Learning to Detect an Oddball Target

Nidhin Koshy Vaidhiyan and Rajesh Sundaresan, *Senior Member, IEEE*

***Abstract*— We consider the problem of detecting an odd process among a group of Poisson point processes, all having the same rate except the odd process. The actual rates of the odd and non-odd processes are unknown to the decision maker. We consider a time-slotted sequential detection scenario where, at the beginning of each slot, the decision maker can choose which process to observe during that time slot. We are interested in policies that satisfy a given constraint on the probability of false detection. We propose a variation on a sequential policy based on the generalised likelihood ratio statistic. The policy, via suitable thresholding, can be made to satisfy the given constraint on the probability of false detection. Furthermore, we show that the proposed policy is asymptotically optimal in terms of the conditional expected stopping time among all policies that satisfy the constraint on the probability of false detection. The asymptotic is as the probability of false detection is driven to zero. We apply our results to a particular visual search experiment studied recently by neuroscientists. Our model suggests a neuronal dissimilarity index for the visual search task. The neuronal dissimilarity index, when applied to visual search data from the particular experiment, correlates strongly with the behavioural data. However, the new dissimilarity index performs worse than some previously proposed neuronal dissimilarity indices. We explain why this may be attributed to some experiment conditions.**

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

I. Introduction

Consider K homogeneous Poisson point processes. All processes except one, which we call the "odd" process, have the same rate. The actual rates of the odd process and the non-odd processes are unknown. The objective is to detect the odd (or anomalous or outlier) process as quickly as possible, but subject to constraints on the probability of false detection. For simplicity, we assume that time is divided into slots of fixed duration T. During a particular time slot, the decision maker can choose exactly one among the K processes for observation. This choice is made only at slot beginnings.

We cast the above problem into one of sequential detection with control [1], with the underlying parameters of the distributions being unknown [2]. The structural constraints in the problem, that exactly one among the K processes has a distribution different from the others, provides an opportunity to *learn* the underlying distributions from the observations, but the decision maker should learn just enough to make a reliable decision. This problem is a special case of that studied by Albert [2]. We shall discuss Albert's results in [2] in the coming sections.

We adapt the sample complexity result of Kaufmann et al. [3], developed for the best arm identification problem, to our setting and obtain a lower bound on the conditional expected stopping time for any policy that satisfies the constraint on the probability of false detection. This result is already available in Albert [2] and is given only for completeness. The key idea dates back to Chernoff [1]. The lower bound suggests that the conditional expected stopping time is asymptotically proportional to the negative of the logarithm of the probability of false detection. The proportionality constant is obtained as the solution to a max-min optimisation problem of relative entropies between the true system state (index of the odd process, its rate, and the rate of the non-odd processes) and other alternatives. The optimisation problem for the lower bound also suggests the nature of an asymptotically optimal strategy. A strategy proposed by Chernoff [1], its variation in Albert [2], and a further variation in this work will soon be described in the coming pages. Our variation's improvement over Albert's [2] will also be highlighted.

The usual methodology employed in problems with lack of exact knowledge of the underlying distributions is to use tests that are based on generalised likelihood ratios (GLR tests or GLRT). We work with a modification of the GLR statistic. Unlike the usual GLR statistic, we replace the maximum likelihood function in the numerator of the statistic by an average likelihood function, where the average is computed with respect to an artificial prior on the odd and non-odd rates. For the Poisson model, we employ a gamma distribution on the rates of the odd and non-odd processes as the prior, with the shape and rate parameter set to one. Any prior density having full support would suffice. The specific gamma prior allows easier characterisation of the averaged likelihood function. The averaging prevents over-estimation of the likelihood ratio function, and at the same time ensures that, asymptotically, the averaged version is not too far away from the true likelihood function. The modification allows us to design a time-invariant and simple threshold policy that satisfies the probability of false detection constraint. We show that the sampling strategy of the proposed policy (which of the K processes to observe at the beginning of each slot) converges to the sampling strategy suggested by the lower bound, where

Manuscript received August 23, 2015; revised February 15, 2017 and August 17, 2017; accepted November 5, 2017. Date of publication November 29, 2017; date of current version January 18, 2018. This work was supported in part by the Indo-French Centre for the Promotion of Advanced Research under Grant 5100-ITA and in part by the Science and Engineering Research Board, Department of Science and Technology, under Grant EMR/2016/002503.

