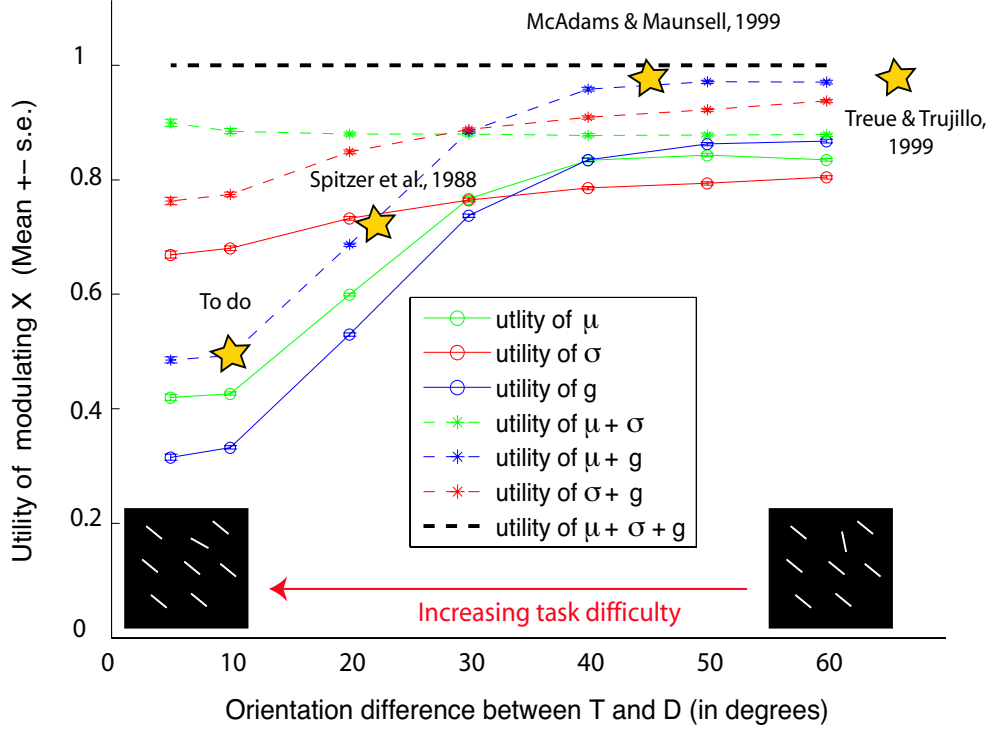
14

Figure 7: Utility of attentional modulation. The x axis shows task difficulty due to target distractor similarity. The y axis shows simulation predictions of utility of modulating preferred features (μ), tuning width (σ), response gains (g), or any combination of these parameters. For easy tasks, we predict that modulating preferred features and gains are useful and sufficient (yielding $0.97 \times$ the best performance). But their combined utility decreases with decreasing orientation difference between the target and distractors (u=0.49), rendering them less useful in difficult tasks. On the other hand, modulating tuning width is more useful and necessary in difficult tasks. A similar trend is observed in separately modulating gains or preferred feature vs. tuning width.

15

1. In classification, when two classes are well separated, all neurons should be widened and gains should be boosted,
2. When two classes are close in feature space, neurons selective to both should be sharpened and their gains should be increased,
3. In easy search, the optimal solution is to widen and boost gain at the target, and sharpen and reduce gain around the distractor (the opposite is seen for tuning width in difficult search),
4. Only in constraint regimen 2 and in difficult search, maximum gain is allocated to the exaggerated neuron as predicted by Navalpakkam & Itti (2007) and seen by Scolari & Serences (2009),
5. Feature selectivity of neurons should be biased toward target features (the two classes in classification) and away from distractors,
6. Optimizing multiple parameters is better than optimizing a single one and joint solutions seem to be combinations of constituent ones,
7. Increasing overlap among neurons worsens SNR, which is more harmful in difficult than in easy search,
8. Uniform noise correlation did not affect our conclusions but more detailed analysis of different noise conditions is encouraged,
9. Task difficulty is a key factor in determining the utility of a neural parameter.

