

Neural Dissimilarity Indices That Predict Oddball Detection in Behaviour

Nidhin Koshy Vaidhiyan, S. P. Arun, and Rajesh Sundaresan, *Senior Member, IEEE*

Abstract—Neuroscientists have recently shown that images that are difficult to find in visual search elicit similar patterns of firing across a population of recorded neurons. The L^1 distance between firing rate vectors associated with two images was strongly correlated with the inverse of decision time in behavior. But why should decision times be correlated with L^1 distance? What is the decision-theoretic basis? In our decision theoretic formulation, we model visual search as an active sequential hypothesis testing problem with switching costs. Our analysis suggests an appropriate neuronal dissimilarity index, which correlates equally strongly with the inverse of decision time as the L^1 distance. We also consider a number of other possibilities, such as the relative entropy (Kullback–Leibler divergence) and the Chernoff entropy of the firing rate distributions. A more stringent test of equality of means, which would have provided a strong backing for our modeling, fails for our proposed as well as the other already discussed dissimilarity indices. However, test statistics from the equality of means test, when used to rank the indices in terms of their ability to explain the observed results, places our proposed dissimilarity index at the top followed by relative entropy, Chernoff entropy, and the L^1 indices. Computations of the different indices require an estimate of the relative entropy between two Poisson point processes. An estimator is developed and is shown to have near unbiased performance for almost all operating regions.

Index Terms—Action planning, active sensing, hypothesis testing, relative entropy, relative entropy estimation, search problems, sequential analysis, visual search.

I. INTRODUCTION

WE INVITE the reader to participate in the following visual search tasks. There are two search tasks on page 4779. Find the oddball image in each of the two configurations. Based on the time taken for each of the tasks, identify which of the two is easier.

Among the two search tasks on page 4779, most subjects find Task 1 the easier, and Task 2 the tougher. Visual search performance, as measured by the time taken to find the oddball

Manuscript received May 10, 2015; revised January 3, 2017; accepted April 11, 2017. Date of publication June 1, 2017; date of current version July 12, 2017. This work was supported by the Department of Science and Technology. This paper was presented in part at the 2012 IEEE International Symposium on Information Theory see [1], and at the 2015 Information Theory and Applications Workshop.

N. K. Vaidhiyan is with Qualcomm India Private Limited, EPIP Zone, Whitefield, Bangalore 560066, India.

S. P. Arun is with the Centre for Neuroscience, Indian Institute of Science, Bangalore 560012, India.

R. Sundaresan is with the Department of Electrical Communication Engineering and with the Robert Bosch Centre for Cyber-Physical Systems, Indian Institute of Science, Bangalore 560012, India.

Communicated by Igor Nikiforov, Associate Editor for Detection and Estimation.

Digital Object Identifier 10.1109/TIT.2017.2707485

image, should depend on the “similarity” of the two images. One has the natural hypothesis:

(H) The more “dissimilar” the two images, the shorter the time taken to find the oddball image.

To test such a hypothesis, one needs a quantification of the notion of “dissimilarity” between two images. Sripathi and Olson [3] proposed one such measure based on neuronal responses (to the images) in the inferotemporal (IT) cortex of the macaque brain. They conducted experiments to 1) find the time taken by human subjects in visual search for a number of image pairs, and 2) record neuronal responses to the same images from the monkey IT cortex. They found quantitative evidence in support of (H) based on their notion of dissimilarity. We now describe their experiments and recall their findings to set the stage for this paper.

The experiments of Sripathi and Olson [3] were the following.¹

- 1) Six human subjects were shown a picture as in Figure 1 on page 4779. Six images were placed at the vertices of a regular hexagon, with one image being different from the others. To be specific, let I_k and I_l be two images. One of these two was picked randomly with equal probability and was placed at one of the six locations randomly, again with equal probability. The other image was placed in the remaining five locations. The subjects were required to identify the correct half (left or right) of the plane where the oddball image was located. The subjects were advised to indicate their decision “as quickly as possible without guessing” [3]. The time taken to make a decision² after the onset of the image was recorded. This experiment was repeated on the same subject and across subjects. The average reaction time across trials, denoted $s(k, l)$, was recorded. Thus $s(k, l)$ is the estimate of the (symmetrised) decision time to distinguish between I_k and I_l . Similar estimates were obtained for several pairs of images.
- 2) For capturing neuronal responses to images, Sripathi and Olson conducted a set of experiments on macaque monkeys. See [3] for details. A single image I_k (respectively, I_l) was displayed on the screen, and the neuronal firings elicited by I_k (respectively, I_l) on a set of IT neurons were recorded across multiple sessions.

¹The neuronal and behavioral data used in this study was collected by one of the authors (S. P. Arun) while he was at the laboratory of Prof. Carl Olson, Carnegie Mellon University.

²A baseline motor reaction time, the time required to make a keypress response, was estimated for each subject in a separate experiment. This was then subtracted to get an estimate of the time to make a decision. See [3] for details.

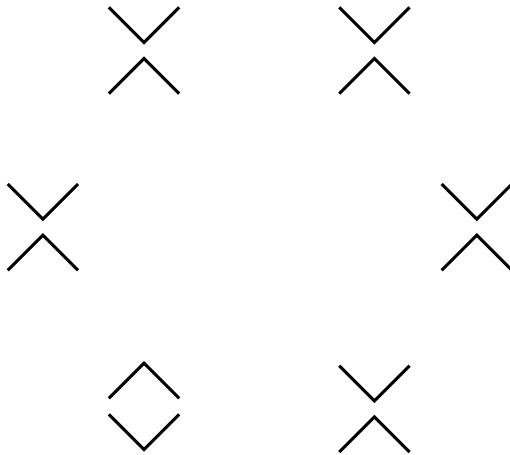


Fig. 1. Task 1.

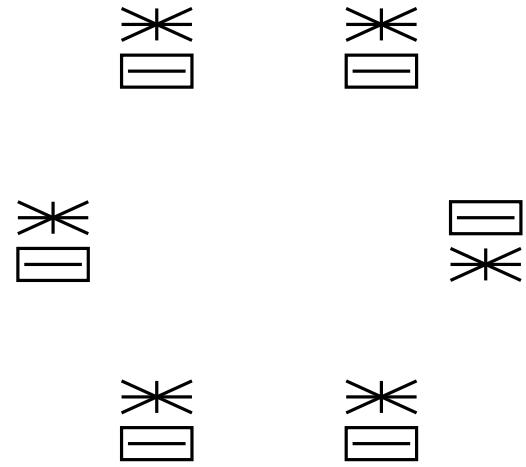


Fig. 2. Task 2.

The neuronal representation of the image I_k was taken to be the vector of average firing rates indexed by the neurons. This is denoted $\mathbf{R}_k = (R_k(1), R_k(2), \dots, R_k(d))$, where d is the number of tapped neurons. Similarly, the neuronal representation of image I_l was estimated and denoted as the vector \mathbf{R}_l . The measure of dissimilarity between the two images I_k and I_l was then taken to be the L^1 -distance normalised by the number of neurons:

$$\|\mathbf{R}_k - \mathbf{R}_l\|_1 = \frac{1}{d} \sum_{m=1}^d |R_k(m) - R_l(m)|. \quad (1)$$

They obtained the scatter plot $(s(k, l))^{-1}, \|\mathbf{R}_k - \mathbf{R}_l\|_1$ shown in Fig. 3, where (k, l) varied across image pairs, and observed a remarkably high correlation ($r = 0.95$), thereby providing evidence in support of a quantitative version of (H) .

For a detailed discussion of how neural activity in monkey visual cortex can be used to predict human search performance, we refer the reader to [3], [4]–[6].

The experiments of Sripathi and Olson [3] and Figure 3 suggest a natural question of interest to researchers in information and decision theory. One does anticipate that $s(k, l)$ is negatively correlated with some notion of dissimilarity between \mathbf{R}_k and \mathbf{R}_l , say $\text{diff}(\mathbf{R}_k, \mathbf{R}_l)$. Figure 3 suggests

$$s(k, l) \cdot \|\mathbf{R}_k - \mathbf{R}_l\|_1 = \text{constant}. \quad (2)$$

However, we know of no decision theoretic basis for $\text{diff}(\mathbf{R}_k, \mathbf{R}_l)$ to be $\|\mathbf{R}_k - \mathbf{R}_l\|_1$. What is an appropriate $\text{diff}(\mathbf{R}_k, \mathbf{R}_l)$?

Familiarity with Wald's Sequential Probability Ratio Test [7] immediately suggests that a relation like (2) should arise, but with perhaps relative entropy,³ or its variant, in place of $\|\mathbf{R}_k - \mathbf{R}_l\|_1$. A variant may be called for because of the possibility of controlled actions. To see why, let us summarize the decision problem in the form of a question:

One of the six images is odd. What would the decision center of the brain do if it got observations (firings of neurons) from the human analogue of the

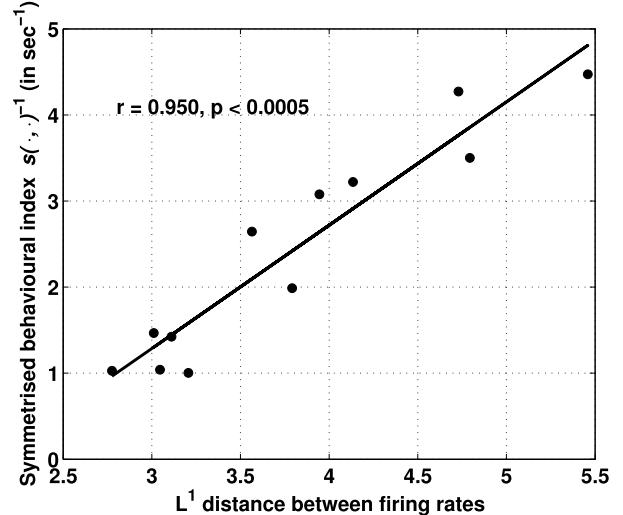


Fig. 3. Scatter plot of $(s(k, l))^{-1}, \|\mathbf{R}_k - \mathbf{R}_l\|_1$. Sripathi and Olson [3] observed a high correlation of 0.95 between the inverse of reaction time and their proposed L^1 distance between the neuronal firing vectors.

IT cortex, and could control the eye (to gaze at one of the six objects)? The goal is to minimise the time to decide the oddball image and its location, yet keep errors within desired limits.

One can model this decision problem as a sequential hypothesis testing problem with control. Naghshvar and Javidi, earlier in [8] and more recently in [9], call such a problem active sequential hypothesis testing (ASHT). ASHT suggests a natural candidate that we shall propose for $\text{diff}(\mathbf{R}_k, \mathbf{R}_l)$. There is however one important modeling issue that we wish to bring to the attention of the reader. Figure 3 shows that the average reaction times in the experiments are between 250 ms and 1000 ms. However, it is known that a switch in focus of gaze from one search location to another⁴ has a cost per switch that ranges from tens of ms to sometimes even higher than 100 ms [10]. To account for this, we extend ASHT to a setting with switching costs, and show that the $\text{diff}(\mathbf{R}_k, \mathbf{R}_l)$ appropriate for the setting

³This refers to relative entropy of the probability measure of a set of d Poisson point processes with rate vector \mathbf{R}_k taken with respect to the probability measure associated with rate vector \mathbf{R}_l .

⁴This rapid eye movement is called a saccade.

TABLE I
CORRELATION WITH DIFFERENT INFORMATION MEASURES

Information Measure	Correlation ($1/s(\cdot)$) vs. Discrimination index)	<i>p</i> -value
Proposed	0.94	5.2×10^{-12}
KL	0.93	8.5×10^{-11}
Chernoff	0.94	7.8×10^{-12}
L^1	0.94	6.1×10^{-12}

without switching costs works equally well with switching costs.

As with L^1 distance, so with our proposed $\text{diff}(\mathbf{R}_k, \mathbf{R}_j)$, and indeed, with other natural dissimilarity indices like relative entropy and Chernoff entropy,⁵ Table I indicates that all these dissimilarity measures have similar high correlation with the behavioural index.⁶ Given that all these dissimilarity indices yield high correlation with the reaction times, does our proposed diff candidate stand out in some way? It is certainly grounded in a decision-theoretic framework as we shall soon see. But is there some experimental evidence in favour of our proposed diff candidate? We address this question as well and propose a method to rank order the dissimilarity measures in their ability to explain the experimental data of Sripati and Olson [3].

A. Prior Work on the ASHT Model

Chernoff [11] studied ASHT in the context of designing optimal experiments. His performance criterion was the total cost of sampling, which is proportional to delay, plus a penalty for false detection. Chernoff proposed a policy, the so-called *Procedure A*, and showed its asymptotic optimality as the cost of sampling went to zero. *Procedure A* maintains a posterior distribution on the set of hypotheses and, at each instant, selects actions according to the hypothesis with the highest posterior probability.

There has been a flurry of recent activity extending Chernoff's work in other directions. In a series of works, Naghshvar and Javidi [8], [9], [12]–[14] studied ASHT from a Bayesian cost minimization perspective. The total cost was the sum of decision delay and a penalty for false detection. They proposed policies, similar to Chernoff's *Procedure A*, identified bounds on the total cost, and established their proposed policies' asymptotic optimality in the same asymptotic regime as Chernoff's.⁷ Nitinawarat et al. [15] studied active hypoth-

⁵Relative entropy and Chernoff entropy are possible candidates because of the following. Consider a simple hypothesis testing problem where exactly one of two images is displayed and the problem is to identify which. The stopping version of the problem corresponds to Wald's sequential hypothesis testing. The expected stopping time to meet a certain error tolerance criteria ϵ is roughly $\log(1/\epsilon)/($ relative entropy) [7]. When the decision is to be made after a fixed number of samples, where the number of samples is fixed upfront to meet a certain error tolerance criteria, the required number of samples is roughly $\log(1/\epsilon)/($ Chernoff entropy).

⁶In Table I, correlation values are based on scatter plots arising from *ordered* pairs of images. This explains why L^1 correlation value in the table (obtained from 24 points in the scatter plot) is marginally different from the correlation indicated in Figure 3 (and obtained from 12 symmetrised points in the plot).

⁷They also consider the asymptotics where the number of hypotheses is large. This is not of direct relevance to our study.

esis testing in fixed sample size and in sequential settings. They also minimize decision delay subject to a constraint on the conditional probability of false detection. When these conditional probabilities of false detection are driven to zero, the resulting asymptotic regime is the same as Chernoff's. In this asymptotic regime, they obtained results similar to those of Chernoff's but under milder assumptions. They also prove a stronger asymptotic result based on the "risk associated with a decision". Nitinawarat and Veeravalli [16] extended ASHT to Markovian observations and non-uniform costs on actions. Recently, Cohen and Zhao [17] studied anomaly detection from an ASHT perspective. They showed that, in their particular setting, a simple deterministic policy is asymptotically optimal. This is in contrast to random policies advocated in the other works. Further, for their particular setting, they showed the asymptotic optimality of Chernoff's policy under milder assumptions. Srivastava et al. [18] studied the problem of adaptive sensor selection in a binary hypothesis problem. They showed that the optimal policy would put all mass on one of the sensors and none on others. In [19] they extended their study to multiple hypothesis testing problems, where they optimized the sensor selection strategy for different objective functions, namely, min-max detection time, average detection time, etc. They proposed policies similar to Chernoff's *Procedure A*. None of the above works consider switching costs associated with a change in action.