N. K. Vaidhiyan is with Qualcomm India Pvt., Ltd., Bengaluru 560066, India.

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

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

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2017.2778264

the convergence is as the number of slots observed tends to infinity. We show that, asymptotically, the conditional expected stopping time of the proposed policy scales as $-\log(P_e)/D^*$, where P_e is the constraint on the probability of false detection and D^*, a relative entropy based constant, is the optimal scaling factor as suggested by the lower bound. D^* is obtained by solving a max-min problem, as mentioned above, and is made explicit in (5).

The motivation to study this problem comes from a visual search problem studied by Sripati and Olson [4], where a subject has to detect an odd image among a sea of distracter images "as quickly as possible without guessing" [4]. We model the visual search task as an oddball detection problem, as above, and propose D^* as a neuronal dissimilarity index for such visual search tasks. Smaller values of D^* result in harder search tasks and longer search times. We provide an exact asymptotic characterisation of the expected search time. We compare the performance of the proposed dissimilarity index with other dissimilarity indices proposed earlier by Vaidhiyan *et al.* [5]. In that paper, it was assumed that the odd and the non-odd rates were known. Here, we do not make this assumption and recognise that information about the non-odd and the odd processes must be gleaned along the way.

A. Prior Work

Chernoff [1] studied sequential hypothesis testing with control, termed active sequential hypothesis testing (ASHT) in [6] and [7], 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 called *Procedure A* which can be employed in very general contexts including the setting of this paper. Procedure A, at each time instant, chooses actions so that the relative entropy between the maximum likelihood estimate of the underlying parameter and its nearest alternative is maximised. Chernoff also proposed many variations – see [1, pp. 763 and 769]. However, Chernoff proved asymptotic optimality of Procedure A, as the cost of sampling went to 0, only in the restricted setting of a finite number of states of nature. Asymptotic optimality of Procedure A, or of other procedures, for more general settings with a continuum of parameter values, as in this paper, was left open by Chernoff [1]. The following works, relevant to this paper, extended Chernoff's work in various directions as we now highlight.

Albert [2] extended Chernoff's study to the case when the underlying true states of nature belonged to an infinite set. Since this is most relevant to us, we now discuss Albert's results in some detail. Albert assumed that the parameter space can be embedded in a compact topological space with certain additional properties. Albert introduced a relaxed version of Chernoff's Procedure A. The first relaxation parameter $\gamma_1 > 0$ provided a guaranteed minimum sampling rate of γ_1 for each action. This ensures that the maximum likelihood estimator (or its relaxation considered below) can be guaranteed to be consistent. The second parameter $\gamma_2 > 0$ made the stopping criterion more stringent by requiring the likelihood to be higher by a factor $1 + \gamma_2$. A third parameter $\rho \in (0, 1)$ relaxed the choice of the maximum likelihood location and picked any state of nature whose likelihood was within ρ of the maximum likelihood for further examination. This is essential because the maximum likelihood parameter estimate for a hypothesised oddball location may not exist. When $\gamma_1 = \gamma_2 = 1 - \rho = 0$, one recovered Chernoff's Procedure A from Albert's relaxed procedure. Albert's results in [2] can be summarised as follows:

Let c be the cost of sampling and let $R(\theta)$ be the total risk (sum of sampling cost and penalty for false detection) for a policy under the true state of nature θ.

- Converse: Albert showed that for any policy that guarantees a risk $R(\theta) = O(-c \log c)$, for every parameter θ, the risk is lower bounded by $R(\theta) \geq (1 + o(1)) (-c \log c)/D^*$. (The quantity D^* when specialised to our setting is given by (5).)
- Bound on probability of false detection: Albert's relaxed policy guaranteed a probability of false detection $P_e(\theta) \leq W(\theta, \gamma_1, \gamma_2, \rho)c$, where $W(\theta, \gamma_1, \gamma_2, \rho)$ is a constant dependent on $\theta, \gamma_1, \gamma_2$, and ρ. This dependence, particularly the dependence on θ and ρ, is something we are able to remove for our specific Poisson setting. We guarantee a probability of false detection independent of the underlying true state of nature θ and the relaxation parameter ρ.
- Bound on total risk: For any $\epsilon > 0$, there exist sufficiently small γ_1 and γ_2 such that $R(\theta) = (1 + \epsilon + o(1))(-c \log c)/D^*$ for the corresponding policy.

Albert [2, pp. 797–798] suggested how one may compactify the parameter space. His suggested procedure applies to the Poisson model under consideration, but the impact on the bounds, particularly on the nature of $W(\theta, \gamma_1, \gamma_2, \rho)$, are not clear. In the specific Poisson case, we are able to exploit the structure of the Poisson distribution and its conjugate prior to provide a precise guarantee on the probability of error at stoppage.