Our theoretical investigation sheds new light on the ongoing controversy of attentional modulation, by indicating that the reported discrepancies in the literature may be due to differences in task difficulty (Figure 6). For instance, previous physiological studies that reported gain modulation (McAdams & Maunsell, 1999; Treue & Trujillo, 1999) used easy tasks: McAdams & Maunsell used an angular difference of 45° or 90° between target and distractor, while Treue & Martinez-Trujillo used either no distractor or one 180° from the target. Previous studies that found preferred feature modulation also used easy tasks: Womelsdorf *et al.* (Womelsdorf et al., 2008) used a spatial attention task where monkeys attended to a target location in the absence of distractors. In such easy tasks, as predicted by our theoretical analysis, modulation of gains and preferred features (which is most useful) is observed, while tuning width modulation (not useful) is not observed. One of the few previous studies (Spitzer et al., 1988) that reported tuning width modulation, observed it in more difficult discrimination tasks (smaller angular difference of 22.5°). Nevertheless, as tuning width modulation remains a controversial issue (e.g., Treue & Trujillo (1999)), our main goal here it to show how tuning width modulation is an optimal strategy when the task is difficult.

It is difficult to disentangle the effect of gain and tuning width modulation behaviorally (see Ling et al. (2009)). We suggest neurophysiology experiments for this purpose by systematically controlling for task difficulty. An ideal task for testing tuning width modulation would be when the monkey attends to a target feature in the presence of flanking distractor (e.g., attend to a 45° oriented moving random dot pattern (RDT) among 50° and 40° oriented RDTs). In such a task, modulating preferred features or gains will not suffice as neurons responding to the target will also respond to similar distractors. Instead, sharpening the tuning curve will help the target-sensitive neurons by decreasing interference from distractors, hence better resolving the difference between target and distractor. In contrast, when the target and flanking distractor are very different (e.g., more than 45° apart), modulating tuning widths is not useful, and thus modulation of preferred features and gains should be observed.

Our model generalizes over previous gain-only models: guided search theory (Wolfe et al., 1989), feature-similarity gain principle (Treue & Trujillo, 1999; Martinez-Trujillo & Treue, 2004), and optimal gain theory (Navalpakkam & Itti, 2007). The guided search theory revises the feature integration theory (FIT) and suggests that top-down attention acts as a linear weighted combination of multiple features which in effect makes an object of interest more salient among distractors and decreases the search time. However, similar to FIT, this theory only attempts to explain the behavior of the organism. In the the feature similarity gain model, gain modulation is a function of similarity between the neuron's preferred feature and the target feature. This theory does not consider target-distractor similarity. The optimal gain theory, combines information from both the target and distracting clutter to maximize the relative salience of the target. Interestingly, this model predicts that it is sometimes optimal to enhance the nontarget features (e.g., Fig. 2, panel G). Here, we considered three neural parameters and showed how distribution of target and distractors can be used to optimally tune all these parameters and make the target salient.

In addition to gain, our model offers testable predictions for tuning width modulation and shifts in selectivity (seen by David et al. (2008) and Ipata et al. (2012) in area V4). Our model differs from the well-established normalization model of attention (Reynolds & Heeger, 2009) in one main aspect: the normalization model commits to explain low-level attentional mechanisms, while our model offers a high-level theoretical account for optimal attention over a population of neurons, considering task difficulty, and stimulus statistics. Obviously, our model has limited prediction power. It may need to be further expanded to account for optimal spatial attention, when deployed jointly with feature-based attention in hybrid spatial/feature tasks. We encourage future neurophysiology studies, with our theoretical framework in hand, to further explore such tasks, which will give new insights for developing unified models of spatial and feature-based attention.

In summary, we investigated three attentional mechanisms, namely attentional modulation of neural response gain, preferred feature and tuning width. Reports from different laboratories differ on whether attention modulates tuning width or gain or preferred feature. We have proposed a simple computational model that reconciles the above differences by predicting that task-difficulty (due to target-distractor similarity) plays a critical role in determining attentional modulation. Our model predicts that gain and preferred feature modulation is useful in easy tasks, while tuning width modulation is useful in difficult tasks – a prediction that is in good qualitative agreement with reported data. This unified model illuminates the similarities and differences in reported data from various laboratories, and provides guidelines for future experiments.

Commercial relationships: none.
Corresponding author: Ali Borji.
Email: borji@usc.edu.
University of Southern California, Hedco Neuroscience Building, 3641 Watt Way, LA, CA 90089, USA

References

Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews Neuroscience*, *7*, 358–366.

Baldassi, S., & Verghese, P. (2005). Attention to locations and features: Different top-down modulation of detector weights. *Journal of Vision*, *5*.

Baluch, F., & Itti, L. (2010). Training top-down attention improves performance on a triple-conjunction search task. *PloS one*, *5*, e9127.

Baluch, F., & Itti, L. (2011). Mechanisms of top-down attention. *Trends in neurosciences*, *34*, 210–224.

Beauchamp, M. S., Cox, R. W., & DeYoe, E. A. (1997). Graded effects of spatial and featural attention on human area MT and associated motion processing areas. *J Neurophysiol*, *78*, 516–520.