B. Diffusion Models for Human Decision Making

Diffusion models have been used to understand human decision making in simple two alternative forced choice (TACF) tasks [20]–[23]. According to the diffusion models for a TACF task, the brain accumulates evidences against the two different choices and a decision is made when the evidence crosses a threshold. The evidence is modelled as a diffusion process with positive drift in one case and a negative drift in the alternate case. The threshold is chosen so as to satisfy the speed-accuracy tradeoff. The larger the threshold, the more accurate is the decision, but at the cost of increased decision delay. The diffusion models match well with many experimental data on TACF tasks. Our approach is also based on the notion of accumulation of evidences. But our 'active' setting is more general than the TACF setting not only because we consider multiple choices, but also because we can control the observation quality during the task and optimize costs associated with switching of actions. Action planning is an important aspect of our model. Intuitively, the decision maker should choose observations such that the drift of the accumulated evidence in favor of the correct choice is maximized. Our work differs from the above works in that we provide a decision theoretic analysis of the optimal action planning and the associated drift.

C. Our Contribution

Broadly, our contribution is a reinterpretation of the experimental results of Sripati and Olson [3] from a decision-theoretic standpoint. The following highlight some specific contributions.

- We formulate the visual search problem as an ASHT problem with switching costs. We show that a modification of Chernoff's *Procedure A*, one that we call *Sluggish Procedure A*, is asymptotically optimal even with switching costs. Further, we show that the growth rate of the total cost, as the probability of false detection is driven to zero, can be made arbitrarily close to that without switching costs.
- We propose a neuronal dissimilarity index for the diff functional in lieu of the L^1 distance between the two vectors (Sripati and Olson's proposed dissimilarity index in [3]). Our proposed dissimilarity index is based on, but is not the same as, the relative entropy between two Poisson point processes with the specified firing rate vectors.
- We test the goodness of this neuronal dissimilarity index with respect to L^1 by examining which comes closest to satisfying

$$s(k, l) \cdot \text{diff}(\mathbf{R}_k, \mathbf{R}_l) = \text{constant}, \quad (3)$$

and which is farthest. We propose a comparison statistic based on the “equality of means” testing. We use three different equality of means tests to arrive at three different statistics. The first is the familiar ANOVA's F -statistic. The second is natural too, and is the analogue of the F -statistic associated with the family of Gamma distributed random variables instead of Gaussians. The Gamma distribution, as we will later discuss, provides a better fit for the delay data. The third is similar to the second, but assumes a known shape parameter. All three methods' rankings are consistent: our proposed dissimilarity index comes out as the best, with relative entropy coming a close second, in answer to the question: Which neural dissimilarity measure based on firing rates would be optimal from a decision-theoretic point of view? We must however add a sobering note that all three equality of means tests reject, in a rather spectacular fashion, the null hypothesis of equal means in (3) at any reasonable level of statistical significance. So we emphasise that the test statistics are merely used to rank order the dissimilarity measures.

- Our estimation of the proposed neuronal dissimilarity index requires a near unbiased estimate of relative entropy as an intermediate step. We suggest a procedure to arrive at a nearly unbiased estimate. This maybe of independent value.

D. Organisation

The rest of the paper is organised as follows. Section II studies the ASHT problem with costs for switching actions. Section III applies the results of Section II to the visual search problem. Section IV develops the proposed neuronal dissimilarity index and discusses its performance through correlation studies and “equality of means” testing. Section V provides some summarising conclusions. The proofs are relegated to appendices A and B. Appendix C details the

technique used to get a near unbiased estimate of relative entropy of one Poisson point process with respect to another.

II. THE ASHT ABSTRACTION

In this section, we describe our mathematical model for visual search and collect all the relevant theoretical results. The development will be somewhat abstract. But we shall relate the model to visual search and shall apply the results to that setting in Section III. The main contribution of this section is the asymptotic growth rate of cost. In Section IV, we shall see how this suggests an appropriate diff function for plugging into (3).

A. The ASHT Model

1) *The Model Description*: Let us begin by setting up some notation.

Let $H_i, i = 1, 2, \dots, M$ denote the M hypotheses of which exactly one, denoted H , holds true. In this section, we do not assume a prior on the hypotheses. Let \mathcal{A} be the set of all possible actions which we take as finite: $|\mathcal{A}| = K < \infty$. Let \mathcal{X} be the observation space. Let $(X_n)_{n \geq 1}$ and $(A_n)_{n \geq 1}$ denote the observation process and the control process respectively. We write X^n for (X_1, \dots, X_n) and similarly A^n for (A_1, \dots, A_n) . We also write $\mathcal{P}(\mathcal{A})$ for the set of probability distributions on \mathcal{A} .

A policy π is a sequence of action plans that at time n looks at the history X^{n-1}, A^{n-1} and prescribes a composite action that is either (stop, δ) or $(\text{continue}, \lambda)$ as explained next. If the composite action is (stop, δ) , then the controller stops taking further samples (or retires) and indicates δ as its decision on the hypothesis; $\delta \in \{1, 2, \dots, M\}$. If the composite action is $(\text{continue}, \lambda)$, the controller picks the next action A_n according to the distribution $\lambda \in \mathcal{P}(\mathcal{A})$. Let $\tau(\pi)$ be the stopping time

$$\tau(\pi) := \inf\{n \geq 1 | A_n = (\text{stop}, \cdot)\}.$$

Consider a policy π . Conditioned on action A_n and the true hypothesis H , we assume that X_n is conditionally independent of previous actions $A^{n-1} = (A_1, A_2, \dots, A_{n-1})$, previous observations $X^{n-1} = (X_1, X_2, \dots, X_{n-1})$, and the policy. Let q_i^a be the conditional probability density function, with respect to some reference measure μ , of the observation X_n under action a when $H = H_i$. Let $D(q_i^a \| q_j^a)$ denote the relative entropy⁸ between the conditional probability measures associated with the observations under hypothesis H_i and under hypothesis H_j , upon action a . Denote by $\text{unif}(\mathcal{A})$ the uniform distribution on \mathcal{A} . Let $q_i^\pi(x^n, a^n)$ be the probability density function of observations and actions (x^n, a^n) till time n under policy π , with respect to the common reference measure $\mu^{\otimes n} \times \text{unif}(\mathcal{A})^{\otimes n}$. Let $Z_i^\pi(n)$ denote the log-likelihood process of hypothesis H_i , i.e.,

$$Z_i^\pi(n) = \log q_i^\pi(X^n, A^n). \quad (4)$$

⁸By an abuse of notation, we use the densities of the probability measures as the arguments of the relative entropy function.

Going forward, for ease of notation, we drop the superscript π while describing q_i^π , Z_i^π , and other variables, but their dependence on the underlying policy should be kept in mind, and the policy under consideration will be clear from the context. Define $Z(n) = (Z_1(n), Z_2(n), \dots, Z_M(n))$. Let $Z_{ij}(n)$ denote the log-likelihood ratio (LLR) process of H_i with respect to H_j , i.e.,

$$\begin{aligned} Z_{ij}(n) &= Z_i(n) - Z_j(n) \\ &= \log \frac{q_i(X^n, A^n)}{q_j(X^n, A^n)} \\ &= \sum_{l=1}^n \log \frac{q_i^{A_l}(X_l)}{q_j^{A_l}(X_l)}. \end{aligned}$$

Let E_i denote the conditional expectation and let P_i denote the conditional probability measure under $H = H_i$. (More formally, these should be represented E_i^π and P_i^π . But as done above, we omit the superscript π .)

Given an error tolerance vector $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_M)$ with $0 < \alpha_i < 1$, let $\Pi(\alpha)$ be the set of policies

$$\Pi(\alpha) = \{\pi : P_i(\delta \neq i) \leq \alpha_i, \forall i\}.$$

These are policies that meet a specified tolerance for the conditional probability of false detection. We define $\|\alpha\| := \max_i \alpha_i$.

We define λ_i to be the best mixed action that guards H_i against its nearest alternative,⁹ i.e., $\lambda_i \in \mathcal{P}(\mathcal{A})$ such that

$$\lambda_i := \arg \max_{\lambda \in \mathcal{P}(\mathcal{A})} \left[\min_{j \neq i} \sum_{a \in \mathcal{A}} \lambda(a) D(q_i^a \| q_j^a) \right]. \quad (5)$$

If there are several maximizers, pick one arbitrarily. Further, define

$$D_i := \max_{\lambda \in \mathcal{P}(\mathcal{A})} \left[\min_{j \neq i} \sum_{a \in \mathcal{A}} \lambda(a) D(q_i^a \| q_j^a) \right]. \quad (6)$$

Let $\mathcal{A}_{ij} := \{a \in \mathcal{A} : D(q_i^a \| q_j^a) > 0\}$, the set of all actions that can differentiate hypothesis H_i from hypothesis H_j . Since $D(q_i^a \| q_j^a) = 0 \Leftrightarrow D(q_j^a \| q_i^a) = 0$, we have $\mathcal{A}_{ij} = \mathcal{A}_{ji}$.

2) *Assumptions:* Throughout, we make the following assumptions.

$$(I) \quad E_i \left[\left(\log \frac{q_i^a(X)}{q_j^a(X)} \right)^2 \right] < \infty \quad \forall i, j, a.$$

(IIa) $\mathcal{A}_{ij} \neq \emptyset \quad \forall i, j$ such that $i \neq j$, and

$$(IIb) \quad \beta := \min \left\{ \sum_{a \in \mathcal{A}_{ij}} \lambda_k(a) \mid 1 \leq i, j, k \leq M, i \neq j \right\} > 0.$$

Assumption (I) implies that $D(q_i^a \| q_j^a) < \infty$, which in turn ensures that no single observation can result in a reliable decision. Assumption (I) is used in proving the lower bound on the expected number of samples needed to satisfy the tolerance criterion. This is also assumed by Chernoff [11] and Nitinawarat et al. [15].

⁹This suffices because the probability of error is dominated by the nearest alternative hypothesis.

Assumption (IIa) ensures that for any distinct i and j , there is at least one control that can help distinguish the hypotheses H_i from H_j . If $\mathcal{A}_{ij} = \emptyset$ for some i and j , it is impossible to distinguish them from each other. Assumption (IIb) is a stronger assumption than, and implies, Assumption (IIa). Assumption (IIb) ensures that if actions are taken according to any of the λ_k in (5) then, for any pair of hypotheses H_i and H_j , there is a positive probability of choosing an action that can discriminate the pair. We shall use Assumption (IIb) in the achievability proofs of our policies. It allows for easier proofs for our policies, and makes the presentation simpler. However one can work with Assumption (IIa) as well, and construct asymptotically optimal policies, with minor modifications to our policies. We will describe the needed modifications later in this section.

3) *Switching Cost and Total Cost:* The costs are as follows.

Switching Cost: Let $g(a, a')$ denote the cost of switching from action a to action a' . Throughout, we make the following additional assumptions.

$$(III) \quad g(a, a') \geq 0 \quad \forall a, a' \in \mathcal{A}, \quad g(a, a) = 0 \quad \forall a \in \mathcal{A}, \text{ and} \\ g_{\max} := \max_{a, a'} g(a, a') < \infty.$$

The assumption in the middle says no switching incurs zero cost. This assumption will play a crucial role towards our eventual conclusion that switching costs do not matter in the asymptotics considered in this paper.

Total cost: For a policy $\pi \in \Pi(\alpha)$, the total cost $C(\pi)$ is taken to be the sum of the stopping time (delay) and the net switching cost, i.e.,

$$C(\pi) := \tau(\pi) + \sum_{l=1}^{\tau(\pi)-1} g(A_l, A_{l+1}).$$

4) *Asymptotics:* We shall be interested in the asymptotics of the minimum expected total cost $E_i[C(\pi)]$, minimized over policies in $\Pi(\alpha)$, as $\|\alpha\| \rightarrow 0$. Note that there are M such conditional expected total costs, one for each hypothesis.

B. Results on the ASHT Model

We collect all the main results in this section. We first identify a lower bound.

1) *The Converse - Lower Bound:* The following proposition gives a lower bound on the conditional expectation of the stopping time, given hypothesis $H = H_i$, for all policies belonging to $\Pi(\alpha)$.

Proposition 1: Assume (I). For each i , we have

$$\liminf_{\|\alpha\| \rightarrow 0} \frac{E_i[\tau(\pi)]}{\log \|\alpha\|} \geq \frac{1}{D_i}, \quad (7)$$

where D_i is given in (6).

Proof: Since only expected time to stop is considered, proof of [11, Th. 2, p. 766] applies. \blacksquare

We then have the following corollary.

Corollary 2: Assume (I). For each i , we have

$$\liminf_{\|\alpha\| \rightarrow 0} \frac{E_i[C(\pi)]}{\log \|\alpha\|} \geq \frac{1}{D_i}. \quad (8)$$

Proof: With switching costs added, we have $C(\pi) \geq \tau(\pi)$, and the corollary follows from Proposition 1. \blacksquare

2) *Achievability - A Modification to Chernoff's Procedure A*: Chernoff [11] proposed a policy termed *Procedure A* and showed that it has asymptotically optimal expected decision delay. We now describe *Procedure A*.

Policy Procedure A: $\pi_{PA}(L)$

Fix $L > 0$.

At time n :

- Let $\theta(n) = \arg \max_i Z_i(n)$, the index with the largest log-likelihood at the current time. Ties are resolved uniformly at random.
- If $Z_{\theta(n),j}(n) < \log((M-1)L)$ for some $j \neq \theta(n)$ then A_{n+1} is chosen according to $\lambda_{\theta(n)}$, i.e.,

$$\Pr(A_{n+1} = a) = \lambda_{\theta(n)}(a). \quad (9)$$

- If $Z_{\theta(n),j}(n) \geq \log((M-1)L)$ for all $j \neq \theta(n)$ then the test retires and declares $H_{\theta(n)}$ as the true hypothesis.

We now describe a modified policy that comes arbitrarily close to being asymptotically optimal in the presence of switching costs. We introduce a switching parameter η , $0 < \eta \leq 1$, which determines the maximum transition rate out of a given action. When $\eta = 1$, we will have the original *Procedure A*. When η approaches zero, the rate of jumping out of the current action approaches zero.

Policy Sluggish Procedure A: $\pi_{SA}(L, \eta)$

Fix $L > 0$, $0 < \eta \leq 1$.

At time n :

- Let $\theta(n) = \arg \max_i Z_i(n)$. Ties are resolved uniformly at random.
- If $Z_{\theta(n),j}(n) < \log((M-1)L)$ for some $j \neq \theta(n)$ then A_{n+1} is chosen as follows.
 - Generate U_{n+1} , a Bernoulli(η) random variable, independent of all other random variables.
 - If $U_{n+1} = 0$, then $A_{n+1} = A_n$.
 - If $U_{n+1} = 1$, then generate A_{n+1} according to distribution $\lambda_{\theta(n)}$.
- If $Z_{\theta(n),j}(n) \geq \log((M-1)L)$, for all $j \neq \theta(n)$, then the test retires and declares $H_{\theta(n)}$ as the true hypothesis.

We also consider two variants of $\pi_{SA}(L, \eta)$ which are useful in the analysis.

- *Policy* $\pi_{SA}^i(L, \eta)$: This is the same as $\pi_{SA}(L, \eta)$, but stops only at decision i when $\min_{j:j \neq i} Z_{ij}(n) \geq \log(L(M-1))$.
- *Policy* $\tilde{\pi}_{SA}(\eta)$: This is the same as $\pi_{SA}(L, \eta)$, but never stops, and hence L is irrelevant.

Under a fixed hypothesis $H = H_i$, and the triplet of policies $(\pi_{SA}(L, \eta), \pi_{SA}^i(L, \eta), \tilde{\pi}_{SA}(\eta))$, it is easily seen that there is a common underlying probability measure with respect to which the processes $(X_n, A_n)_{n \geq 1}$ associated with the three policies are naturally coupled, with only the stopping times being different. Under this coupling, the following are true:

$$\begin{aligned} \tau(\pi_{SA}^i(L, \eta)) &\geq \tau(\pi_{SA}(L, \eta)), \\ \{\tau(\pi_{SA}(L, \eta)) > n\} &\subset \{\tau(\pi_{SA}^i(L, \eta)) > n\} \\ &\subset \left\{ \min_{j:j \neq i} Z_{ij}(n) < \log(L(M-1)) \right\}. \end{aligned}$$

Policy $\pi_{SA}(L, \eta)$ is designed to stop only when the posteriors suggest a reliable decision. This is formalized now.