Chernoff [1, p. 768] remarked that his criterion of asymptotic optimality may not be relevant when the cost of sampling is high. One way to address this issue is to study the discounted cost setting. Another way is to study nonasymptotic cost minimisation. In a series of works, Naghshvar and Javidi [6]–[10] took the latter approach and studied ASHT from a Bayesian cost minimisation perspective. Similarly, Nitinawarat *et al.* [11], [12] studied ASHT from the perspective of minimising the conditional expected cost (generally stopping delay) subject to constraints on the probability of false detection. The latter works assumed knowledge of the underlying distributions under different hypotheses.

In the context of fixed sample size testing, as against the current context of sequential testing, Li *et al.* [13] studied outlier detection under unknown typical and outlier distributions, and thus dealt with composite hypotheses. They assumed that the observation space is finite. They also assumed simultaneous observability of all processes at each observation instance. They then proposed a GLRT with thresholds varying with sample size, and showed that it has, asymptotically, the same error exponent as that of an optimal algorithm possessing knowledge of the underlying distributions. The asymptotics was as the number of processes available for observation tended to infinity. They termed such algorithms *asymptotically exponentially consistent*. Further, they extended their study to the setting where there are more than one outlier processes. They showed that their extended algorithm is asymptotically exponentially consistent in the new setting. As for the sequential setting, Li *et al.* [14] studied sequential versions of [13] and showed that another modified GLRT that keeps sampling until the test statistic crosses a threshold is universally consistent as the threshold is increased to infinity. In both these works, unlike the ASHT setting and unlike our setting, observations from all processes are available to the decision maker at each observation instance.

Nitinawarat and Veeravalli [15] studied an outlier detection problem in a setting similar to ours, where at each observation instance, the decision maker is allowed to observe only one of the processes. But different from our setting, they assume knowledge of the typical (or non-odd) distribution. They proposed an algorithm that was shown to have vanishing probability of false detection as the threshold is increased to infinity. Further, the proposed algorithm was shown to have, asymptotically, the same error exponent as that of an optimal policy with knowledge of the atypical (odd) distribution. Recently, Cohen and Zhao [16] studied a problem similar to ours, but restricted their study to the setting when the atypical (odd) and typical (non-odd) distributions belonged to disjoint sets. Consequently, in their setting, the optimal action at each decision instance is to observe the process that has the generalised maximum likelihood with respect to the set of atypical (odd) distributions. Their proposed policy also had a threshold based stopping criterion. They showed that their policy has the same asymptotic scaling for the conditional expected stopping time as for an optimal policy with knowledge of the distributions. Unlike the results in [15] and [16], we shall see that the information structures on the odd and non-odd distributions in this paper are such that lack of knowledge of the exact distributions leads to a distinct loss in performance.

A related problem, which has seen a resurgence in interest in the machine learning community, is the problem of identification of the *best arm* for multi-armed bandits. Indeed, Chernoff [1, p. 758] and Albert [2, p. 775] had already used this as their prototype example to describe their results. Asymptotic optimality of Procedure A was however not established. Kaufmann *et. al.* [3] recently studied the sample complexity of the best arm identification problem. The problem considered in this paper can be cast as an odd-arm identification problem as against the best-arm identification problem. The composite hypotheses structures in the problems are different.

B. Our Contribution

Our asymptotically optimal algorithm and results differ from prior works in the following aspects:

- Unlike the works on ASHT [6]–[12], we do not assume complete knowledge of the underlying distribution under different hypotheses. Unlike the analysis available in Chernoff [1], but as in Albert [2], our analysis is for the setting where the number of states of nature is a continuum. However, we consider an adaptation of Chernoff's Procedure A that is different from Albert's [2].
- Our modification to Procedure A involves a modified GLR that uses a fictitious Bayesian prior. Testing with this modified GLR ensures that a given false detection constraint can be met when the statistic is compared with a fixed threshold. The false detection rate is independent of the true state of nature. This is a significant difference over Albert's work [2]. We believe our proposed modification may be useful in other settings as well. See the last bullet.
- We show that, unlike in Albert [2], there is no need for an extra parameter γ_1 that forces an exploration of the action space for consistency of the estimates. Instead, we show that for the Poisson case, the inherent structure automatically ensures this exploration without need for an additional parameter; see Figure 1.
- Unlike the works of Li *et al.* [13], [14], our observations are limited by the chosen actions. There is then a clear exploration versus exploitation tradeoff.
- Unlike the work of Nitinawarat and Veeravalli [15], we do not assume knowledge of the atypical (odd) distribution, nor do we assume the typical (non-odd) distribution.
- Unlike the work of Cohen and Zhao [16], we do not assume that the atypical and typical distributions belong to disjoint sets.
- We specifically consider the setting of Poisson point processes mainly because of our desire to explain the experimental observations of Sripati and Olson [4] on neuronal data which are modelled as Poisson point processes in [5]. Nevertheless, we believe that the same ideas may carry forward to other class of distributions, especially exponential families. Indeed, there has already been a more recent work that uses our modification to the GLRT to solve the best arm identification problem in the more general setting of exponential families [17].