Bejjanki, V. R., Beck, J. M., Lu, Z.-L. L., & Pouget, A. (2011). Perceptual learning as improved probabilistic inference in early sensory areas. *Nature neuroscience*, *14*, 642–648.

Berens, P., Ecker, A. S., Gerwinn, S., Tolias, A. S., & Bethge, M. (2011). Reassessing optimal neural population codes with neurometric functions. *Proceedings of the National Academy of Sciences*, *108*.

Bichot, N. P., Rossi, A. F., & Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area v4. *Science*, *308*, 529–534.

Bichot, N. P., & Schall, J. D. (2002). Priming in macaque frontal cortex during popout visual search: feature-based facilitation and location-based inhibition of return. *J Neurosci*, *22*, 4675–4685.

Bisley, J. W., & Goldberg, M. E. (2003). The role of the parietal cortex in the neural processing of saccadic eye movements. *Adv Neurol*, *93*, 141–157.

Borji, A., & Itti, L. (2012). State-of-the-art in modeling visual attention. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, *35*, 185–207.

Borji, A., & Itti, L. (2013). Defending yarbus: Eye movements reveal observers' task. *Journal of Vision.*, .

Boynton, G. M. (2009). A framework for describing the effects of attention on visual responses. *Vision Research*, *49*, 1129–1143.

Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the 'spotlight' of visual attention. *Nat Neurosci*, *2*, 370–374.

Bushnell, M. C., Goldberg, M. E., & Robinson, D. L. (1981). Behavioral enhancement of visual responses in monkey cerebral cortex. i. modulation in posterior parietal cortex related to selective visual attention. *J Neurophysiol*, *46*, 755–772.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*, 1484–1525.

Carrasco, M., & McElree, B. (2001). Covert attention accelerates the rate of visual information processing. *Proc Natl Acad Sci U S A*, *98*, 5363–5367.

Carrasco, M., Penpeci-Talgar, C., & Eckstein, M. (2000). Spatial covert attention increases contrast sensitivity across the csf: support for signal enhancement. *Vision Res*, *40*, 1203–1215.

Chawla, D., Rees, G., & Friston, K. J. (1999). The physiological basis of attentional modulation in extrastriate visual areas. *Nat Neurosci*, *2*, 671–676.

Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *J Neurophysiol*, *80*, 2918–2940.

Chen, Z. (2012). Object-based attention: A tutorial review. *Attention, Perception, & Psychophysics*, *74*, 784–802.

Cohen, E. H., & Tong, F. (2013). Neural mechanisms of object-based attention. *Cerebral Cortex*, .

Colby, C. L., Duhamel, J. R., & Goldberg, M. E. (1996). Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. *J Neurophysiol*, *76*, 2841–2852.

Connor, C. E., Gallant, J. L., Preddie, D. C., & Essen, D. C. V. (1996). Responses in area v4 depend on the spatial relationship between stimulus and attention. *J Neurophysiol*, *75*, 1306–1308.

Connor, C. E., Preddie, D. C., Gallant, J. L., & Essen, D. C. V. (1997). Spatial attention effects in macaque area v4. *J Neurosci*, *17*, 3201–3214.

Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., & Petersen, S. E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science*, *248*, 1556–1559.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci*, *3*, 201–215.

Cover, T. M., & Thomas, J. A. (1991). *Elements of Information Theory*. New York, NY, USA: Wiley-Interscience.

Crick, F. (1984). Function of the thalamic reticular complex: the searchlight hypothesis. *Proc Natl Acad Sci USA*, *81*, 4586–4590.

Crick, F., & Koch, C. (1998). Constraints on cortical and thalamic projections: the no-strong-loops hypothesis. *Nature*, *391*, 245–50.

David, S., Hayden, B., Mazer, J., & Gallant, J. (2008). Attention to stimulus features shifts spectral tuning of V4 neurons during natural vision. *Neuron*, *59*, 509–521.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu Rev Neuro*, .

Duncan, J. (1984). Selective attention and the organization of visual information. *J Exp Psychol Gen*, *113*, 501–17.

Duncan, J. (1996). *Cooperating brain systems in selective perception and action.*. Cambridge, MA: Attention and Performance XVI.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Rev*, .

Eckstein, M. P. (2011). Visual search: A retrospective. *Journal of Vision*, *11*.

Eckstein, M. P., Peterson, M. F., Pham, B. T., & Droll, J. A. (2009). Statistical decision theory to relate neurons to behavior in the study of covert visual attention. *Vision Research*, *49*.

Fecteau, J. H., & Munoz, D. P. (2006). Salience, relevance, and firing: a priority map for target selection. *Trends Cogn Sci*, *10*, 382–390.