Proposition 3: Assume (I) and (IIb). For Policy $\pi_{SA}(L, \eta)$, the conditional probability of error under hypothesis H_i is upper bounded by $P_i(\delta \neq i) \leq 1/L$.

See Appendix A.1 for a proof. As a consequence we have $\pi_{SA}(L, \eta) \in \Pi(\alpha)$ if $\alpha_i \geq 1/L$ for every i .

We now state the time-delay performance of the policy $\pi_{SA}(L, \eta)$.

Theorem 4: Assume (I) and (IIb). Consider the policy $\pi_{SA}(L, \eta)$. The conditional expected time to make a decision, for each i , satisfies

$$\lim_{L \rightarrow \infty} \frac{E_i[\tau(\pi_{SA}(L, \eta))]}{\log L} \leq \frac{1}{D_i}. \quad (10)$$

See Appendix A.2 for a detailed proof. This result will be crucial because the policy $\pi_{SA}(L, \eta)$, despite its sluggishness induced by η , remains asymptotically optimal when only the stopping time $\tau(\pi_{SA}(L, \eta))$ is considered as cost. We now leverage this to show that, if η is sufficiently small, $\pi_{SA}(L, \eta)$ is near optimal when switching costs are also taken into account.

Proposition 5: Assume (I), (IIb), and (III). Consider the policy $\pi_{SA}(L, \eta)$. We then have, for each i ,

$$\lim_{L \rightarrow \infty} E_i \left[\frac{C(\pi_{SA}(L, \eta))}{\log L} \right] \leq \frac{1}{D_i} + \frac{g_{\max} \eta}{D_i}. \quad (11)$$

Proof: We can write the following chain of inequalities.

$$\begin{aligned} E_i[C(\pi_{SA}(L, \eta))] &= E_i \left[\tau(\pi_{SA}(L, \eta)) + \sum_{l=1}^{\tau(\pi_{SA}(L, \eta))-1} g(A_l, A_{l+1}) \right] \\ &\leq E_i[\tau(\pi_{SA}(L, \eta))] + g_{\max} E_i \left[\sum_{l=1}^{\tau(\pi_{SA}(L, \eta))-1} 1_{\{A_l \neq A_{l+1}\}} \right] \\ &\leq E_i[\tau(\pi_{SA}(L, \eta))] + g_{\max} E_i \left[\sum_{l=1}^{\tau(\pi_{SA}(L, \eta))-1} U_{l+1} \right] \\ &= E_i[\tau(\pi_{SA}(L, \eta))] + g_{\max} \eta E_i[\tau(\pi_{SA}(L, \eta)) - 1] \\ &\leq E_i[\tau(\pi_{SA}(L, \eta))] (1 + g_{\max} \eta). \end{aligned} \quad (12)$$

In the above chain, the second inequality follows from Assumption (III). The penultimate equality holds because of Wald's equation [24]. Dividing by $\log L$, letting $L \rightarrow \infty$, and using Theorem 4, we see that (11) holds. ■

3) Asymptotic Optimality: Corollary 2 and Proposition 5 show that, when the conditional probability of false detection is driven to zero, the proposed policy $\pi_{SA}(L, \eta)$ has nearly the same growth rate for cost as an asymptotically optimal policy without switching costs. We now make the above statement precise. The parameter η should be suitably chosen to get sufficiently close to asymptotic optimality.

Theorem 6: Assume (I), (IIb), and (III). Consider a sequence of vectors $(\alpha^{(n)})_{n \geq 1}$, where $\alpha^{(n)}$ is the n^{th} tolerance vector, such that $\lim_{n \rightarrow \infty} \|\alpha^{(n)}\| = 0$ and

$$\lim_{n \rightarrow \infty} \frac{\|\alpha^{(n)}\|}{\min_k \alpha_k^{(n)}} < B \quad (13)$$

for some B . Then, for each n , the policy $\pi_{SA}(L_n, \eta)$ with $\log L_n = -\log \min_k \alpha_k^{(n)}$ belongs to $\Pi(\alpha^{(n)})$. Furthermore, for each i ,

$$\liminf_{n \uparrow \infty} \frac{E_i[C(\pi)]}{\log L_n} = \lim_{\eta \downarrow 0} \lim_{n \uparrow \infty} \frac{E_i[C(\pi_{SA}(L_n, \eta))]}{\log L_n} = \frac{1}{D_i}. \quad (14)$$

Proof: The fact that $\pi_{SA}(L_n, \eta) \in \Pi(\alpha^{(n)})$ is evident from Proposition 3, and $1/L_n \leq \alpha_k^{(n)}$, $k = 1, 2, \dots, M$. We then have the following chain of inequalities:

$$\begin{aligned} \frac{1}{D_i} &\leq \liminf_{n \uparrow \infty} \frac{E_i[C(\pi)]}{\log \|\alpha^{(n)}\|} \\ &= \liminf_{n \uparrow \infty} \frac{E_i[C(\pi)]}{\log L_n} \\ &\leq \lim_{\eta \downarrow 0} \lim_{n \uparrow \infty} \frac{E_i[C(\pi_{SA}(L_n, \eta))]}{\log L_n} \\ &\leq \frac{1}{D_i}. \end{aligned}$$

The first inequality follows from Corollary 2. The next equality follows from the fact that

$$\lim_{n \rightarrow \infty} \frac{\log \|\alpha^{(n)}\|}{\log L_n} = 1,$$

which in turn is true due to the assumption (13). The third inequality follows because $\pi_{SA}(L_n, \eta)$ is one specific policy in $\Pi(\alpha^{(n)})$. The last inequality follows from Proposition 5 after letting $\eta \downarrow 0$. Consequently, all inequalities must be equalities. ■

C. Discussion on Assumption (IIb)

Chernoff's result on the asymptotic optimality of *Procedure A* [11] was proved under a stronger assumption than Assumption (IIb), namely, Chernoff required

$$D(q_i^a \| q_j^a) > 0 \text{ for all } a \text{ and for all pairs } i \neq j. \quad (15)$$

Assumption (IIb) ensures that, at all times, and for any pair of hypotheses i and j , $i \neq j$, there is a positive probability of choosing an action that can distinguish the two hypotheses. This suffices for Chernoff's proofs to go through. Specifically, we shall use Assumption (IIb) to prove the exponential decay result in Proposition 13 of Appendix A. Nitinawarat et al. [15] proposed a modified *Procedure A* that sampled actions randomly at intervals $\lceil v^l \rceil_{l \geq 1}$, $v > 1$, and showed that their proposed policy is asymptotically optimal under the weaker Assumption (IIa). The random sampling enabled them to obtain a polynomial decay counterpart of Proposition 13 of Appendix A. Recently, Cohen and Zhao [17] showed the asymptotic optimality of *Procedure A* under the weaker Assumption (IIa) for an active anomaly detection problem, which is a specific ASHT problem. We conjecture that Chernoff's *Procedure A* is asymptotically optimal under the weaker Assumption (IIa) for all ASHT problems. A proof of this claim has remained elusive. Nevertheless, policies whose performances are provably arbitrarily close to the optimum can be designed. We make the above claim precise in the next proposition.

Proposition 7: Assume (I) and (IIa). Fix $\epsilon > 0$. Then there exists a sequence of policies $\{\pi_\epsilon(L)\}$ that satisfies $\pi_\epsilon(L) \in \Pi(\frac{1}{L}, \frac{1}{L}, \dots, \frac{1}{L})$ and

$$\lim_{L \rightarrow \infty} E_i \left[\frac{\tau(\pi_\epsilon(L))}{\log L} \right] \leq \frac{1}{(1 - \epsilon)D_i}. \quad (16)$$

We omit the proof because the needed modifications to the proof of Theorem 4 are straightforward. Policy $\{\pi_\epsilon(L)\}$ can be constructed as a variant of *Procedure A* that, at each instant n , chooses an action according to $\text{unif}(\mathcal{A})$ with probability ϵ or according to (9) with probability $(1 - \epsilon)$. Note that, at each time n , the modified policy $\{\pi_\epsilon(L)\}$ uses a randomisation on the actions of the form $\tilde{\lambda}_{\theta(n)} = (1 - \epsilon)\lambda_{\theta(n)} + \epsilon\text{unif}(\mathcal{A})$. It can be shown that, under hypothesis H_i , $\theta(n) = i$ in finite time with probability 1, and thereby the asymptotic log likelihood ratio rate between H_i and any other H_j will be lower bounded by $(1 - \epsilon)D_i$. Thus, at the cost of a small penalty, we can design nearly asymptotically optimal policies under the weaker Assumption (IIa). A similar argument holds true with switching costs, just as Theorem 4 is extended in Theorem 6, albeit with a corresponding but arbitrarily small increase in the total cost. Again, we omit the proof of this claim with switching costs. The conclusion is that Assumption (IIa) suffices for the asymptotic growth rate to be $\frac{1}{D_i}$.

III. BACK TO VISUAL SEARCH

We now return to the visual search problem. In the visual search task, a subject has to identify an oddball image from amongst W images displayed on a screen ($W = 6$ in Figures 1 - 2). For the purpose of modeling, we make the following assumptions. The subject can focus attention on only one of the W positions, and the field of view is restricted to the image at that position alone. Further, we assume that time is slotted¹⁰ and each slot is of duration T . The subject can change the focus of his attention to any of the W image locations, but only at the slot boundaries. A switch in focus of attention (saccade) requires an integer number of slots for the operation, and no sensing is possible during such a saccade. The lost time during saccades are modeled as switching costs (delays), and hence the total decision time is the sum of sensing delay and switching delays. We assume that the subject would have indeed found the exact location and identity of the oddball image before mapping it to a “left” or “right” decision. These are clearly oversimplifying assumptions, but enable easier analysis and provide valuable insights.

If the image in the focused location is I_k , we assume that a set of d neurons react accordingly to produce spike trains. These constitute the observations. Specifically, these are modeled as d independent Poisson point processes of duration T with rates given by the components of the rate vector $\mathbf{R}_k = (R_k(1), R_k(2), \dots, R_k(d))$. More formally, let \mathcal{X} be the space of counting processes in $[0, T]$ with an associated σ -algebra. Let $\mu_{1,T}$ be the standard rate 1 Poisson point process and let $\mu_{1,T}^{\otimes d}$ be its d -fold product measure. Let $\mu_{\mathbf{R}_k, T}$ denote

¹⁰One could also consider an extension to the continuous-time setting. But all essential ideas are best described in the slotted setting.

the probability measure P_k , so that density of $\mu_{\mathbf{R}_k, T}$ with respect to $\mu_{1, T}^{\otimes d}$ is given by

$$f_k := \frac{d\mu_{\mathbf{R}_k, T}}{d\mu_{1, T}^{\otimes d}},$$

with a similar definition for f_l corresponding to image I_l .

We now describe two possible settings. Case 2 will turn out to be closer to the experiment of Sripati and Olson [3].

Case 1: The subject has knowledge that the oddball image is I_k and that the distractors are I_l . Since there are W locations, and $1 \leq i \leq W$, there are W hypotheses.

The visual search problem under Case 1 can be formulated as an ASHT problem as follows.

- Hypotheses: H_i is the hypothesis that the oddball image (I_k) is at location i , $1 \leq i \leq W$.
- Actions: The subject may focus on any one of the W locations, and so $\mathcal{A} = \{1, 2, \dots, W\}$.
- Observations: The conditional probability density function q_i^a of the observations, under hypothesis H_i and when action a is chosen, is:

$$q_i^a = \begin{cases} f_k & \text{if } a = i \\ f_l & \text{if } a \neq i. \end{cases}$$

In words, under Hypothesis H_i , the oddball image is I_k and is at location i . If the action is to focus on location i , i.e., $a = i$, then the subject views the oddball image I_k , and so the observations have density f_k . If $a \neq i$, then the subject views the distractor image I_l , and so the observations have density f_l .

The relative entropies for the various combinations of hypotheses pairs (i, j) , with $i \neq j$, and actions are as follows:

$$D(q_i^a \| q_j^a) = \begin{cases} D(f_k \| f_l) & a = i \\ D(f_l \| f_k) & a = j \\ 0 & a \neq i, a \neq j. \end{cases} \quad (17)$$

Proposition 8: For the setting of Case 1, the λ_i and D_i of (5) and (6), respectively, are as follows.

If $D(f_k \| f_l) > D(f_l \| f_k)/(W - 1)$ then

$$\lambda_i(i) = 1, \quad \lambda_i(j) = 0 \quad \forall j \neq i, \quad \text{and } D_i = D(f_k \| f_l).$$

If $D(f_k \| f_l) < D(f_l \| f_k)/(W - 1)$ then

$$\lambda_i(i) = 0, \quad \lambda_i(j) = \frac{1}{(W - 1)} \quad \forall j \neq i, \quad \text{and } D_i = \frac{D(f_l \| f_k)}{(W - 1)}.$$

If $D(f_k \| f_l) = D(f_l \| f_k)/(W - 1)$ then

$$\lambda_i(i) = c, \quad \lambda_i(j) = \frac{1 - c}{(W - 1)} \quad \forall j \neq i, \quad \text{and any } 0 \leq c \leq 1$$

$$\text{and } D_i = \frac{D(f_l \| f_k)}{(W - 1)}.$$

Proof: We can upper bound D_i as follows:

$$D_i = \max_{\lambda \in \mathcal{P}(\mathcal{A})} \min_{j \neq i} \sum_{a \in \mathcal{A}} \lambda(a) D(q_i^a \| q_j^a) \quad (18)$$

$$= \max_{\lambda \in \mathcal{P}(\mathcal{A})} \min_{j \neq i} [\lambda(i) D(f_k \| f_l) + \lambda(j) D(f_l \| f_k)] \quad (19)$$

$$= \max_{\lambda \in \mathcal{P}(\mathcal{A})} \left[\lambda(i) D(f_k \| f_l) + \min_{j \neq i} \lambda(j) D(f_l \| f_k) \right] \quad (20)$$

$$\leq \max_{\lambda \in \mathcal{P}(\mathcal{A})} \left[\lambda(i) D(f_k \| f_l) + \frac{1 - \lambda(i)}{W - 1} D(f_l \| f_k) \right] \quad (21)$$

$$= \begin{cases} D(f_k \| f_l) & \text{if } D(f_k \| f_l) > \frac{D(f_l \| f_k)}{W - 1}, \\ & \text{by setting } \lambda(i) = 1, \\ D(f_k \| f_l) & \text{if } D(f_k \| f_l) = \frac{D(f_l \| f_k)}{W - 1}, \\ & \text{by setting } \lambda(i) = c, \quad 0 \leq c \leq 1, \\ \frac{D(f_l \| f_k)}{W - 1} & \text{if } D(f_k \| f_l) < \frac{D(f_l \| f_k)}{W - 1}, \\ & \text{by setting } \lambda(i) = 0. \end{cases} \quad (22)$$

Here (19) follows from (17), (20) follows after taking the minimisation inside, (21) follows because the minimum of a set of numbers is upper bounded by their arithmetic mean, and (22) follows by maximising the linear objective function in (21). Finally, (21) can be made an equality by choosing all $\lambda_j, j \neq i$ to be identical. This proves the Proposition. ■

Thus, under H_i , to distinguish H_i from its nearest alternative, one either focuses only at the oddball location or at any of the other locations with equal probability depending on whether $D(f_k \| f_l) > D(f_l \| f_k)/(W - 1)$ or not.