C. Organisation

In Section II, we develop the required notation and describe the model. In Section III, we provide a lower bound on the conditional expected stopping time for any policy that satisfies the probability of false detection constraint. The nature of the lower bound suggests a candidate asymptotically optimal policy. In the same section, we make some observations on some structural properties of the suggested policy. In Section IV, we formally propose the policy and show that it is asymptotically optimal. In Section V, we discuss some simulation results to corroborate our theoretic results. In Section VI, we apply the theory to visual search. We show that the proposed neuronal dissimilarity index is strongly correlated with the behavioural data. In Section VII, we make some concluding remarks and discuss possible extensions. Most proofs are relegated to appendices VII and VII.

II. Model

In this section we develop the required notation and describe the model.

Let $K \geq 3$ denote the number of Poisson point processes under consideration. Conditioned on the rates, the processes are assumed to be independent of each other. Let H, $1 \leq H \leq K$, denote the index of the odd process. Let $R_1 > 0$ denote the unknown rate of the odd process, and let $R_2 > 0$ denote the unknown rate of the non-odd processes. We assume $R_1 \neq R_2$. Let the triplet $\Psi = (H, R_1, R_2)$ denote the configuration of the processes, where the first component represents the index of the odd process, while the second and third components represent the odd and non-odd rates respectively. Let T denote the time duration of a time slot. Without loss of generality we can assume $T = 1$, the analysis holds for general T with an appropriate scaling of the rates. The analysis can be done in continuous time as well, but we shall take the simpler slotted time approach.

Given the Poisson process assumption, a sufficient statistic for the observed process during a time slot is the number of jumps observed in that time slot. Let $A_n \in \{1, 2, \ldots, K\}$ denote the index of the process chosen for observation in time slot n, and let $X_n \in \mathbb{Z}_+$ denotes the number of jumps observed in the process during time slot n. 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 $(X_1, X_2, \ldots, X_n)$ and A^n for $(A_1, A_2, \ldots, A_n)$. We also write $\mathcal{P}(K)$ for the set of probability distributions on $\{1, 2, \ldots, K\}$.

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 CA_n that is either $(stop, \delta)$ or $(continue, \lambda)$ as explained next. If the composite action is $(stop, \delta)$, then the detector stops taking further samples (or retires) and indicates δ as its decision on the hypotheses; $\delta \in \{1, 2, \ldots, K\}$. If the composite action is $(continue, \lambda)$, the detector picks the next process to observe A_n according to the distribution $\lambda \in \mathcal{P}(K)$. The stopping time is defined as

$$\tau := \inf\{n \geq 1 : CA_n = (stop, \cdot)\}.$$

Consider a policy π. Conditioned on action A_n, the true hypothesis H, and the odd and non-odd rates R_1 and R_2, we assume that the observation X_n is independent of previous actions A^{n-1}, previous observations X^{n-1}, and the policy. The conditional distribution of X_n, given the current action A_n, the configuration $\Psi = (H, R_1, R_2)$, the history X^{n-1}, A^{n-1}, and the Poisson assumption, is given by

$$\begin{aligned} P(X_n = l|\Psi = (H, R_1, R_2), A_n, X^{n-1}, A^{n-1}) \\ = P(X_n = l|\Psi = (H, R_1, R_2), A_n) \quad (1) \\ = \begin{cases} \frac{R_1^l e^{-R_1}}{l!} & \text{if } A_n = H \\ \frac{R_2^l e^{-R_2}}{l!} & \text{if } A_n \neq H, \end{cases} \quad (2) \end{aligned}$$

where $l \in \mathbb{Z}_+$.