Fries, P., Reynolds, J. H., Rorie, A. E., & Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*, *291*, 1560–1563.

Gottlieb, J. P., Kusunoki, M., & Goldberg, M. E. (1998). The representation of visual salience in monkey parietal cortex. *Nature*, *391*, 481–484.

Haenny, P., & Schiller, P. (1988). State dependent activity in monkey visual cortex. single cell activity in v1 and v4 on visual tasks. *Exp Brain Res*, *69*, 245–259.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, *9*, 188–194.

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, *383*, 334–337.

Huk, A. C., & Heeger, D. J. (2000). Task-related modulation of visual cortex. *J Neurophysiol*, *83*.

Ipata, A. E., Gee, A. L., & Goldberg, M. E. (2012). Feature attention evokes task-specific pattern selectivity in V4 neurons., . (pp. 16778–16785).

Itti, L., & Koch, C. (2001). Computational modeling of visual attention. *Nature Reviews Neuroscience*, *2*, 194–203.

J. F. Jehee, F. T., D. K. Brady (2011). Attention improves encoding of task-relevant features in the human visual cortex. *Journal of neuroscience*, .

James, W. (1890). MA: Harvard University Press.

Jazayeri, M., & Movshon, J. A. (2006). Optimal representation of sensory information by neural populations. *Nature Neuroscience*, *9*, 690–696.

Kanwisher, N., & Wojciulik, E. (2000). Visual attention: insights from brain imaging. *Nat Rev Neurosci*, *1*, 91–100.

Kastner, S., Pinsk, M. A., Weerd, P. D., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, *22*.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Hum Neurobiol*, *4*, 219–27.

Kowler, E. (2011). Eye movements: The past 25years. *Vision Research*, *51*, 1457–1483.

Lee, D. K., Itti, L., Koch, C., & Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nat Neurosci*, *2*, 375–381.

Ling, S., Liu, T., & Carrasco, M. (2009). How spatial and feature-based attention affect the gain and tuning of population responses. *Vision Res.*, .

Liu, T., Fuller, S., & Carrasco, M. (2006). Attention alters the appearance of motion coherence. *Psychon Bull Rev*, *13*, 1091–1096.

Lu, Z. L., & Dosher, B. A. (1998). External noise distinguishes attention mechanisms. *Vision Res*, .

Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas v1, v2, and v4 of macaque visual cortex. *J Neurophysiol*, *77*, 24–42.

Macknik, S. L., King, M., Randi, J., Robbins, A., Teller, Thompson, J., & Martinez-Conde, S. (2008). Attention and awareness in stage magic: turning tricks into research. *Nature Reviews Neuroscience*, *9*, 871–879.

Martinez, A., Anllo-Vento, L., Sereno, M. I., Frank, L. R., Buxton, R. B., Dubowitz, D. J., Wong, E. C., Hinrichs, H., Heinze, H. J., & Hillyard, S. A. (1999). Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nat Neurosci*, *2*, 364–369.

Martinez-Trujillo, J., & Treue, S. (2002). Attentional modulation strength in cortical area mt depends on stimulus contrast. *Neuron*, *35*, 365–370.

Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Curr Biol*, *14*, 744–751.

Maunsell, J., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends Neurosci.*, *29*, 317–322.

McAdams, C. J., & Maunsell, J. H. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area v4. *J Neurosci*, *19*, 431–441.

Moore, T., & Fallah, M. (2004). Microstimulation of the frontal eye field and its effects on covert spatial attention. *J Neurophysiol*, *91*, 152–162.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*, 782–4.

Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas v1, v2, and v4 in the presence of competing stimuli. *J Neurophysiol*, *70*, 909–19.

Munoz, D. P., Pelisson, D., & Guitton, D. (1991). Movement of neural activity on the superior colliculus motor map during gaze shifts. *Science*, *251*, 1358–1360.

Nakayama, K., & Martini, P. (2011). Situating visual search. *Vision Research*, *51*, 1526–1537.

Navalpakkam, V., & Itti, L. (2007). Search goal tunes visual features optimally. *Neuron*, *53*.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, .

O'Connor, D., Fukui, M., Pinsk, M., & Kastner, S. (2002). Attention modulates responses in the human lateral geniculate nucleus. *Nature Neuroscience*, *5*, 1203–1209.

O'Craven, K. M., Rosen, B. R., Kwong, K. K., Treisman, A., & Savoy, R. L. (1997). Voluntary attention modulates fMRI activity in human MT-MST. *Neuron*, *18*, 591–598.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Exp Psy.*, .

Pouget, A., Denève, S., Ducom, J.-C., & Latham, P. E. (1999). Narrow vs wide tuning curves: What's best for a population code? *Neural Computation*, .