Case 2: The subject has knowledge of the two competing images I_k and I_l , but does not know which of the two is the oddball image.

This visual search problem can be formulated as a $2W$ hypothesis testing problem as follows.

- Hypotheses:

H_i with $i \leq W$: The oddball image is I_k and is at location i . All other locations have image I_l .

H_i with $i > W$: The oddball image is I_l and is at location $i - W$. All other locations have image I_k .

- Actions: The subject can focus on any one of the W locations, and so $\mathcal{A} = \{1, 2, \dots, W\}$.

- Observations: The conditional probability density function q_i^a of the observations, under hypothesis H_i and when action a is chosen, is:

$$q_i^a = \begin{cases} f_k & i \leq W, a = i \\ f_l & i \leq W, a \neq i \\ f_l & i > W, a = i - W \\ f_k & i > W, a \neq i - W. \end{cases}$$

In words, under Hypothesis H_i with $i \leq W$, the oddball image is I_k and is at location i . If the action is to focus on location i , i.e., $a = i$, then the subject views image I_k and so the observations have density f_k corresponding to I_k . The outcome of other actions for this hypothesis are explained similarly. An analogous description holds for outcomes of actions under H_i when $i > W$.

The relative entropies for the various combinations of hypotheses pairs $(i \neq j)$ and actions are as follows. The expressions are self-explanatory.

$$(i) \quad i \leq W, \quad j \leq W : \quad (23)$$

$$D(q_i^a \| q_j^a) = \begin{cases} D(f_k \| f_l) & a = i \\ D(f_l \| f_k) & a = j \\ 0 & a \neq i, a \neq j. \end{cases} \quad (24)$$

$$(ii) \quad i \leq W, \quad j = i + W : \quad (25)$$

$$D(q_i^a \| q_j^a) = \begin{cases} D(f_k \| f_l) & a = i \\ D(f_l \| f_k) & a \neq i. \end{cases} \quad (26)$$

$$(iii) \quad i \leq W, \quad j > W, \quad j \neq i + W: \quad (27)$$

$$D(q_i^a \| q_j^a) = \begin{cases} 0 & a = i \\ 0 & a = j - W \\ D(f_l \| f_k) & a \neq i, a \neq j - W. \end{cases} \quad (28)$$

(iv) For $i > W$, the expressions for $j > W$, $j = i - W$, or $j < W$ but $j \neq i - W$ are similar to (i), (ii), and (iii) above, respectively, but with f_k and f_l interchanged.

We now identify the structure of λ_i and D_i for $i = 1, 2, \dots, 2W$. We refer the reader to Appendix B for a proof.

Proposition 9: Let $W \geq 3$. Let $i \leq W$. For the setting of Case 2, the optimum λ_i and D_i of (5) and (6), respectively, are as follows. If $D(f_k \| f_l) > D(f_l \| f_k)/(W-1)$ then

$$\begin{aligned} \lambda_i(i) &= \frac{(W-3)D(f_l \| f_k)}{(W-1)D(f_k \| f_l) + (W-3)D(f_l \| f_k)}, \\ \lambda_i(j) &= \frac{D(f_k \| f_l)}{(W-1)D(f_k \| f_l) + (W-3)D(f_l \| f_k)} \quad \forall j \neq i, \end{aligned}$$

and

$$D_i = \frac{(W-2)D(f_k \| f_l)D(f_l \| f_k)}{(W-1)D(f_k \| f_l) + (W-3)D(f_l \| f_k)}. \quad (29)$$

If $D(f_k \| f_l) < D(f_l \| f_k)/(W-1)$ then

$$\lambda_i(i) = 0, \quad \lambda_i(j) = \frac{1}{(W-1)} \quad \forall j \neq i, \quad \text{and} \quad D_i = \frac{D(f_l \| f_k)}{(W-1)}.$$

If $D(f_k \| f_l) = D(f_l \| f_k)/(W-1)$ then

$$\begin{aligned} \lambda_i(i) &= c, \quad \text{for any } 0 \leq c \leq \frac{W-3}{W-2}, \\ \lambda_i(j) &= \frac{1-c}{(W-1)} \quad \forall j \neq i, \quad \text{and} \\ D_i &= \frac{D(f_l \| f_k)}{(W-1)}. \end{aligned}$$

For $i > W$, λ_i and D_i have the same structure as above, but with f_l and f_k interchanged.

IV. PROPOSAL FOR A NEURONAL DISSIMILARITY INDEX

We now apply the results obtained in the previous section to the data from the experiments of Sripathi and Olson [3]. The visual search experiments of Sripathi and Olson [3] on human subjects correspond closely with Case 2 of the previous section. Similar to Case 2, the subjects in the experiments had no prior information on which of the two images I_k and I_l was the oddball image and which the distractor. But different from Case 2, the subjects in the experiments had to learn about the images I_k and I_l on-the-go, while in Case 2 we assume that the subject knows that the oddball and distractor images come from the set $\{I_k, I_l\}$. A more accurate modeling that takes the learning aspect into account is available in [30]. Here, we shall proceed with the Case 2 model.

Recall that T is the slot duration during which the subject focuses attention on a particular image. First, we calculate the relative entropy $D(f_k \| f_l)$ when f_k and f_l are densities of vector Poisson point processes of duration T with rates $\mathbf{R}_k = (R_k(1), R_k(2), \dots, R_k(d))$ and $\mathbf{R}_l = (R_l(1),$

$R_l(2), \dots, R_l(d))$. Under the assumption that the neurons fire independently with the specified rates,¹¹ the relative entropy decomposes into a sum:

$$\begin{aligned} &D(\mu_{\mathbf{R}_k, T} \| \mu_{\mathbf{R}_l, T}) \\ &= E_{\mu_{\mathbf{R}_k, T}} \left[\log \frac{d\mu_{\mathbf{R}_k, T}}{d\mu_{\mathbf{R}_l, T}} \right] \\ &= \sum_{m=1}^d E_{\mu_{\mathbf{R}_k(m), T}} \left[\log \frac{d\mu_{\mathbf{R}_k(m), T}}{d\mu_{\mathbf{R}_l(m), T}} \right] \\ &= T \sum_{m=1}^d \left[R_k(m) \log \left(\frac{R_k(m)}{R_l(m)} \right) - R_k(m) + R_l(m) \right], \end{aligned}$$

where the term within square brackets in the last summation is the relative entropy of the Poisson point processes with rate $R_k(m)$ taken with respect to another such process with rate $R_l(m)$.

In Case 2, if the number of locations $W = 6$, if I_k is the oddball image, if I_l is the distractor image, and if $D(\mu_{\mathbf{R}_k, T} \| \mu_{\mathbf{R}_l, T}) > D(\mu_{\mathbf{R}_l, T} \| \mu_{\mathbf{R}_k, T})/(W-1)$ which is the case when $D(\mu_{\mathbf{R}_l, T} \| \mu_{\mathbf{R}_k, T})$ is close to $D(\mu_{\mathbf{R}_k, T} \| \mu_{\mathbf{R}_l, T})$, then from Proposition 9 we have

$$D_{kl} = \frac{4D(\mu_{\mathbf{R}_k, T} \| \mu_{\mathbf{R}_l, T})D(\mu_{\mathbf{R}_l, T} \| \mu_{\mathbf{R}_k, T})}{5D(\mu_{\mathbf{R}_k, T} \| \mu_{\mathbf{R}_l, T}) + 3D(\mu_{\mathbf{R}_l, T} \| \mu_{\mathbf{R}_k, T})}. \quad (30)$$

Similarly, if I_l is the oddball image, if I_k is the distractor image, and if $D(\mu_{\mathbf{R}_l, T} \| \mu_{\mathbf{R}_k, T}) > D(\mu_{\mathbf{R}_k, T} \| \mu_{\mathbf{R}_l, T})/(W-1)$, we have

$$D_{lk} = \frac{4D(\mu_{\mathbf{R}_l, T} \| \mu_{\mathbf{R}_k, T})D(\mu_{\mathbf{R}_k, T} \| \mu_{\mathbf{R}_l, T})}{5D(\mu_{\mathbf{R}_l, T} \| \mu_{\mathbf{R}_k, T}) + 3D(\mu_{\mathbf{R}_k, T} \| \mu_{\mathbf{R}_l, T})}. \quad (31)$$

Let us normalize D_{kl} per unit time and per neuron and denote it \tilde{D}_{kl} :

$$\tilde{D}_{kl} = \frac{1}{dT} D_{kl}. \quad (32)$$

The subset of experimental data gathered by Sripathi and Olson that we use in our analysis consisted of the following.

1) Neuronal firing rate vectors were obtained from the IT cortex of rhesus macaque monkeys for twenty four images. The number of neurons ranged from 78 to 174, the variation was due to experimental constraints. But the sets of neurons tapped were identical for images that were to be paired in the decision time experiments on human subjects, which we describe next.

2) Decision time statistics for detection of the oddball image were obtained from experiments on human subjects. For oddball image I_k and distractors I_l , data was collected as follows. Six subjects participated and each was shown twelve stimuli. In each stimulus, the oddball location was picked uniformly at random from the $W = 6$ locations. The decision

¹¹As the tapped neurons constituted a sparse sampling of the neurons of the infero-temporal cortex, we assume that the neurons spike independently of each other. The independence assumption enables easy computation of the neuronal metric. That the numerical predictions made with this assumption match up with the experimental data suggests that the independence assumption is a good one. From a conceptual stand-point, however, the basic theoretical results of the paper continue to hold even without the independence assumption.

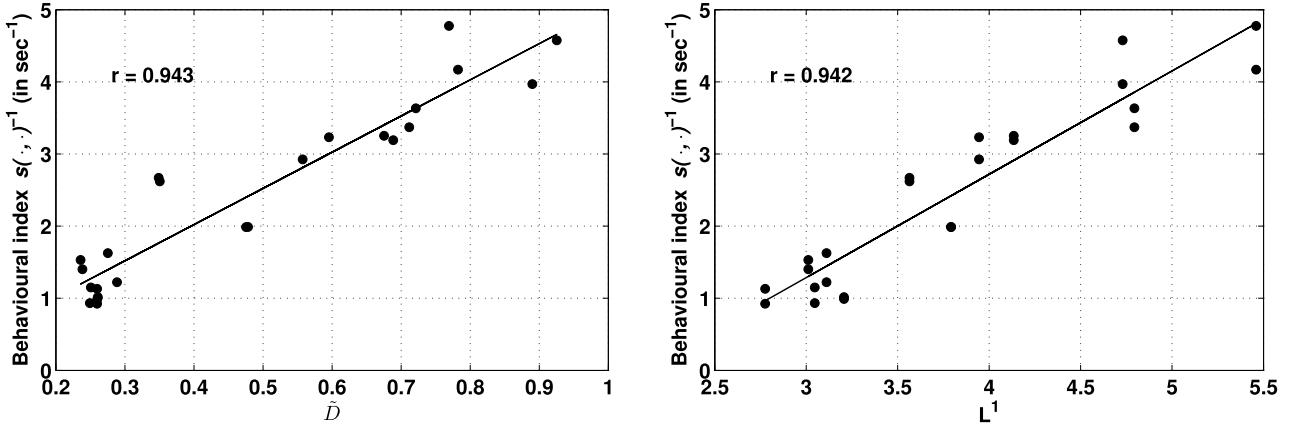


Fig. 4. The observed behavioural dissimilarity index versus the proposed neuronal dissimilarity index (\tilde{D}) and the L^1 -neuronal index.

time were averaged across various subjects and across stimuli instances to get $s(k, l)$. The first argument k stands for the oddball image I_k .

Recall from Case 2 that H_i , when $i \leq W = 6$, is the hypothesis that the oddball image is I_k and the distractor images are I_l . Taking a cue from Theorem 6, assuming a sufficiently stringent error tolerance vector of $(\alpha, \alpha, \dots, \alpha)$ for α sufficiently small, and assuming nearly optimal decision making, we predict that

$$E_i[C(\pi)] \approx \frac{\log(1/\alpha)}{D_{kl}},$$

where $C(\pi)$ models the total decision time, the sum of sensing delay and switching delays. Averaging across $i = 1, \dots, W$, i.e., averaging across all those stimuli where I_k is the oddball image and I_l is the distractor, we get

$$\begin{aligned} E[C(\pi) \mid I_k \text{ is the oddball and } I_l \text{ is the distractor}] &\approx \frac{\log(1/\alpha)}{D_{kl}} \\ &\approx \frac{\log(1/\alpha)}{\tilde{D}_{kl}} \\ &= \frac{(1/dT) \log(1/\alpha)}{\tilde{D}_{kl}}, \end{aligned}$$

or in other words

$$s(k, l) \cdot \tilde{D}_{kl} \approx \text{constant}.$$

For $i > W = 6$, one similarly has

$$s(l, k) \cdot \tilde{D}_{lk} \approx \text{constant}.$$

This naturally leads to the proposal

$$\text{diff}(\mathbf{R}_k, \mathbf{R}_l) = \tilde{D}_{kl}. \quad (33)$$

A. Correlation Study

The *behavioural dissimilarity index* for an ordered pair of images (k, l) is the inverse $s(k, l)^{-1}$ of the average decision time $s(k, l)$, and gives an indication of the speed of discrimination. In Figure 4, we plot the behavioural dissimilarity index $s(k, l)^{-1}$ against the proposed neuronal dissimilarity index \tilde{D}_{kl} and against the L^1 dissimilarity index for various ordered pairs (k, l) . We observe a strong correlation of 0.94 for \tilde{D}_{kl}

TABLE II

CORRELATION WITH DIFFERENT NEURONAL DISSIMILARITY INDICES

Information Measure	Correlation ($1/s$ vs. Neuronal index)	p -value	Correlation (s vs Neuronal index $^{-1}$)	p -value
Proposed \tilde{D}	0.94	5.2×10^{-12}	0.89	4.3×10^{-09}
KL	0.93	8.5×10^{-11}	0.90	3.1×10^{-09}
Chernoff	0.94	7.8×10^{-12}	0.88	2.1×10^{-08}
L^1	0.94	6.1×10^{-12}	0.88	1.1×10^{-08}

which is the same as the correlation between the behavioural dissimilarity index and the L^1 distance $\|\mathbf{R}_k - \mathbf{R}_l\|_1$.

Now that we have discovered that our proposed neuronal dissimilarity index and the L^1 index are both equally well-correlated with the behavioural dissimilarity index, it is natural to ask if there is some basis to choose one over the other. The point that our proposed dissimilarity index has a “microscopic basis” (grounded in decision theory and based on the ASHT framework) that explains the “macroscopic observations” (speed of discrimination) is certainly in our favour. But there are other related dissimilarity indices such as relative entropy (KL) and Chernoff entropy that have similar correlation with the behavioural dissimilarity index. Table II summarises the correlations (second column) along with their p -values (third column). It is therefore natural to ask if a finer examination of the experimental data can help us identify the “best” among these neuronal indices. We shall pursue this in the next subsection and shall propose a method to rank order the indices in terms of their ability to explain the experimental data.

A more basic question, and one that is motivated by our expectation that $s(k, l) \cdot \text{diff}(\mathbf{R}_k, \mathbf{R}_l) = \text{constant}$, is whether it is more appropriate to correlate $s(k, l)$ versus $\text{diff}(\mathbf{R}_k, \mathbf{R}_l)^{-1}$ as opposed to what is done in Figure 4 which correlates $s(k, l)^{-1}$ versus $\text{diff}(\mathbf{R}_k, \mathbf{R}_l)$. Table II reports these correlations (fourth column) along with the corresponding p -values (fifth column), and Figure 5 provides the correlation plot. The new correlations, though high, are lower than those reported in the second column.