Let E^π denote the conditional expectation and let P^π denote the conditional probability measure, given Ψ, under the policy π. Given an error tolerance vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, with $0 < \alpha_i < 1$, let $\Pi(\alpha)$ be the set of desirable policies defined as

$$\begin{aligned} \Pi(\alpha) := \{\pi : P^\pi(\delta \neq i|\Psi = (H, R_1, R_2), H = i) \leq \alpha_i, \\ \text{for all } i \text{ and for all } \Psi \text{ such that } R_1 \neq R_2\}. \quad (3) \end{aligned}$$

Let $\|\alpha\|$ denote $\max_i \alpha_i$.

For ease of notation, we drop the superscript π while writing E^π, P^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.

III. The Converse-Lower Bound

In this section we develop a lower bound on the conditional expected stopping time for any policy that belongs to $\Pi(\alpha)$. While optimal results for finite $\|\alpha\|$ are difficult to characterise, the asymptotic $\|\alpha\| \to 0$ puts us in the regime of a large number of samples where asymptotically optimal results are more easily available (see [1, p. 755]). We show that, as $\|\alpha\| \to 0$, the lower bound scales as $-\log(\|\alpha\|)/D^*$. We also characterise D^* in detail in this section. Section IV provides an asymptotically optimal policy whose conditional expected stopping time is close to this lower bound.

The following proposition gives the lower bound. Its proof may be seen as an application of the data processing inequality ([18, p. 16], [19]) for relative entropy.

Proposition 1: Fix α, with $0 < \alpha_i < 1$ for each i. Let $\Psi = (i, R_1, R_2)$ be the true configuration. For any $\pi \in \Pi(\alpha)$, we have

$$E^\pi[\tau|\Psi] \geq \frac{d_b(\|\alpha\|, 1 - \|\alpha\|)}{D^*(i, R_1, R_2)}, \quad (4)$$

where $d_b(\|\alpha\|, 1 - \|\alpha\|)$ is the binary relative entropy function defined as

$$d_b(x, 1 - x) := x \log(x/(1 - x)) + (1 - x) \log((1 - x)/x),$$

and $D^*(i, R_1, R_2)$ is defined as

$$\begin{aligned} D^*(i, R_1, R_2) \\ := \max_{\lambda \in \mathcal{P}(K)} \min_{R_1' > 0, R_2' > 0, j \neq i} \big[\lambda(i) D(R_1 \| R_2') \\ + \lambda(j) D(R_2 \| R_1') + (1 - \lambda(i) - \lambda(j)) D(R_2 \| R_2')\big], \end{aligned} \quad (5)$$

where $D(x\|y) := x \log(x/y) - x + y$ is the KL-divergence or relative entropy between two Poisson random variables with means x and y.

Let $\lambda^*(i, R_1, R_2)$ denote the $\lambda \in \mathcal{P}(K)$ that maximises (5), i.e.,

$$\begin{aligned} \lambda^*(i, R_1, R_2) \\ = \arg \max_{\lambda \in \mathcal{P}(K)} \min_{R_1', R_2', j \neq i} \big[\lambda(i) D(R_1 \| R_2') \\ + \lambda(j) D(R_2 \| R_1') + (1 - \lambda(i) - \lambda(j)) D(R_2 \| R_2')\big]. \end{aligned} \quad (6)$$

We can interpret $D^*(i, R_1, R_2)$ as the minimum among relative entropy rates between the true configuration $\Psi = (i, R_1, R_2)$ and all other possible alternate configurations

$\Psi' = (j, R_1', R_2')$, with $j \neq i$, but maximised over all policies (action strategies) that pick actions in an independent and identically distributed (i.i.d.) manner. It can also be interpreted as the max-min-drift of the log likelihood ratio process between the true configuration and other error configurations, the minimum being over all possible error configurations, and the maximum being over all i.i.d. policies. $D^*(i, R_1, R_2)$ is the key information quantity in this paper. Since $d_b(\|\alpha\|, 1 - \|\alpha\|)/\log(\|\alpha\|) \to -1$ as $\|\alpha\| \to 0$, Proposition 1 shows that the conditional expected stopping time of the optimal policy scales at least as $-\log(\|\alpha\|)/D^*(i, R_1, R_2)$ as the probability of false detection constraint $\|\alpha\| \to 0$. In Section IV we will describe a policy that is upper bounded by, and therefore achieves, a similar scaling, though only asymptotically as $\|\alpha\| \to 0$.