Reynolds, J., Pasternak, T., & Desimone (2000). Attention increases sensitivity of v4 neurons. *Neuron*, .

Reynolds, J. H., Alborzian, S., & Stoner, G. R. (2003). Exogenously cued attention triggers competitive selection of surfaces. *Vision Research*, *43*, 59–66.

Reynolds, J. H., & Heeger, D. J. (2009). The normalization model of attention. *Neuron*, .

Reynolds, J. H., & Stoner, G. R. (2009). The normalization model of attention. *Neuron*, *61*, 168–185.

Roelfsema, P. R., Lamme, V. A., & Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature*, *395*, 376–381.

Saenz, M., Buracas, G. T., & Boynton, G. M. (2002). Global effects of feature-based attention in human visual cortex. *Nat Neurosci*, *5*, 631–632.

Saenz, M., Buracas, G. T., & Boynton, G. M. (2003). Global feature-based attention for motion and color. *Vision Res*, *43*, 629–637.

Saproo, S., & Serences, J. (2010). Spatial attention improves the quality of population codes in human visual cortex. *Journal of Neurophysiology*, *104*.

Schütz, A., Braun, D., & Gegenfurtner, K. (2011). Eye movements and perception: A selective review. *Journal of Vision.*, *11*, 1–30.

Scolari, M., Byers, A., & Serences, J. T. (2012). Optimal deployment of attention gain during fine discriminations. *Journal of Neuroscience*, *32*.

Scolari, M., & Serences, J. T. (2009). Adaptive allocation of attentional gain. *J. neuroscience*, .

Scolari, M., & Serences, J. T. (2010). Basing perceptual decisions on the most informative sensory neurons. *Journal of Neurophysiology*, *104*.

Serences, J. T., & Boynton, G. M. (2007). Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron*, *55*, 301–312.

Seriès, P., Latham, P. E., & Pouget, A. (2004). Tuning curve sharpening for orientation selectivity: coding efficiency and the impact of correlations. *Nature Neuroscience*, *7*, 1129–1135.

Sohn, W., Chong, S. C., Papathomas, T. V., & Vidnyánszky, Z. (2005). Cross-feature spread of global attentional modulation in human area mt+. *Neuroreport*, *16*, 1389–1393.

Spitzer, H., Desimone, R., & Moran, J. (1988). Increased attention enhances both behavioral and neuronal performance. *Science*, *240*, 338–340.

Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? *Cereb Cortex*, *17*, 679–90.

Tatler, B., Hayhoe, M., Land, M., & Ballard, D. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of vision*, *11*, 1–23.

Treisman, A., & Gelade, G. (1980a). A feature integration theory of attention. *Cognitive Psychology.*, *12*, 97–136.

Treisman, A., & Gelade, G. (1980b). A feature integration theory of attention. *Cognitive Psychology*, .

Treue, S., & Maunsell, J. H. (1996). Attentional modulation of visual motion processing in cortical areas mt and mst. *Nature*, *382*, 539–41.

Treue, S., & Trujillo, J. C. M. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, *399*, 575–579.

Tsotsos, J. K. (1992). On the relative complexity of active vs passive visual-search. *International Journal of Computer Vision*, *7*, 127–41.

Verghese, P. (2001). Visual search and attention: a signal detection theory approach. *Neuron*, *31*.

Verghese, P., Kim, Y.-J., & Wade, A. R. (2012). Attention selects informative neural populations in human v1. *The Journal of Neuroscience*, *32*, 16379–16390.

Wang, Z., Stocker, A., & Lee, D. (2012). "Optimal Neural Tuning Curves for Arbitrary Stimulus Distributions: Discrimax, Infomax and Minimum L_p Loss". In *NIPS*.

Watanabe, T., Sasaki, Y., Miyauchi, S., Putz, B., Fujimaki, N., Nielsen, M., Takino, R., & Miyakawa, S. (1998). Attention-regulated activity in human primary visual cortex. *J Neurophysiol*, *79*, 2218–2221.

Williford, T., & Maunsell, J. (2006). Effects of spatial attention on contrast response functions in macaque area V4. *J. Neurophysiol.*, *96*, 40–54.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: an alternative to the feature integration model of visual search. *J Exp Psychol [Hum Perc Perf]*, *15*, 4–433.

Womelsdorf, T., Anton-Erxleben, K., Pieper, F., & Treue, S. (2006). Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. *Nat. Neurosci.*, *9*, 1156–1160.

Womelsdorf, T., Anton-Erxleben, K., & Treue, S. (2008). Receptive field shift and shrinkage in macaque middle temporal area through attentional gain modulation. *J. Neurosci.*, *28*, 8934–8944.