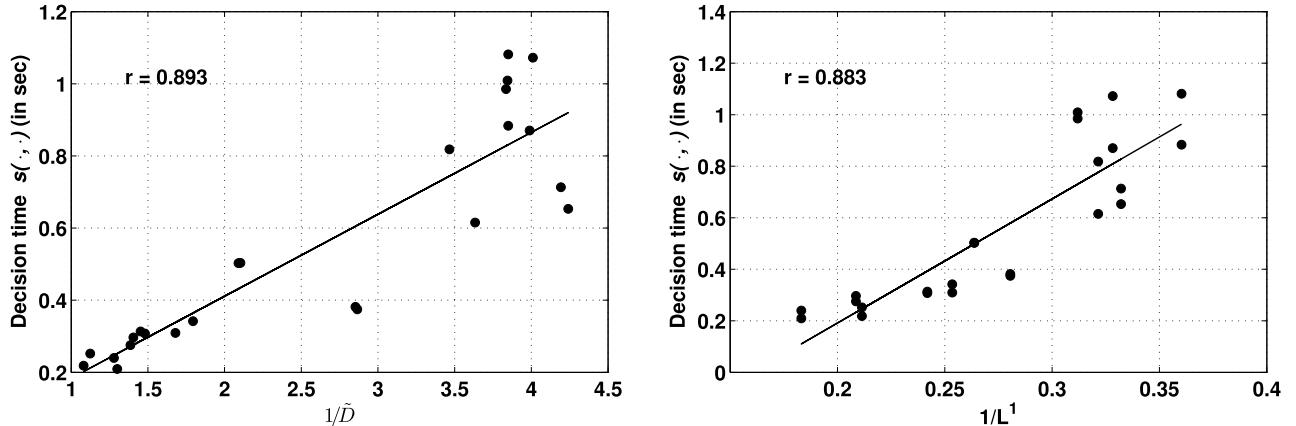


Fig. 5. The observed decision time versus the inverse of the proposed neuronal index ($1/\tilde{D}$) and the inverse of the L^1 -neuronal index.

We do not have a clear-cut answer on which of the two scatter plots –

$$(s(k, l), \text{diff}(\mathbf{R}_k, \mathbf{R}_l)^{-1}) \text{ or } (s(k, l)^{-1}, \text{diff}(\mathbf{R}_k, \mathbf{R}_l)) \quad (34)$$

– and the corresponding correlations is more appropriate. However, recall that Pearson's test for rejecting the null hypothesis that a bivariate normal has independent components is that the correlation statistic r arising from independent and identically distributed samplings of the bivariate normal has $|r|$ exceeding a threshold. Given that $s(k, l)$ is the arithmetic mean of $n = 72$ experimentally measured decision time, when centred and scaled, $s(k, l)$ is likely to be closer to normal than its inverse. We therefore believe the correlation of $(s(k, l), \text{diff}(\mathbf{R}_k, \mathbf{R}_l)^{-1})$, the one that leads to lower correlations, is more appropriate. The indicated p -values, shown in Figures 4 and 5 and in Tables I and II, are the probabilities that the correlation statistic equals or exceeds the indicated observed levels when the null hypothesis is true (independent components).

B. Model Testing via Three “Equality of Means” Tests

In Section I, we posed the question of identifying a suitable diff function that satisfies

$$s(k, l) \cdot \text{diff}(\mathbf{R}_k, \mathbf{R}_l) = \text{constant}. \quad (35)$$

In the previous section, we modeled visual search as an ASHT problem and proposed the diff given in (33), denoted \tilde{D} . However, we also saw that the candidates L^1 , Chernoff entropy, relative entropy, and \tilde{D} , all yielded high correlation with the behavioural dissimilarity index. We now address the question of which of these dissimilarity indices best explain the data.

Our methodology is as follows. Consider a fixed $\text{diff}(\mathbf{R}_k, \mathbf{R}_l)$ function. Let us test the new null hypothesis:

$$(H_0) : E[C(\pi) \mid I_k \text{ is the oddball and } I_l \text{ is the distractor}] \times \text{diff}(\mathbf{R}_k, \mathbf{R}_l) = \text{constant},$$

where $C(\pi)$ is the decision time for a fixed error tolerance on the ordered image pair (I_k, I_l) . The decision time data across subjects and across multiple stimuli that have I_k as the oddball and I_l as the distractor images constitute one group

associated with the ordered pair (k, l) . H_0 hypothesises that the diff -scaled means is constant across groups. Let us identify the diff indices for which the corresponding null hypothesis is accepted for a desired significance level. If the test passes for $\text{diff}(\mathbf{R}_k, \mathbf{R}_l) = \tilde{D}_{kl}$, then there is significant evidence that the data is well-explained by our theory.

To perform this test, we must do the following for each diff candidate.

- Identify a test statistic $\mathcal{T}(\text{diff})$ for testing equality of means of the diff -scaled decision times. Note each diff leads to a separate hypothesis test.
- Accept or reject the corresponding null hypothesis for a desired level of significance.

Let $\tau_{k,l}(j)$ be the j th sample in the group indexed by (k, l) . Let n denote the common number of samples in each group, and let g be the number of groups. The experimental data of Sripathi and Olson had 24 groups and 72 samples per group; $n = 72$ and $g = 24$. The number of samples in each group was identical.

1) *Test 1 - Oneway ANOVA*: : The one-way analysis of variance (ANOVA) statistic [25] is often used to test equality of means across groups when the samples are Gaussian and when the variances across groups are the same. This test is known to be robust to the Gaussian assumption. It is also known to be robust to the equality of variances assumption so long as the number of samples is the same across groups [26, p.243]. As we will soon see, we neither have Gaussianity nor equality of variances across groups. But since the number of samples is the same across the groups, we may still use the oneway ANOVA test.

Let $T_{k,l}(j) := \tau_{k,l}(j) \cdot \text{diff}(\mathbf{R}_k, \mathbf{R}_l)$. Write the sample means, the mean across groups, and the pooled variance as follows.

$$\begin{aligned} \bar{T}_{k,l} &= \frac{1}{n} \sum_{j=1}^n T_{k,l}(j), \\ \bar{\bar{T}} &= \frac{1}{g} \sum_{(k,l)} \bar{T}_{k,l}, \\ S_p^2 &= \frac{1}{g(n-1)} \sum_{(k,l)} \sum_{j=1}^n (T_{k,l}(j) - \bar{T}_{k,l})^2. \end{aligned}$$

TABLE III
EQUALITY OF MEANS TEST. VARIOUS STATISTICS

diff	ANOVA statistic	ANOVA p -values	Gamma GLR	$s \times \log(\text{AM}/\text{GM})$
Proposed \tilde{D}	06.30	9.35×10^{-19}	0.0533	0.0600
KL	06.68	2.88×10^{-20}	0.0561	0.0633
Chernoff	06.74	1.61×10^{-20}	0.0663	0.0756
L^1	24.00	3.42×10^{-87}	0.1652	0.2061

Note that these depend on the diff index under consideration. The oneway ANOVA test [25, p.533] is as follows: Reject H_0 (associated with the diff under consideration) if

$$\mathcal{T}(\text{diff}) := \frac{\sum_{(k,l)} n \left(\bar{T}_{k,l} - \bar{\bar{T}} \right)^2}{S_p^2} > (g-1) F_{g-1, g(n-1), \alpha},$$

where α is the desired significance level and $F_{g-1, g(n-1), \alpha}$ is the corresponding threshold.¹²

Look at the second and third columns of Table III. The first row contains the value of the ANOVA statistic with $\text{diff}(\mathbf{R}_k, \mathbf{R}_l) = \tilde{D}_{kl}$ and the corresponding p -value. The p -value is so small that we must summarily reject the null hypothesis H_0 associated with $\text{diff} = \tilde{D}$ (for, say, a typical significance level of 5%). The situation is the same for the other dissimilarity indices, as can be seen from the remaining rows of Table III. In each test, the null hypothesis is rejected for, say, the typical 5% significance level.

Observe that the values of $\mathcal{T}(\tilde{D})$, $\mathcal{T}(\text{KL})$, and $\mathcal{T}(\text{Chernoff})$ are close to each other while $\mathcal{T}(L^1)$ is significantly larger. This suggests that one could use $\mathcal{T}(\cdot)$ to rank the different dissimilarity measures in their ability to explain the observed data. The oneway ANOVA statistic suggests the ranking

$$\tilde{D} > \text{KL} > \text{Chernoff} > L^1. \quad (36)$$

We shall return to this observation after trying out two other refinements of the equality of means test.

2) *Equality of Means for Gamma Distributions*: We began with the oneway ANOVA statistic because it is known to be robust to the Gaussian assumption. We checked for Gaussianity anyway. Lilliefor's test for Gaussianity is a variation on the Kolmogorov-Smirnov test when the null hypothesis does not specify the parameters of the Gaussian distribution. None of the 24 groups of data passed the test of Gaussianity at the 5% significance level.

We next looked for features in the data that may suggest other distributions. First, the decision times are positive random variables. Next, Figure 6 shows the standard deviation versus the mean decision time for the 24 image pairs. Observe the linear relation, with $y = 0.61x$ being the best linear fit. A class of distributions on \mathbb{R}_+ whose standard deviation is a linear function of its mean is the family of Gamma distributions with a fixed shape parameter. The Gamma density

¹²The threshold at which the cdf of the F -distribution with $(g-1, g(n-1))$ degrees of freedom equals $1 - \alpha$.

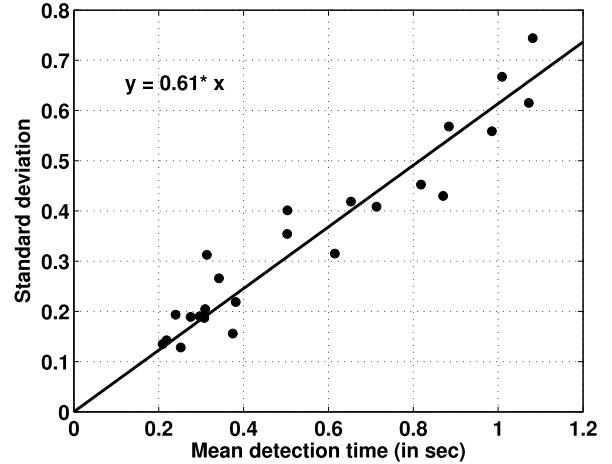


Fig. 6. Standard deviation of the decision times versus mean decision times, across image pairs.

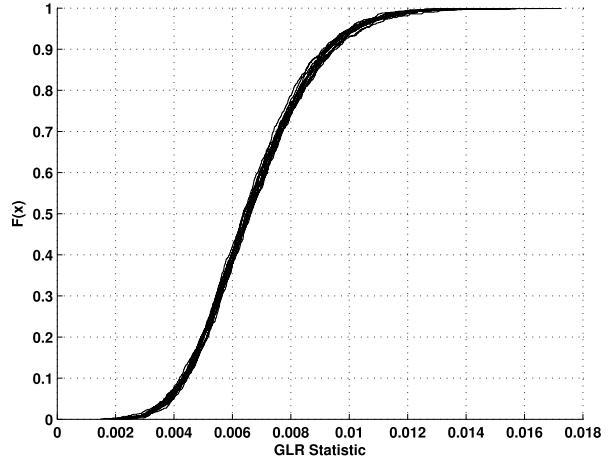


Fig. 7. Superposition of CDF of GLRT statistic under the Gamma assumption for the equality of means test. Each CDF consists of 1000 sample points. Each CDF corresponds to a random instance of mean and shape parameter. Mean was uniformly sampled from [0.2, 1.2]. Shape parameter was uniformly sampled from [2, 5]. Each sample point consisted of 24 groups, and 72 samples per group, same as that for the experimental data for decision times. The indicated intervals for mean and shape were based on the experimental data.

with shape parameter s and scale m is

$$(m\Gamma(s))^{-1} (x/m)^{s-1} e^{-x/m}.$$

The mean is ms and the standard deviation is $m\sqrt{s}$ so that standard deviation to mean ratio is $1/\sqrt{s}$. The slope of 0.61 in Figure 6 suggests a shape parameter of $1/(0.61)^2 = 2.7$. The Kolmogorov-Smirnov test on the data against Gamma distributions with shape parameter 2.7 and mean set to the sample mean accepts 18 of the 24 image pairs and rejects 6 out of 24 image pairs at 5% significance level. This suggests that Gamma is a better fit to the data than Gaussian.

We conducted a generalised likelihood ratio test (GLRT) for equality of means under the Gamma assumption, and under a constant shape parameter assumption. This corresponds to an “equality of scales” test; Shiue *et al.* [27] suggest a statistic analogous to the oneway ANOVA but for Gamma distributions. In Figure 7 we plot the CDF of the GLRT statistic under the null hypothesis and under equal mean and

equal shape parameter assumptions. Note that the CDF of this statistic is robust to the shape parameter.

In column 4 of Table III we provide the GLRT statistics for the decision time data. If we compare the GLRT statistics from Table III against GLRT CDF in Figure 7, we observe that the statistics is well beyond the 5% significance point. Again, we must summarily reject each of the equality of means hypotheses. Indeed, each GLRT statistic is off the chart in Figure 7. However, direct ordering of the statistics suggests the ranking (36), the same as that obtained with ANOVA.

3) *Equality of Means for Gamma Distributions With a Fixed and Known Shape Parameter:* Consider the setting where the shape parameter is known to be s across the groups. The GLRT for equality of means under this setting can be straightforwardly shown to be $s \log(\text{AM}/\text{GM})$ where AM and GM are the arithmetic and geometric means across groups defined as follows:

$$\text{AM} = \bar{\bar{T}} = \frac{1}{g} \sum_{(k,l)} \bar{T}_{k,l} \text{ and } \text{GM} = \left(\prod_{(k,l)} \bar{T}_{k,l} \right)^{1/g}.$$

The last column of Table III shows this statistic for the various dissimilarity measures. Yet again, the statistics are off the chart (CDF not plotted) with so small p -values that the null hypothesis of equality of means must be rejected. Once again, direct ordering of statistics suggests the ranking (36).

4) *A Lesser Goal - Ranking:* We saw that all three equality of means tests reject all four dissimilarity measures. In retrospect, this might have been anticipated. If the test associated with \tilde{D} had passed, that would have been a spectacular confirmation of our theory, which we really did not expect due to the crudeness of our modeling. Nevertheless, the equality of means test statistic provides a means to check which of the four dissimilarity measures best explains the data.

A little thought will inform us that all three equality of means tests check how clustered the sample means are, across various groups. This is clearest in the third equality of means test (for Gamma distributions with a known and common shape parameter across groups) where the statistic is a monotone function of the ratio AM/GM. The test passes if AM/GM is small, that is if the group means are close to each other, and fails if it is large.

The columns of Table III corresponding to each test statistic suggest that the data points are most clustered under \tilde{D} scaling and least clustered with L^1 scaling. The ranking is as in (36).

Let us note \tilde{D} , KL, and Chernoff are close to each other, and L^1 a distant fourth. Indeed, the vector $(\bar{T}_{(k,l)} / \bar{\bar{T}})_{(k,l)}$ associated with the L^1 dissimilarity measure *majorised* ([28, Defn.A.1]) the other three. There was no such ordering among \tilde{D} , KL, and Chernoff.

V. DISCUSSION

We modelled the visual search task of Sripati and Olson [3] as an active sequential hypothesis testing problem (ASHT). We extended the ASHT results of Chernoff [11] to the case with switching costs. We showed that adding switching costs does not affect the asymptotic growth rate of the total cost, provided ‘no switching’ incurs ‘zero cost’. This is an

interesting result in itself that may have wider applicability in sleep-wake cycling scenarios with switching costs where a change in the state of a node attracts a cost, but ‘no switching’ incurs ‘zero cost’. See [29] for a result along these lines.