Proof of Proposition 1: Assume $E^\pi[\tau|\Psi]$ is finite, for otherwise (4) is trivially true. We apply the sample complexity result of Kaufmann *et al.* [3, Lemma 1] to our setting. Let $N_j(\tau) = \sum_{k=1}^{\tau} 1_{\{A_k = j\}}$ denote the number of samples from process j observed till the stopping time τ. Clearly, $\tau = \sum_{j=1}^{K} N_j(\tau)$. Kaufmann et al. [3, Lemma 1] showed that, for any $\pi \in \Pi(\alpha)$, conditioned on the true configuration $\Psi = (i, R_1, R_2)$, and for any alternate configuration $\Psi' = (j, R_1', R_2')$, $j \neq i$, the conditional expected sample sizes satisfy

$$E^\pi[N_i(\tau)|\Psi]\, D(R_1\|R_2') + E^\pi\left[N_j(\tau)|\Psi\right] D(R_2\|R_1') + \left(\sum_{k\neq i, k\neq j} E^\pi[N_k(\tau)|\Psi]\right) D(R_2\|R_2') \geq d_b(\|\alpha\|, 1 - \|\alpha\|). \tag{7}$$

This is a consequence of the convexity of relative entropy and the data processing inequality. Multiplying and then dividing the left-hand side by $E^\pi[\tau|\Psi]$, we get

$$\begin{aligned} &d_b(\|\alpha\|, 1 - \|\alpha\|) \\ &\leq E^\pi[\tau|\Psi]\Bigg[\frac{E^\pi[N_i(\tau)|\Psi]}{E^\pi[\tau|\Psi]} D(R_1\|R_2') + \frac{E^\pi\left[N_j(\tau)|\Psi\right]}{E^\pi[\tau|\Psi]} D(R_2\|R_1') \\ &\quad + \left(1 - \frac{E^\pi[N_i(\tau)|\Psi] + E^\pi\left[N_j(\tau)|\Psi\right]}{E^\pi[\tau|\Psi]}\right) D(R_2\|R_2')\Bigg]. \end{aligned} \tag{8}$$

Since (8) holds for any R_1', R_2' and $j \neq i$, and since $E^\pi[\tau|\Psi]$ does not depend on R_1', R_2' and $j \neq i$, we can choose the tightest bound and get

$$\begin{aligned} &d_b(\|\alpha\|, 1 - \|\alpha\|) \\ &\leq E^\pi[\tau|\Psi] \min_{R_1', R_2', j\neq i} \Bigg[\frac{E^\pi[N_i(\tau)|\Psi]}{E^\pi[\tau|\Psi]} D(R_1\|R_2') + \frac{E^\pi\left[N_j(\tau)|\Psi\right]}{E^\pi[\tau|\Psi]} D(R_2\|R_1') \\ &\quad + \left(1 - \frac{E^\pi[N_i(\tau)|\Psi] + E^\pi\left[N_j(\tau)|\Psi\right]}{E^\pi[\tau|\Psi]}\right) D(R_2\|R_2')\Bigg] \end{aligned} \tag{9}$$

$$\begin{aligned} &\leq E^\pi[\tau|\Psi] \max_{\lambda\in\mathcal{P}(K)} \min_{R_1', R_2', j\neq i} \big[\lambda(i) D(R_1\|R_2') \\ &\quad + \lambda(j) D(R_2\|R_1') + (1 - \lambda(i) - \lambda(j)) D(R_2\|R_2')\big]. \end{aligned} \tag{10}$$

The last inequality follows because maximisation over all $\lambda \in \mathcal{P}(K)$ only increases the right-hand side. This completes the proof. ■

We now describe some simplifications for $D^*(i, R_1, R_2)$ and $\lambda^*(i, R_1, R_2)$. We show that the K-dimensional optimisation in (5) can be reduced to a one-dimensional optimisation, which can be easily solved via say a simple line search. This has positive implications on the complexity of a policy achieving the above lower bound, which we will discuss in the next section.

Proposition 2: Consider K Poisson point processes with configuration $\Psi = (i, R_1, R_2)$. The quantity $D^*(i, R_1, R_2)$ of (5) can be equivalently expressed as

$$D^*(i, R_1, R_2) = \max_{0\leq\lambda(i)\leq 1}\left[\lambda(i) D(R_1\|\tilde{R}) + (1 - \lambda(i))\frac{(K-2)}{(K-1)} D(R_2\|\tilde{R})\right], \tag{11}$$

where

$$\tilde{R} = \frac{\left(\lambda(i)R_1 + (1 - \lambda(i))\frac{(K-2)}{(K-1)} R_2\right)}{\left(\lambda(i) + (1 - \lambda(i))\frac{(K-2)}{(K-1)}\right)}. \tag{12}$$