Womelsdorf, T., & Fries, P. (2007). The role of neuronal synchronization in selective attention. *Current opinion in neurobiology*, *17*, 154–160.

Womelsdorf, T., Schoffelen, J.-M., Oostenveld, R., Singer, W., Desimone, R., Engel, A. K., & Fries, P. (2007). Modulation of neuronal interactions through neuronal synchronization. *science*, *316*, 1609–1612.

Yarbus, A. (1967). *Eye Movements and Vision*. Plenum Press, New York.

Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature*, *396*, 72–75.

Zhang, K., & Sejnowski, T. J. (1999). Neuronal tuning: To sharpen or broaden. *Neural Computation*, .

Zhang, Y., Meyers, E. M., Bichot, N. P., Serre, T., Poggio, T. A., & Desimone, R. (2011). Object decoding with attention in inferior temporal cortex. *PNAS*, .

5. Appendix A

Figure 8 shows the optimal attentional modulation for heterogeneous classification and visual search (i.e., a single target among two distractors) over a population of 6 neurons. Please see section 3.1.

6. Appendix B

Figure 9 shows simulation results over 12 neurons for classification and visual search. Please see section 3.1.

7. Appendix C

Figure 10 shows simulation results over 6 neurons with correlated noise in visual search. Please see the discussion section 3.1 for the explanation of results.
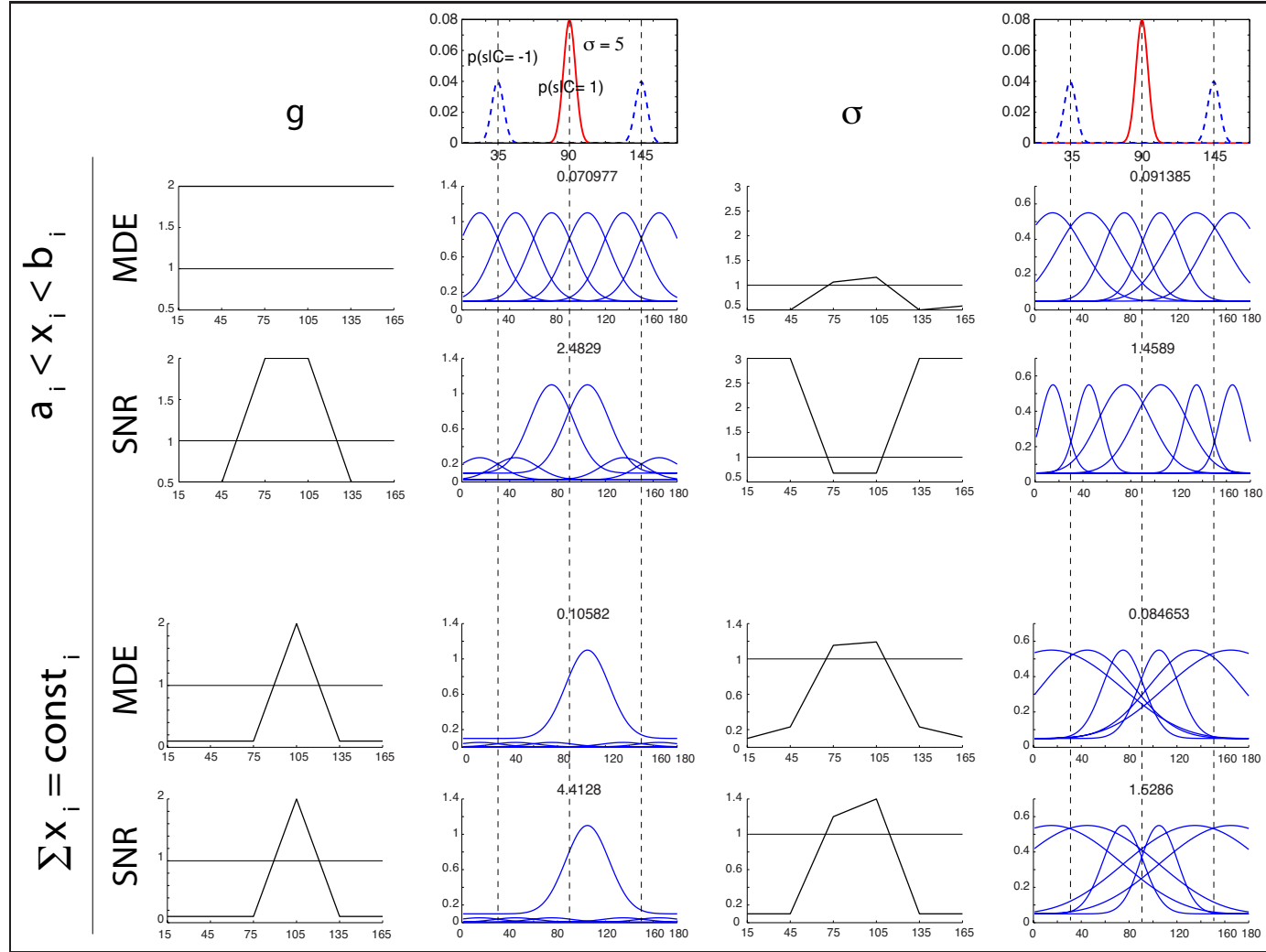
Figure 8: Optimal attentional modulation for heterogeneous classification and visual search (one target among two distractors) over a population of 6 neurons.
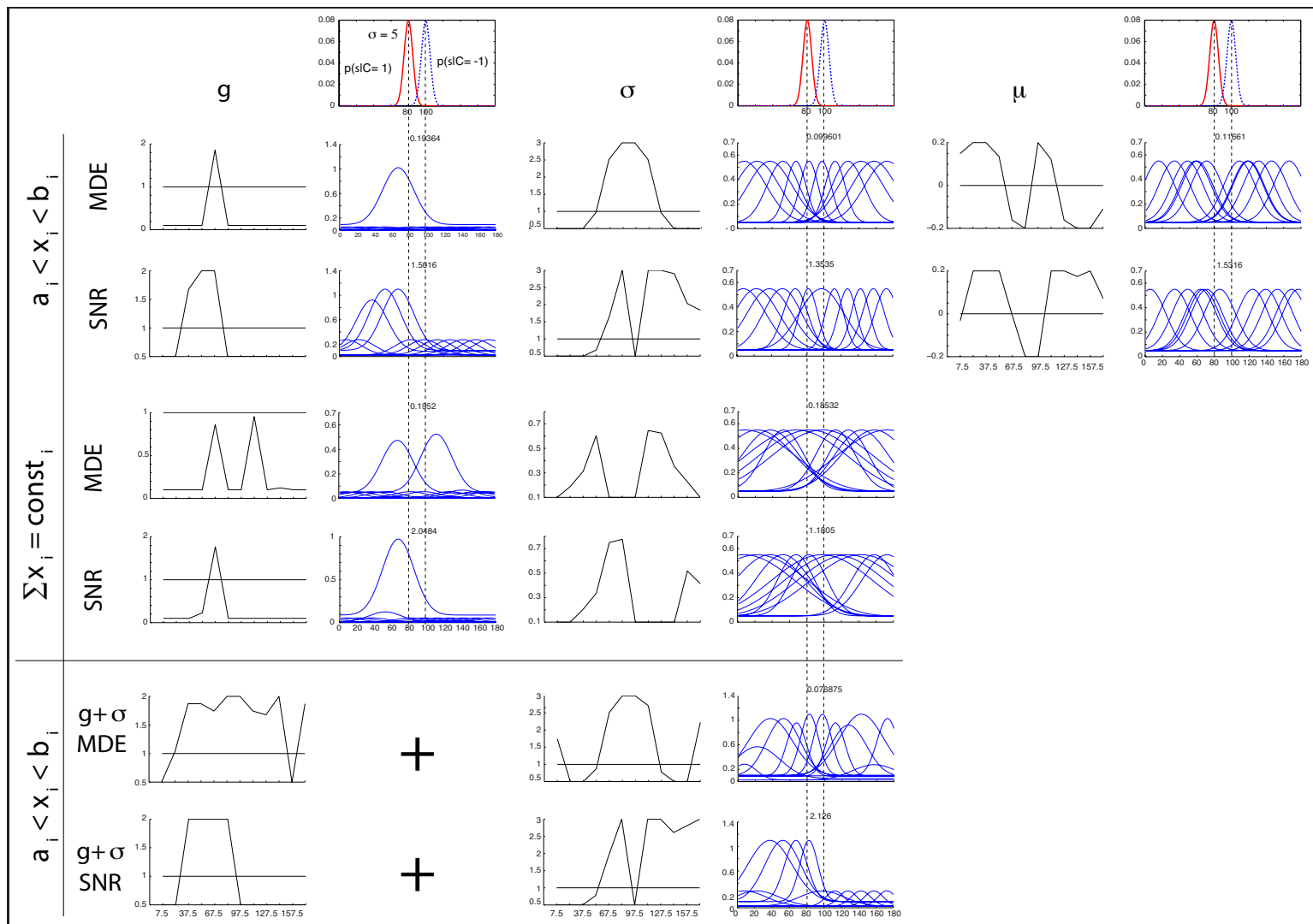