The ASHT model suggests a dissimilarity index between pairs of images. The inverse of the asymptotic growth rate of the total cost in the ASHT model is proposed as a dissimilarity index between pairs of images. We derived expressions for computing the proposed dissimilarity index for the specific search task considered by Sripati and Olson [3]. The proposed dissimilarity index is a function of the neuronal firing rates elicited by the images in the infero temporal cortex of macaque monkeys.

Correlation study indicated that the proposed index is as good as L^1 and other dissimilarity measures such as the Chernoff entropy and the relative entropy (KL). Equality of means testing indicated that the equality of means hypothesis should be rejected, and this can be done with overwhelming confidence.

Equality of means testing procedure is perhaps a rather stringent test. What would be an appropriate test if, say, we can leave one group out? Does our proposed neuronal dissimilarity index pass such a less stringent test? Can we leave two groups out? Which two? We do not yet have a principled way to address these questions and instead decided to stick to the strictest test.

The statistics associated with the equality of means testing, however, suggested a ranking of the dissimilarity measures. We proposed three different statistics. Each measures the spread across groups of the group sample means. One of them is the familiar AM/GM ratio. The ranking was consistent across the three different statistics. Our proposed index was ranked first, relative entropy (KL divergence) and Chernoff entropy were a close second and third, and L^1 was a somewhat distant fourth.

The decision times were tested for the Gamma distribution and the test passed for two-third of the groups. The shape parameter for the distributions of delay, estimated via the method of moments, was close to 3.

In our work, we took only valid trials, i.e., those where the decisions made by the subjects were correct. We also assumed that the error probability tolerance were the same across subjects. It would be interesting to model speed-accuracy tradeoffs and see how they vary across individuals. It would also be interesting to explore how they vary for a single subject under different incentive settings.

Extension of ASHT to the case when no prior information is available about the images, where the subject has to actively learn \mathbf{R}_k and \mathbf{R}_l on-the-fly, is an interesting problem that we have addressed in a related paper [30].

APPENDIX A PROPERTIES OF LOG-LIKELIHOOD RATIO PROCESSES UNDER $\pi_{SA}(L, \eta)$

We will now show some desirable properties of the log-likelihood ratio processes under the policy $\pi_{SA}(L, \eta)$. These properties are analogous to those of classical sequential hypothesis testing, but their analyses are more involved

because actions introduce 1) dependency in the log-likelihood ratio increments, and 2) the increments are no longer identically distributed. The properties we will establish will be useful in forthcoming proofs.

Define $\Delta Z_{ji}(n) = Z_{ji}(n) - Z_{ji}(n-1)$. We then have $\Delta Z_{ji}(n) = -\Delta Z_{ij}(n)$. Here, $\Delta Z_{ji}(n)$ is the increment in the process associated with the log-likelihood ratio of H_i with respect to H_i at time n . We now show that under Assumptions (I) and (IIb), and under policy $\pi_{SA}(L, \eta)$, the log-likelihood ratio processes are well behaved in the following sense: the log-likelihood ratio of the true hypothesis H_i with respect to any other hypothesis H_j has a positive drift. This will be made precise in Proposition 13. Towards that, we first establish the following lemmas.

Lemma 10: Assume (I) and (IIb). Fix i, j such that $j \neq i$. Let $a \in \mathcal{A}_{ij}$. We then have, for all $0 < s < 1$,

$$\rho_{ij}^a(s) := E_i \left[e^{s\Delta Z_{ji}(n)} | A_n = a \right] < 1 \quad \forall n. \quad (37)$$

Proof: The following sequence of inequalities hold:

$$\begin{aligned} E_i \left[e^{s\Delta Z_{ji}(n)} | A_n = a \right] &= \int_{x \in \mathcal{X}} \left(\frac{q_j^a(x)}{q_i^a(x)} \right)^s q_i^a(x) dx \\ &= \int_{x \in \mathcal{X}} (q_j^a(x))^s (q_i^a(x))^{1-s} dx \\ &< \left(\int_{x \in \mathcal{X}} q_j^a(x) dx \right)^s \left(\int_{x \in \mathcal{X}} q_i^a(x) dx \right)^{1-s} \\ &= 1. \end{aligned} \quad (38)$$

The strict inequality in (38) follows from Hölder's inequality and the fact that $a \in \mathcal{A}_{ij}$ implies q_i^a and q_j^a are not linearly related. ■

The above result was obtained by conditioning on the action A_n to lie in the desirable set \mathcal{A}_{ij} . The result is independent of the underlying policy, because when conditioned on the current action A_n , the observation is independent of the policy.

Recall that $\tilde{\pi}_{SA}(\eta)$ is the non-stopping variant of $\pi_{SA}(L, \eta)$. Further, recall from Assumption (IIb) that we have $\beta = \min \left\{ \sum_{a \in \mathcal{A}_{ij}} \lambda_k(a) \mid 1 \leq i, j, k \leq M, i \neq j \right\} > 0$. Now we show that, under Assumption (IIb) and policy $\tilde{\pi}_{SA}(\eta)$, a similar result holds, but without conditioning on the action A_n . First, let us define

$$\rho_{ij}(s) := \eta\beta \left(\max_{a \in \mathcal{A}_{ij}} \rho_{ij}^a(s) \right) + (1 - \eta\beta). \quad (39)$$

The fact that $\rho_{ij}(s) < 1$ is evident from Lemma 10.

Lemma 11: Assume (I) and (IIb). Consider the policy $\tilde{\pi}_{SA}(\eta)$. Fix i . We then have, for all $0 < s < 1$,

$$E_i \left[e^{s\Delta Z_{ji}(n)} | X^{n-1}, A^{n-1} \right] \leq \rho_{ij}(s) < 1 \quad \forall n, \forall j \neq i.$$

Proof: The following sequence of inequalities hold as described after the last inequality.

$$\begin{aligned} E_i \left[e^{s\Delta Z_{ji}(n)} | X^{n-1}, A^{n-1} \right] &= E_i \left[E_i \left[e^{s\Delta Z_{ji}(n)} | X^{n-1}, A^{n-1}, A_n \right] | X^{n-1}, A^{n-1} \right] \\ &= \sum_{a \in \mathcal{A}} P_i(A_n = a | X^{n-1}, A^{n-1}) E_i \left[e^{s\Delta Z_{ji}(n)} | A_n = a \right] \end{aligned} \quad (40)$$

$$\begin{aligned} &\leq P_i(A_n \in \mathcal{A}_{ij} | X^{n-1}, A^{n-1}) \max_{a \in \mathcal{A}_{ij}} E_i \left[e^{s\Delta Z_{ji}(n)} | A_n = a \right] \\ &\quad + (1 - P_i(A_n \in \mathcal{A}_{ij} | X^{n-1}, A^{n-1})) \end{aligned} \quad (41)$$

$$\begin{aligned} &\leq \eta\beta \left(\max_{a \in \mathcal{A}_{ij}} \rho_{ij}^a(s) \right) + (1 - \eta\beta) \\ &< 1. \end{aligned} \quad (42)$$

Equality (40) holds because conditioned on $A_n = a$, $\Delta Z_{ij}(n)$ is independent of the remaining history. Inequality (41) holds because, when $a \notin \mathcal{A}_{ij}$, we have $\Delta Z_{ij}(n) \equiv 0$. The penultimate inequality is a consequence of the fact that, under $\pi_{SA}(L, \eta)$, one will choose an action $a \in \mathcal{A}_{ij}$ with probability at least $\eta\beta$. ■

We now proceed to show an inequality analogous to the Chernoff bound for the log-likelihood ratio. In classical sequential hypothesis testing, due to independence of samples across time, the expectation of the likelihood ratio can be split as the product of the expectation of the likelihood ratio increments, as follows:

$$E_i \left[e^{sZ_{ji}(n)} \right] = \prod_{k=1}^n E_i \left[e^{s\Delta Z_{ji}(n)} \right].$$

The same decomposition is not valid in ASHT because actions introduce dependency in the likelihood ratio increments across time. However, we can obtain an upper bound of the product form.

Lemma 12: Assume (I) and (IIb). Consider policy $\tilde{\pi}_{SA}(\eta)$. Fix i . We then have, for all $0 < s < 1$,

$$E_i \left[e^{sZ_{ji}(n)} \right] \leq (\rho_{ij}(s))^n \quad \forall n, \forall j \neq i.$$

Proof: Once again, we proceed through the chain of inequalities all of which are now self-evident:

$$\begin{aligned} E_i \left[e^{sZ_{ji}(n)} \right] &= E_i \left[E_i \left[e^{sZ_{ji}(n-1)} e^{s\Delta Z_{ji}(n)} | X^{n-1}, A^{n-1} \right] \right] \\ &= E_i \left[e^{sZ_{ji}(n-1)} E_i \left[e^{s\Delta Z_{ji}(n)} | X^{n-1}, A^{n-1} \right] \right] \\ &= \rho_{ij}(s) E_i \left[e^{sZ_{ji}(n-1)} \right] \text{ (from Lemma 11)} \\ &\leq (\rho_{ij}(s))^n, \end{aligned}$$

where the last inequality follows by induction. ■

We now show an exponential decay property of the log-likelihood process which primarily stems from the anticipated negative drift in $Z_{ji}(n)$ for $j \neq i$. Let us alert the reader that in the following Proposition we deal with $Z_{ij}(n) = -Z_{ji}(n)$.

Proposition 13: Assume (I) and (IIb). Consider policy $\tilde{\pi}_{SA}(\eta)$. Fix i . There exist constants $C_K > 0$ and $\gamma > 0$ such that

$$P_i \left(\min_{j \neq i} Z_{ij}(n) \leq K \right) < C_K e^{-\gamma n}. \quad (43)$$

C_K is independent of i , but γ may depend on i .

Proof: This follows from the previous lemmas via the following :

$$P_i \left(\min_{j \neq i} Z_{ij}(n) \leq K \right) = P_i \left(\max_{j \neq i} Z_{ji}(n) \geq -K \right)$$

$$\leq \sum_{j \neq i} P_i (Z_{ji}(n) \geq -K) \quad (44)$$

$$\leq \sum_{j \neq i} e^{sK} E_i \left[e^{sZ_{ji}(n)} \right] \quad (45)$$

$$\leq e^{sK} \sum_{j \neq i} (\rho_{ij}(s))^n \quad (46)$$

$$\leq e^{sK} \cdot (M-1) \cdot \max_{j \neq i} (\rho_{ij}(s))^n \\ = C_K e^{-\gamma n},$$

where $\max_{j \neq i} \rho_{ij}(s) = e^{-\gamma}$, and $C_K = M e^{sK}$. The inequality in (44) is due to the union bound, the inequality in (45) is due to Chernoff's bound with $0 < s < 1$, and the inequality in (46) is due to Lemma 12. \blacksquare

We now show that under the hypothesis $H = H_i$, the $\theta(n)$ process eventually settles at i . Indeed we show something stronger. Let us define

$$T_i := \inf\{n : \theta(n') = i, \forall n' \geq n\}, \quad (47)$$

the time at which $\theta(n)$ meets its eventuality of settlement at i . This random variable has a tail that decays exponentially fast, as shown next.

Lemma 14: Assume (I) and (IIb). Consider policy $\tilde{\pi}_{SA}(\eta)$. Fix i . Then there exist $C > 0$ and $b > 0$, both finite and possibly dependent on i , such that

$$P_i(T_i > n) < C e^{-bn}. \quad (48)$$

Proof: By the union bound

$$\begin{aligned} P_i(T_i > n) &= P_i(\theta(n') \neq i \text{ for some } n' \geq n) \\ &\leq \sum_{n' \geq n} P_i(\theta(n') \neq i) \\ &\leq \sum_{n' \geq n} P_i \left(\min_{j \neq i} Z_{ij}(n') \leq 0 \right). \end{aligned}$$

The assertion now follows from Proposition 13. \blacksquare

Thus far we have considered the policy $\tilde{\pi}_{SA}(\eta)$ which never stops. We now show that the policy $\pi_{SA}(L, \eta)$ stops in finite time.

Proposition 15: Assume (I) and (IIb). Consider the policy $\pi_{SA}(L, \eta)$. Fix i . We then have

$$P_i(\tau(\pi_{SA}(L, \eta)) < \infty) = 1.$$

Proof: We consider $\pi_{SA}^i(L, \eta)$ for analysis. Recall that $\tau(\pi_{SA}(L, \eta)) \leq \tau(\pi_{SA}^i(L, \eta))$, and hence it is sufficient to show that

$$P_i(\tau(\pi_{SA}^i(L, \eta)) < \infty) = 1. \quad (49)$$

From Proposition 13, we know that, for a suitable constant \tilde{C} ,

$$P_i \left(\min_{j \neq i} Z_{ij}(n) < \log(L(M-1)) \right) < \tilde{C} e^{-\gamma n}.$$

Since this bound is summable, by the Borel-Cantelli lemma,

$$P_i \left(\min_{j \neq i} Z_{ij}(n) < \log(L(M-1)) \text{ infinitely often} \right) = 0,$$

which is stronger than the assertion (49). \blacksquare

Propositions 13 and 15 are the ones that will be used in the sequel.

A. Proof of Proposition 3

The proof relies on a standard change of measure argument. Let Δ_j denote the event that the policy $\pi_{SA}(L, \eta)$ declares H_j as the true hypothesis.

$$\begin{aligned} P_i(\delta \neq i) &= \sum_{j \neq i} P_i(\delta = j) + P_i(\tau(\pi_{SA}(L, \eta)) = \infty) \\ &= \sum_{j \neq i} \sum_{n > 0} \int_{\omega^n \in \Delta_j} dP_i(\omega^n) + 0 \\ &= \sum_{j \neq i} \sum_{n > 0} \int_{\omega^n \in \Delta_j} \frac{dP_i}{dP_j}(\omega^n) dP_j(\omega^n) \\ &\leq \sum_{j \neq i} \sum_{n > 0} \int_{\omega^n \in \Delta_j} \frac{1}{(M-1)L} dP_j(\omega^n) \\ &\leq \frac{1}{(M-1)L} \sum_{j \neq i} P_j(\Delta_j) \\ &\leq \frac{1}{L}. \end{aligned} \quad (50)$$

The second equality holds because we have shown in Proposition 15 that the stopping time is finite with probability 1. The inequality (50) follows because $\omega^n \in \Delta_j$ implies $Z_{ji}(n)(\omega^n) \geq \log((M-1)L)$, that is, $\frac{dP_i}{dP_j}(\omega^n) \leq \frac{1}{(M-1)L}$. \blacksquare

B. Proof of Theorem 4: Achievability

We assume (I) and (IIb). All statements in this proof are under $H = H_i$ and under *Sluggish Procedure A*. We follow the proof technique of Chernoff [11, Lem. 2]. Chernoff's proof technique does not go through completely because unlike in *Procedure A*, the next action in *Sluggish Procedure A* is not conditionally independent of the previous action, given the current likelihood values. A similar issue was addressed by Nitinawarat and Veeravalli in [16], in the context of Markovian observation model, and we will adapt their proof technique to our setting.

Let us first setup some notation. Fix $\epsilon > 0$. Define

$$D_{ij} := \sum_{a \in \mathcal{A}} \lambda_i(a) D(q_i^a \| q_j^a),$$

where λ_i is as defined in (5). Let D_i be as defined by (6), i.e., $D_i = \min_{j \neq i} D_{ij}$. Under the *Sluggish Procedure A*, the transition probability matrix $TP(\theta(n))$ of the action process A_n at time n is given by

$$TP(\theta(n)) = (1 - \eta) \mathbf{I} + \eta \left(\mathbf{1} \lambda_{\theta(n)}^T \right). \quad (51)$$

It is easy to verify that the stationary distribution associated with $TP(\theta(n))$ is $\lambda_{\theta(n)}$. Define $\mathcal{F}_{k-1} := \sigma(X^{k-1}, A^{k-1})$, the σ -field generated by the random variables (X^{k-1}, A^{k-1}) .