Also, $\lambda^*(i, R_1, R_2)$ is of the form

$$\lambda^*(i, R_1, R_2)(j) = \begin{cases} \lambda^*(i, R_1, R_2)(i) & \text{if } j = i \\ (1 - \lambda^*(i, R_1, R_2)(i))/(K-1) & \text{if } j \neq i. \end{cases} \tag{13}$$

Proof: Consider (5). Observe that R_1' appears only in the middle term on the right-hand side. This is minimised when $R_1' = R_2$ and the minimum value is zero. We therefore have

$$\begin{aligned} D^*(i, R_1, R_2) &= \max_{\lambda\in\mathcal{P}(K)} \min_{R_2', j\neq 1} \big[\lambda(i) D(R_1\|R_2') \\ &\qquad + (1 - \lambda(i) - \lambda(j)) D(R_2\|R_2')\big] \qquad (14) \\ &= \max_{0\leq\lambda(i)\leq 1} \min_{R_2'} \Bigg[\lambda(i) D(R_1\|R_2') \\ &\qquad + (1 - \lambda(i))\frac{(K-2)}{(K-1)} D(R_2\|R_2')\Bigg]. \qquad (15) \end{aligned}$$

Equation (15) follows from the fact that the λ that maximises (14) will have equal mass on all locations other than i, i.e., the maximiser λ^* will satisfy $\lambda^*(j) = (1 - \lambda^*(i))/(K-1)$, for all $j \neq i$.

For a fixed $\lambda(i)$, to find the R_2' that minimises the term within brackets in (15) which is a strictly convex function of R_2', we take its derivative with respect to R_2' and equate it to zero. We then see that the minimising R_2' satisfies the equation

$$\lambda(i) D'(R_1\|R_2') + (1 - \lambda(i))\frac{(K-2)}{(K-1)} D'(R_2\|R_2') = 0, \tag{16}$$

where $D'(x\|y)$ is the derivative of $D(x\|y)$ with respect to the second argument y, which turns out to be $1 - x/y$. The R_2' thus obtained is

$$R_2' = \frac{\left(\lambda(i)R_1 + (1-\lambda(i))\frac{(K-2)}{(K-1)}R_2\right)}{\left(\lambda(i) + (1-\lambda(i))\frac{(K-2)}{(K-1)}\right)}. \tag{17}$$

This completes the proof. ■

As we will see, $\lambda^*(i, R_1, R_2)$ can be interpreted as the distribution on the set of actions of the optimal i.i.d. policy that achieves $D^*(i, R_1, R_2)$. Heuristically, a good policy would attempt to have an action process whose empirical measure on the set of actions approaches the distribution $\lambda^*(i, R_1, R_2)$, as $\|\alpha\| \to 0$. A closed form expression for $\lambda^*(i, R_1, R_2)$ is not available. But we now describe some structural properties of $\lambda^*(i, R_1, R_2)$. In particular, we show for any configuration Ψ, all components of $\lambda^*(\Psi)$ are strictly bounded away from zero.

Proposition 3: Fix $K \geq 3$. Let λ^* be as in (6). There exists a constant $c_K \in (0, 1)$, independent of (k, θ_1, θ_2) but dependent on K, such that

$$\lambda^*(k, \theta_1, \theta_2)(j) > c_K > 0$$

for all $j \in \{1, 2, \ldots, K\}$ and for all (k, θ_1, θ_2) such that $\theta_1 > 0, \theta_2 > 0$ and $\theta_1 \neq \theta_2$.

Proof: See Appendix A. ■

Proposition 3 suggests that a good policy should sample each process at least c_K fraction of the time. Estimates of the rate of each process should then converge to the corresponding true rate. We will make use of this fact in the analysis of our proposed algorithm, which is to come shortly.

An explicit expansion of the objective function in (11) shows that $\lambda^*(k, \theta_1, \theta_2)(k)$ is invariant to joint scaling of (θ_1, θ_2). It can therefore be expressed in terms of the ratio $\nu = \theta_1/(\theta_1 + \theta_2)$ as $\lambda^*(k, \nu, 1-\nu)(k)$. Fig. 1 shows the value of $\lambda^*(k, \theta_1, \theta_2)(k)$ for different values of ν and for different K, K varying from 3 to 1000 and ∞. We observe the following:

1) $\lambda^*(k, \theta_1, \theta_2)(k)$ is lower bounded by ~ 0.3 for all ν and for all K, and $\lambda^*(k, \theta_1, \theta_2)(k)$ attains its minimum at $\nu = 1$ and for $K = 3$.
2) $\lambda^*(k, \theta_1, \theta_2)(k)$ is upper bounded by ~ 0.7 for all ν and for all K, and the maximum is approached at $\nu = 0$ and as $K \to \infty$.
3) At $\nu = 1/2$, we have $R_1 = R_2$; the objective function in (11) is identically zero, and any $\lambda(k)$ works. We may take $\lambda^*(k)$ to be the continuous extension of $\lambda^*(k)$ as $\nu \to 1/2$.