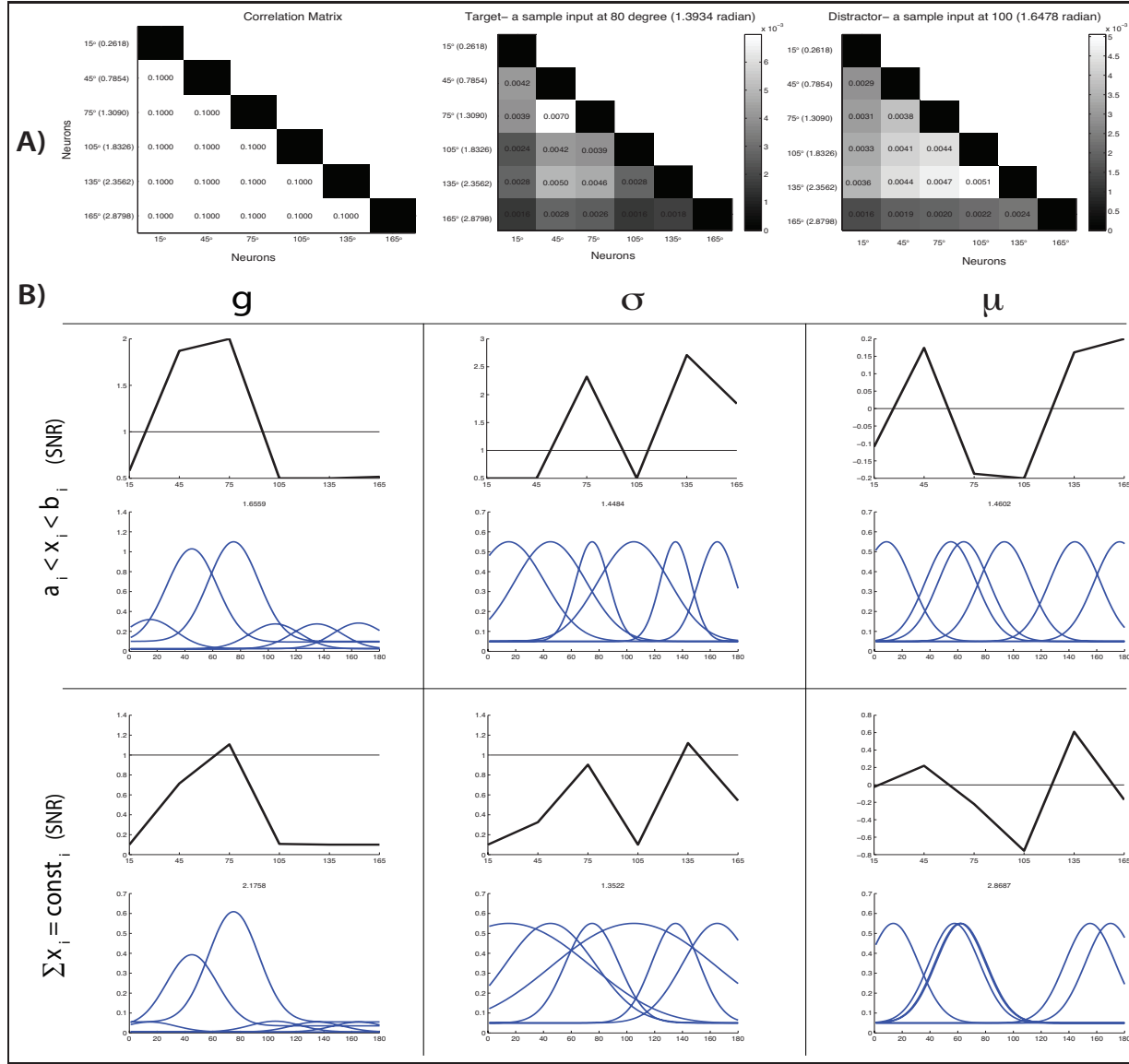Figure 9: Simulation results over 12 neurons for classification and visual search.

Figure 10: Simulation results over 6 neurons with correlated noise in visual search (target at 80, distractor at 100, $\sigma = 5$ corresponding to difficult task with low uncertainty in Fig. 2). A) Correlation coefficient matrix (for parameter c=0.1 in Berens et al. (2011)'s implementation), and population correlations for a sample draw from the target distribution at 1.3934 (radian) and a sample draw from the distractor distribution at 1.6478. Diagonal entries are set to 0 for better visualization. B) Optimal gain, tuning curve, and feature selectivity for constraint regimens 1 and 2. Compare these plots with their corresponding plots in Fig. 2.