We now upper bound the expected time to make a decision under *Sluggish Procedure A* as follows:

$$\begin{aligned} E_i[\tau(\pi_{SA}(L, \eta))] &\leq E_i[\tau(\pi_{SA}^i(L, \eta))] \\ &= \sum_{n \geq 0} P_i(\tau(\pi_{SA}^i(L, \eta)) > n) \\ &\leq \frac{(1 + \epsilon) \log(L(M-1))}{D_i} \end{aligned}$$

$$+ \sum_{n \geq \tilde{n}} P_i \left(\tau(\pi_{SA}^i(L, \eta)) > n \right), \quad (52)$$

where

$$\tilde{n} = \frac{(1 + \epsilon) \log(L(M - 1))}{D_i}.$$

To complete the proof, we will now show that for any $\epsilon > 0$, the second term on the right-hand side of (52) goes to zero as $L \rightarrow \infty$. Indeed, we claim that each term in the summation decays exponentially with n with an exponent that does not depend on L . Assuming the claim, the tail sum vanishes as $L \rightarrow \infty$, because $\tilde{n} \rightarrow \infty$. This suffices to complete the proof of Theorem 4.

We now proceed to prove the claim. Observe that

$$\begin{aligned} & P_i \left(\tau(\pi_{SA}^i(L, \eta)) > n \right) \\ & \leq P_i \left(\min_{j \neq i} Z_{ij}(n) \leq \log(L(M - 1)) \right) \\ & \leq \sum_{j \neq i} P_i \left(Z_{ij}(n) \leq \log(L(M - 1)) \right). \end{aligned}$$

Fix one $j \neq i$. (The same analysis holds for other j .) Then

$$\begin{aligned} & P_i \left(Z_{ij}(n) \leq \log(L(M - 1)) \right) \\ & = P_i \left(\sum_{k=1}^n \Delta Z_{ij}(k) \leq \log(L(M - 1)) \right) \\ & = P_i \left(\sum_{k=1}^n (\Delta Z_{ij}(k) - E_i [\Delta Z_{ij}(k) | \mathcal{F}_{k-1}]) + \epsilon' \right. \\ & \quad \left. + \sum_{k=1}^n (E_i [\Delta Z_{ij}(k) | \mathcal{F}_{k-1}] - D_{ij} + \epsilon') \right. \\ & \quad \left. + n(D_{ij} - 2\epsilon') \leq \log(M - 1)L \right) \\ & \leq P_i \left(\sum_{k=1}^n (\Delta Z_{ij}(k) - E_i [\Delta Z_{ij}(k) | \mathcal{F}_{k-1}]) + \epsilon' < 0 \right) \\ & \quad + P_i \left(\sum_{k=1}^n (E_i [\Delta Z_{ij}(k) | \mathcal{F}_{k-1}] - D_{ij} + \epsilon') < 0 \right) \\ & \quad + P_i \left(n(D_{ij} - 2\epsilon') \leq \log(L(M - 1)) \right). \end{aligned} \quad (53)$$

Look at the first probability term in (53). Each entry within the summation has a positive mean and, from Chernoff's bounding technique in [11, Lem. 2], there exists a $b(\epsilon') > 0$ such that

$$P_i \left(\sum_{k=1}^n (\Delta Z_{ij}(k) - E_i [\Delta Z_{ij}(k) | \mathcal{F}_{k-1}]) + \epsilon' < 0 \right) \leq e^{-nb(\epsilon')}.$$

The third probability term is 0 if we choose an ϵ' small enough such that $n(D_{ij} - 2\epsilon') > \log(L(M - 1))$, for all $n > \tilde{n}$. Indeed, any ϵ' satisfying $0 < \epsilon' < \frac{\epsilon}{1+\epsilon} \frac{D_i}{2}$ suffices. So set $\epsilon' = \frac{\epsilon}{1+\epsilon} \frac{D_i}{4}$.

We now proceed to show that the second term also decays exponentially to zero. Let T_i be as defined in (47). For a suitably chosen ϵ'' , and we will soon indicate how to choose it, we have

$$P_i \left(\sum_{k=1}^n (E_i [\Delta Z_{ij}(k) | \mathcal{F}_{k-1}] - D_{ij} + \epsilon') < 0 \right)$$

$$\begin{aligned} & \leq P_i \left(\sum_{k=1}^n (E_i [\Delta Z_{ij}(k) | \mathcal{F}_{k-1}] - D_{ij} + \epsilon') < 0, T_i \leq n\epsilon'' \right) \\ & \quad + P_i(T_i > n\epsilon''). \end{aligned}$$

From Lemma 14, the second probability term on the right-hand side decays exponentially with n . To show that the first probability term on the right-hand side decays exponentially with n , we use a technique of Nitinawarat and Veeravalli [16, (6.23)].

First, we indicate how to choose ϵ'' . Define

$$\begin{aligned} \tilde{C} & = \min_{a \in \mathcal{A}} E_i [\Delta Z_{ij}(k) | A_k = a] - D_{ij} \\ & = \min_{a \in \mathcal{A}} D(q_i^a \| q_j^a) - D_{ij}. \end{aligned}$$

Since D_{ij} is the λ_i -weighted average of $D(q_i^a \| q_j^a)$, we have $\tilde{C} \leq 0$. Choose ϵ'' small enough so that $\tilde{\epsilon} := \epsilon' + \epsilon'' \tilde{C} > 0$. We then have

$$\begin{aligned} & P_i \left(\sum_{k=1}^n (E_i [\Delta Z_{ij}(k) | \mathcal{F}_{k-1}] - D_{ij} + \epsilon') < 0, T_i \leq n\epsilon'' \right) \\ & = P_i \left(\sum_{k=1}^{\lfloor n\epsilon'' \rfloor} (E_i [\Delta Z_{ij}(k) | \mathcal{F}_{k-1}] - D_{ij} + \epsilon') \right. \\ & \quad \left. + \sum_{k=\lfloor n\epsilon'' \rfloor + 1}^n (E_i [\Delta Z_{ij}(k) | \mathcal{F}_{k-1}] - D_{ij} + \epsilon') < 0, \right. \\ & \quad \left. T_i \leq n\epsilon'' \right) \\ & \leq P_i \left(\lfloor n\epsilon'' \rfloor (\tilde{C} + \epsilon') \right. \\ & \quad \left. + \sum_{k=\lfloor n\epsilon'' \rfloor + 1}^n (E_i [\Delta Z_{ij}(k) | \mathcal{F}_{k-1}] - D_{ij} + \epsilon') < 0, \right. \\ & \quad \left. T_i \leq n\epsilon'' \right) \\ & \leq P_i \left(\sum_{k=\lfloor n\epsilon'' \rfloor + 1}^n (E_i [\Delta Z_{ij}(k) | \mathcal{F}_{k-1}] - D_{ij} + \tilde{\epsilon}) < 0, \right. \\ & \quad \left. T_i \leq n\epsilon'' \right) \\ & \leq \tilde{P}_i \left(\sum_{k=\lfloor n\epsilon'' \rfloor + 1}^n (E_i [\Delta Z_{ij}(k) | \mathcal{F}_{k-1}] - D_{ij} + \tilde{\epsilon}) < 0 \right) \\ & \leq Ce^{-n\tilde{b}(\tilde{\epsilon})}, \end{aligned} \quad (54)$$

for some $C > 0$ and some $\tilde{b}(\tilde{\epsilon}) > 0$. The second inequality follows from the fact that $\tilde{C} \leq E_i [\Delta Z_{ij}(k) | \mathcal{F}_{k-1}] - D_{ij}$, for all k . The third inequality follows from the choice of $\tilde{\epsilon}$ and the fact that

$$\lfloor n\epsilon'' \rfloor (\tilde{C} + \epsilon') + (n - \lfloor n\epsilon'' \rfloor) \epsilon' \geq (n - \lfloor n\epsilon'' \rfloor) \tilde{\epsilon}.$$

\tilde{P}_i is a new measure under which actions are taken according to *Sluggish Procedure A* but assuming $\theta(n) = i \ \forall n$, and

the observations are conditionally independent of past observations and actions, given the current action. Consequently, under \tilde{P}_i , the action process A_n is a stationary Markov Chain with transition probability matrix $TP(i)$. By the ergodic theorem and concentration inequalities for Markov Chains [31], this term also decays exponentially with n , which is (54). ■

APPENDIX B PROOF OF PROPOSITION 9

We will focus only on $i \leq W$ and will determine

$$D_i = \max_{\lambda \in \mathcal{P}(\mathcal{A})} \min_{j \neq i} \sum_{a \in \mathcal{A}} \lambda(a) D(q_i^a \| q_j^a). \quad (55)$$

The case $i > W$ can be handled similarly and is omitted. Using (23) - (27), we can simplify the minimisation in (55) by considering three regions for j as follows:

$$\begin{aligned} D_i &= \max_{\lambda \in \mathcal{P}(\mathcal{A})} \min \left\{ \min_{j \leq W, j \neq i} (\lambda(i) D(f_k \| f_i) + \lambda(j) D(f_i \| f_k)), \right. \\ &\quad \lambda(i) D(f_k \| f_i) + (1 - \lambda(i)) D(f_i \| f_k), \\ &\quad \left. \min_{j > W, j \neq i + W} (1 - \lambda(i) - \lambda(j - W)) D(f_i \| f_k) \right\}, \quad (56) \\ &= \max_{\lambda \in \mathcal{P}(\mathcal{A})} \min \left\{ \lambda(i) D(f_k \| f_i) + \min_{j \neq i} \lambda(j) D(f_i \| f_k), \right. \\ &\quad \lambda(i) D(f_k \| f_i) + (1 - \lambda(i)) D(f_i \| f_k), \\ &\quad \left. (1 - \lambda(i) - \max_{j \neq i} \lambda(j)) D(f_i \| f_k) \right\}. \quad (57) \end{aligned}$$

Observe that the second term is always greater than or equal to the other two terms, and hence can be removed from the minimisation. Thus,

$$\begin{aligned} D_i &= \max_{\lambda \in \mathcal{P}(\mathcal{A})} \min \left\{ \lambda(i) D(f_k \| f_i) + \min_{j \neq i} \lambda(j) D(f_i \| f_k), \right. \\ &\quad \left. (1 - \lambda(i) - \max_{j \neq i} \lambda(j)) D(f_i \| f_k) \right\}. \quad (58) \end{aligned}$$

We now perform the maximisation over λ in two steps. First, let us fix $\lambda(i)$ and optimise over the distribution of $1 - \lambda(i)$ among the other actions. Since

$$\min_{j \neq i} \lambda(j) \leq \frac{1 - \lambda(i)}{W - 1} \leq \max_{j \neq i} \lambda(j),$$

we have

$$\left(\min_{j \neq i} \lambda(j) \right) D(f_i \| f_k) \leq \left(\frac{1 - \lambda(i)}{W - 1} \right) D(f_i \| f_k)$$

and

$$-\max_{j \neq i} \lambda(j) D(f_i \| f_k) \leq -\frac{(1 - \lambda(i))}{W - 1} D(f_i \| f_k).$$

Thus both the terms within braces in (58) are lesser than or equal to the corresponding terms for equal distribution of $1 - \lambda(i)$ among the other actions. The optimisation problem is now reduced to a single variable optimisation of the form

$$\begin{aligned} D_i &= \max_{0 \leq \lambda(i) \leq 1} \min \left\{ \lambda(i) D(f_k \| f_i) + \frac{(1 - \lambda(i))}{W - 1} D(f_i \| f_k), \right. \\ &\quad \left. (1 - \lambda(i)) \frac{W - 2}{W - 1} D(f_i \| f_k) \right\}. \quad (59) \end{aligned}$$

Second, we now perform the optimisation in (59) over $\lambda(i)$. The first term in the minimisation is increasing or non-increasing in $\lambda(i)$ depending on $D(f_k \| f_i) > D(f_i \| f_k)/(W - 1)$ or $D(f_k \| f_i) \leq D(f_i \| f_k)/(W - 1)$, respectively. The second term is decreasing in $\lambda(i)$.

1) Suppose $D(f_k \| f_i) > D(f_i \| f_k)/(W - 1)$, then the two terms viewed as linear functions over $\lambda(i)$ cross each other, and so the maximum will be achieved at the point of equality, i.e.,

$$\begin{aligned} &\lambda(i) D(f_k \| f_i) + \frac{(1 - \lambda(i))}{W - 1} D(f_i \| f_k) \\ &= (1 - \lambda(i)) \frac{W - 2}{W - 1} D(f_i \| f_k). \end{aligned}$$

Solving for $\lambda(i)$ yields

$$\begin{aligned} \lambda(i) &= \frac{(W - 3) D(f_i \| f_k)}{(W - 1) D(f_k \| f_i) + (W - 3) D(f_i \| f_k)}, \\ \lambda(j) &= \frac{D(f_k \| f_i)}{(W - 1) D(f_k \| f_i) + (W - 3) D(f_i \| f_k)}, \quad \forall j \neq i, \\ D_i &= \frac{(W - 2) D(f_k \| f_i) D(f_i \| f_k)}{(W - 1) D(f_k \| f_i) + (W - 3) D(f_i \| f_k)}. \end{aligned}$$

2) Suppose $D(f_k \| f_i) < D(f_i \| f_k)/(W - 1)$, then the maximum is achieved at $\lambda(i) = 0$. Then $\lambda(j) = 1/(W - 1)$, $\forall j \neq i$, and

$$\begin{aligned} D_i &= \min \left\{ \frac{D(f_i \| f_k)}{W - 1}, \frac{(W - 2) D(f_i \| f_k)}{W - 1} \right\} \\ &= \frac{D(f_i \| f_k)}{W - 1}, \end{aligned}$$

since $W > 3$.

3) Suppose $D(f_k \| f_i) = D(f_i \| f_k)/(W - 1)$, then the maximum is achieved at any of the following λ : $\lambda(i) = c$, for any $0 \leq c \leq \frac{W-3}{W-2}$. Then $\lambda(j) = 1 - c/(W - 1)$, $\forall j \neq i$, and

$$\begin{aligned} D_i &= \min \left\{ \frac{D(f_i \| f_k)}{W - 1}, (1 - c) \frac{(W - 2) D(f_i \| f_k)}{W - 1} \right\} \\ &= \frac{D(f_i \| f_k)}{W - 1}. \end{aligned}$$

This completes the proof. ■

APPENDIX C ESTIMATION OF RELATIVE ENTROPY RATE

The computation of our proposed neuronal index requires a computation of the relative entropy rate between two Poisson point processes *from estimates of their rates*. Wang et al. [32] considered the problem of estimating the relative entropy between two continuous distributions based on i.i.d. samples from each distribution. The authors proposed a k -nearest neighbour based estimate for the relative entropy and showed that their estimate is asymptotically unbiased and mean-square consistent. However, the approach in [32] is not directly applicable to our problem because we do not have i.i.d. samples from the distributions. Instead we only have empirical firing rates of the two Poisson point processes.

The relative entropy between two Poisson point processes with rates R_1 and R_2 is

$$\begin{aligned} \frac{1}{T} D(\mu_{R_1, T} \| \mu_{R_2, T}) &= R_1 \log \left(\frac{R_1}{R_2} \right) + R_2 - R_1 \\ &= R_1 \log R_1 - R_1 \log R_2 + R_2 - R_1. \end{aligned} \quad (60)$$

Let $N_i(k, T)$ be the number of spikes observed in time slot k , $1 \leq k \leq n$, of duration T on the i^{th} process, $i = 1, 2$. The empirical firing rate is then $\hat{R}_i = \frac{1}{nT} \sum_{k=1}^n N_i(k, T)$. A natural estimate for (60), based on the observations, would be to substitute R_i , $i = 1, 2$ by their respective empirical estimates \hat{R}_i , $i = 1, 2$, to get

$$\begin{aligned} \hat{D} &= \hat{R}_1 \log \left(\frac{\hat{R}_1}{\hat{R}_2} \right) + \hat{R}_2 - \hat{R}_1 \\ &= \hat{R}_1 \log \hat{R}_1 - \hat{R}_1 \log \hat{R}_2 + \hat{R}_2 - \hat{R}_1. \end{aligned} \quad (61)$$

A little reflection suggests that this is a bad estimate, for there is a positive probability that $\hat{R}_1 > \hat{R}_2 = 0$, yielding $E_{R_1, R_2} [\hat{D}] = \infty$. Estimate (61) is thus biased (though consistent). Our approach is to obtain estimates for each of the terms in (60) with minimal bias.

Unbiased and maximum likelihood estimates for the third and fourth terms on the right hand side of (60) are the respective empirical firing rates themselves. Let us therefore now study the second term. We may assume that the firings are independent, given R_1 and R_2 . Thus we may look for an estimator of the form $-\hat{R}_1 f(\hat{R}_2)$ which has expectation $-R_1 E_{R_2} [f(\hat{R}_2)]$. For this to be close to the desired $-R_1 \log R_2$, we look for a function $f(\hat{R}_2)$ such that $E_{R_2} [f(\hat{R}_2)] \approx \log R_2$. The difficulty is due to the $\log(0) = -\infty$ artifact. We consider a simple fix of adding a nonzero offset to the empirical estimate, i.e., we consider estimates of the form $\log(\hat{R} + \theta)$. Figure 8 shows the optimum offset $\theta^*(R)$ for different firing rates R when $n = T = 1$. The optimum offset $\theta^*(R)$ can be seen to converge to 0.5 for large R . Further, the convergence is quite fast, $\theta^*(R)$ is close to 0.5 for all R greater than 3. Hence in this work we use $\theta = 0.5$ as the offset, thus resulting in an estimator for $\log(R)$ of the form $\log(\hat{R} + 1/2)$. For a general n and T we then have $E [\log(nT\hat{R} + 1/2)] \approx \log(nTR)$, which in turn implies $E [\log(\hat{R} + 1/2nT)] \approx \log(R)$. Thus an estimator for a general n and T would be $\log(\hat{R} + 1/2nT)$. The estimator for the second term in (60) is then $-\hat{R}_1 \log(\hat{R}_2 + 1/2nT)$. One could look for better estimators with the offset being a function of the observed empirical means. In this work we stick to the constant offset estimator, with the constant offset being $\theta = 0.5$, as it is reasonable to assume that the neurons have a firing rate greater than $3/nT = 3/(24 * 0.25) = 0.5$ spikes/second ($n = 24$, $T = 250$ ms), thus putting them in the firing rate regime where $\theta = 0.5$ is a good offset for near unbiasedness. The values $T = 250$ ms and $n = 24$ correspond to the neuronal recording time and the number of repetitions in the neuronal recording experiment of Sripathi and Olson [3].

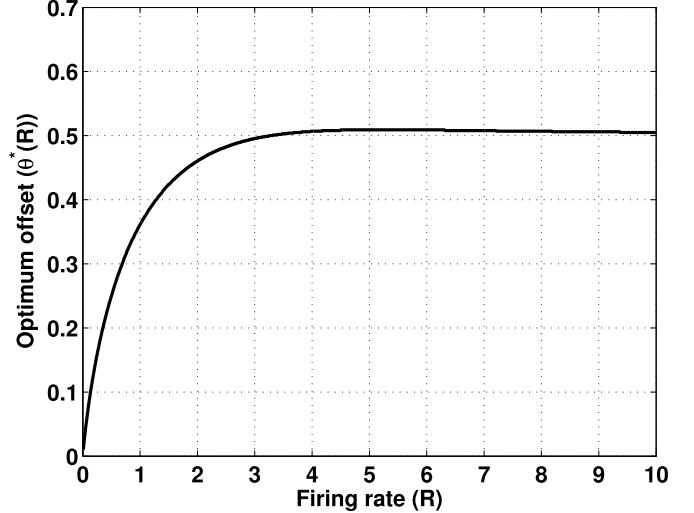


Fig. 8. Optimum offset ($\theta^*(R)$) to minimise bias for different firing rates (R).

To address the first term of (60) we consider estimates of the form $\hat{R}_1 g(\hat{R}_1)$ such that

$$E_{R_1} [\hat{R}_1 g(\hat{R}_1)] \cong R_1 \log R_1.$$

Expanding the expectation above for $n = T = 1$, we obtain,

$$\begin{aligned} E_{R_1} [\hat{R}_1 g(\hat{R}_1)] &= \sum_{k=0}^{\infty} k g(k) \frac{R_1^k e^{-R_1}}{k!} \\ &= R_1 \sum_{k=1}^{\infty} g(k) \frac{R_1^{k-1} e^{-R_1}}{(k-1)!} \\ &= R_1 E_{R_1} [g(\hat{R}_1 + 1)]. \end{aligned}$$

Thus we want a g such that $E_{R_1} [g(\hat{R}_1 + 1)] \cong \log R_1$. From the discussion on the second term, we know that $E_{R_1} [\log(\hat{R}_1 + 1/2)] \sim \log R_1$, and hence a good choice for g would be $g(\hat{R}_1) = \log(\hat{R}_1 + 1/2)$. Thus our estimate for the first term for a general n and T is

$$\begin{cases} \hat{R}_1 \log \left(\hat{R}_1 + \frac{1}{2nT} \right) & \text{if } \hat{R}_1 \geq \frac{1}{2nT}, \\ 0 & \text{otherwise.} \end{cases}$$

i.e., there is atleast one point,

Therefore our combined estimate for the relative entropy rate in (60), based on the average firing rate estimates \hat{R}_1 and \hat{R}_2 and obtained over a time of duration nT is

$$\hat{D}(\hat{R}_1 \| \hat{R}_2) = \begin{cases} \left[\hat{R}_1 \log \left(\frac{\hat{R}_1 - \frac{1}{2nT}}{\hat{R}_2 + \frac{1}{2nT}} \right) + \hat{R}_2 - \hat{R}_1 \right]^+ & \text{if } \hat{R}_1 \geq \frac{1}{2nT} \\ \hat{R}_2 & \text{otherwise.} \end{cases}$$

Relative entropy being a convex function of its arguments, the plug-in estimator of (61) would always have a positive bias. Naturally, an unbiased estimator would have a smaller value than (61), and our proposed estimator does satisfy this requirement.

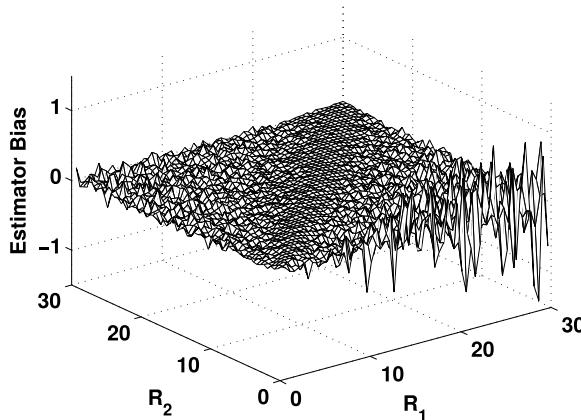


Fig. 9. Estimator bias for the proposed relative entropy rate estimator, $n = 24$, $T = 250$ ms.

In Figure 9 we plot the estimator bias for different (R_1, R_2) pairs for $n = 24$ and $T = 250$ ms, motivated by the specific neuronal experimental data of Sripathi and Olson [3]. From Figure 9 we can see that our proposed estimator has low estimation error for most (R_1, R_2) . Estimation error is relatively large only when R_1 is large and R_2 is close to zero.

REFERENCES

- [1] N. K. Vaidhiyan, S. P. Arun, and R. Sundaresan, "Active sequential hypothesis testing with application to a visual search problem," in *Proc. ISIT*, Jul. 2012, pp. 2201–2205.
- [2] N. K. Vaidhiyan and R. Sundaresan, "Active search with a cost for switching actions," in *Proc. Inf. Theory Appl. Workshop (ITA)*, Feb. 2015, pp. 17–24.
- [3] A. P. Sripathi and C. R. Olson, "Global image dissimilarity in macaque inferotemporal cortex predicts human visual search efficiency," *J. Neurosci.*, vol. 30, no. 4, pp. 1258–1269, Jan. 2010.
- [4] K. Nakayama and P. Martini, "Situating visual search," *Vis. Res.*, vol. 51, no. 13, pp. 1526–1537, 2011.
- [5] J. M. Wolfe and T. S. Horowitz, "What attributes guide the deployment of visual attention and how do they do it?" *Nature Rev., Neurosci.*, vol. 5, no. 6, pp. 495–501, 2004.
- [6] J. Palmer, P. Verghese, and M. Pavel, "The psychophysics of visual search," *Vis. Res.*, vol. 40, no. 10, pp. 1227–1268, 2000.
- [7] A. Wald and J. Wolfowitz, "Optimum character of the sequential probability ratio test," *Ann. Math. Statist.*, vol. 19, no. 3, pp. 326–339, 1948.
- [8] M. Naghshvar and T. Javidi, "Active M -ary sequential hypothesis testing," in *Proc. IEEE Int. Symp. Info. Theory*, Jun. 2010, pp. 1623–1627.
- [9] M. Naghshvar and T. Javidi, "Active sequential hypothesis testing," *Ann. Statist.*, vol. 41, no. 6, pp. 2703–2738, Dec. 2013.
- [10] S. Gezeck, B. Fischer, and J. Timmer, "Saccadic reaction times: A statistical analysis of multimodal distributions," *Vis. Res.*, vol. 37, no. 15, pp. 2119–2131, 1997.
- [11] H. Chernoff, "Sequential design of experiments," *Ann. Math. Statist.*, vol. 30, no. 3, pp. 755–770, 1959.
- [12] M. Naghshvar and T. Javidi, "Information utility in active sequential hypothesis testing," in *Proc. 48th Annu. Allerton Conf.*, Oct. 2010, pp. 123–129.
- [13] M. Naghshvar and T. Javidi, "Performance bounds for active sequential hypothesis testing," in *Proc. IEEE Int. Symp. Inf. Theory*, Aug. 2011, pp. 2666–2670.
- [14] M. Naghshvar and T. Javidi, "Sequentiality and adaptivity gains in active hypothesis testing," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 5, pp. 768–782, Oct. 2013.
- [15] S. Nitinawarat, G. K. Atia, and V. V. Veeravalli, "Controlled sensing for multihypothesis testing," *IEEE Trans. Autom. Control*, vol. 58, no. 10, pp. 2451–2464, Oct. 2013.
- [16] S. Nitinawarat and V. V. Veeravalli, "Controlled sensing for sequential multihypothesis testing with controlled Markovian observations and non-uniform control cost," *Design Methods Appl.*, vol. 34, no. 1, pp. 1–24, Oct. 2013.
- [17] K. Cohen and Q. Zhao, "Active hypothesis testing for anomaly detection," *IEEE Trans. Inf. Theory*, vol. 61, no. 3, pp. 1432–1450, Mar. 2015.
- [18] V. Srivastava, K. Pilar, and F. Bullo, "Adaptive sensor selection in sequential hypothesis testing," in *Proc. 50th IEEE Conf. Decision Control Eur. Control Conf.*, Dec. 2011, pp. 6284–6289.
- [19] V. Srivastava, K. Pilar, and F. Bullo, "Randomized sensor selection in sequential hypothesis testing," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2342–2354, May 2011.
- [20] D. R. J. Laming, *Information Theory of Choice-Reaction Times*. Oxford, U.K.: Academic, 1968.
- [21] R. Ratcliff, "A theory of memory retrieval," *Psychol. Rev.*, vol. 85, no. 2, p. 59, 1978.
- [22] R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen, "The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks," *Psychol. Rev.*, vol. 113, no. 4, p. 700, 2006.
- [23] R. Ratcliff and G. McKoon, "The diffusion decision model: Theory and data for two-choice decision tasks," *Neural Comput.*, vol. 20, no. 4, pp. 873–922, 2008.
- [24] A. Wald, "On cumulative sums of random variables," *Ann. Math. Statist.*, vol. 15, no. 3, pp. 283–296, Sep. 1944.
- [25] G. Casella and R. L. Berger, *Statistical Inference* (Duxbury Advanced Series). Boston, MA, USA: Brooks Cole, 2002.
- [26] G. V. Glass, P. D. Peckham, and J. R. Sanders, "Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance," *Rev. Edu. Res.*, vol. 42, no. 3, pp. 237–288, 1972.
- [27] W.-K. Shieh, L. J. Bain, and M. Engelhardt, "Test of equal gamma-distribution means with unknown and unequal shape parameters," *Technometrics*, vol. 30, no. 2, pp. 169–174, 1988.
- [28] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and its Applications: Theory of Majorization and its Applications*. New York, NY, USA: Springer, 2011.
- [29] S. Krishnaswamy, P. T. Akhil, A. Arapostathis, S. Shakkottai, and R. Sundaresan, "Augmenting max-weight with explicit learning for wireless scheduling with switching costs," in *Proc. IEEE INFOCOM*, May 2017, pp. 352–360.
- [30] N. K. Vaidhiyan and R. Sundaresan. (Aug. 2015). "Learning to detect an oddball target." [Online]. Available: <https://arxiv.org/abs/1508.05572>
- [31] P. Lezaud, "Chernoff-type bound for finite Markov chains," *Ann. Appl. Probab.*, vol. 8, no. 3, pp. 849–867, Aug. 1998.
- [32] Q. Wang, S. R. Kulkarni, and S. Verdú, "Divergence estimation for multidimensional densities via-nearest-neighbor distances," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2392–2405, May 2009.

Nidhin Koshy Vaidhiyan received his B.Tech degree in electronics and communication engineering from the College of Engineering Trivandrum, Master of Engineering and Ph.D. degrees in electrical communication engineering from the Indian Institute of Science Bangalore in 2005, 2009 and 2016, respectively. Since 2015, he is working as a Senior Engineer with Qualcomm India Pvt. Ltd., Bangalore.

S. P. Arun received his B.Tech from IIT Bombay, and MS & PhD from Johns Hopkins University, all in Electrical Engineering. From 2006–2009 he was a postdoctoral fellow at the Center for the Neural Basis of Cognition at the Carnegie Mellon University. Since 2010 he joined the Centre for Neuroscience at the Indian Institute of Science where he is currently an Associate Professor and an associate faculty of the Department of Electrical Communication Engineering. His research interests are in visual perception and object recognition.

Rajesh Sundaresan (S'96–M'00–SM'06) received the B.Tech. degree in electronics and communication from the Indian Institute of Technology Madras, the M.A. and Ph.D. degrees in electrical engineering from Princeton University in 1996 and 1999, respectively. From 1999 to 2005, he worked at Qualcomm Inc. on the design of communication algorithms for wireless modems. Since 2005, he has been with the Indian Institute of Science where he is currently a Professor in the Department of Electrical Communication Engineering and an associate faculty in the Robert Bosch Centre for Cyber-Physical Systems. His interests are in the areas of communication, computation, and control over networks. He was an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY for the period 2012–2015.