From the above observations, for a fixed K, we have $\lambda^*(k, \theta_1, \theta_2)(j) \gtrsim (0.3/(K-1))$ for all j and for all (k, θ_1, θ_2). In Appendix A, where we prove Proposition 3, we obtain a looser bound for $\lambda^*(k, \theta_1, \theta_2)(j)$. We only show that $\lambda^*(k, \theta_1, \theta_2)(j) > 0.1/(K-1)$.

IV. Achievability-Modified GLRT

In this section we describe our proposed asymptotically optimal policy that achieves the lower bound in Proposition 1 as the constraint on the probability of false detection is driven to zero. Our algorithm is an adaptation of Chernoff's

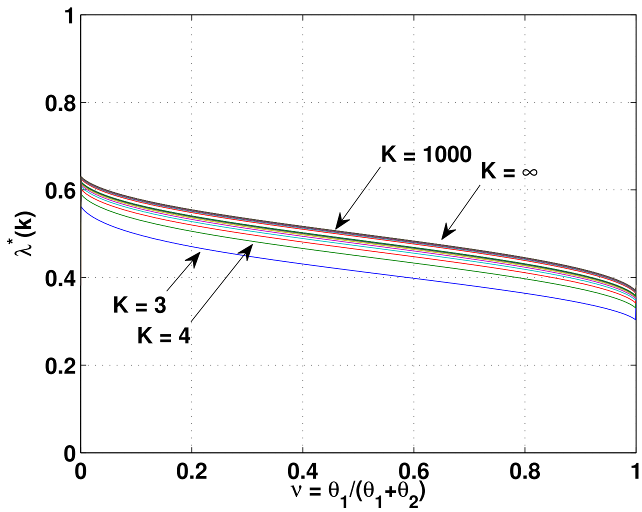

Fig. 1. $\lambda^*(k, \theta_1, \theta_2)(k)$ versus $\nu = \theta_1/(\theta_1 + \theta_2)$ for various K.

Procedure A. The likelihood ratio function in *Procedure A* is replaced by a modified generalised likelihood ratio function in our algorithm. The strategy at each time slot is not only a function of the hypothesis with the largest GLR statistic, but also a function of the maximum likelihood estimates of the odd and non-odd rates.

Before describing the algorithm, we develop some required notation.

Let N_j^n denote the number of times process j was chosen for observation up to time n, i.e., $N_j^n = \sum_{t=1}^n 1_{\{A_t = j\}}$ and so $n = \sum_{j=1}^K N_j^n$. Let Y_j^n denote the number of observed jumps in process j up to time n; $Y_j^n = \sum_{t=1}^n X_t 1_{\{A_t=j\}}$. Let Y^n denote the total number of observed jumps up to time n; $Y^n = \sum_{j=1}^K Y_j^n$.

Let $f(X^n, A^n | \Psi = (j, \theta_1, \theta_2))$ be the likelihood function of the observations and actions up to time n, conditioned on the configuration Ψ, i.e.,

$$\begin{aligned} &f(X^n, A^n | \Psi = (j, \theta_1, \theta_2)) \\ &\quad = \frac{1}{\prod_{t=1}^n (X_t!)} \, \theta_1^{Y_j^n} e^{-N_j^n \theta_1} \, \theta_2^{(Y^n - Y_j^n)} e^{-(n - N_j^n)\theta_2}. \end{aligned} \tag{18}$$

This reflects the assumption that the observations of the number of jumps (neuronal spikes) follow the Poisson distribution. When the parameters are unknown, there is a natural conjugate prior distribution for these parameters which enables easy updates of the posterior distribution based on observations. This prior is the gamma distribution which requires a pair of parameters. Since we have two Poisson distributions, coming from the odd and the non-odd processes, we need a total of four parameters. Let these four parameters be $\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}$, all fixed constants and all greater than zero. Let

$$\begin{aligned} &f_{\beta_{11},\beta_{12},\beta_{21},\beta_{22}}(\Psi = (j, \theta_1, \theta_2) | H = j) \\ &\quad := f_{\beta_{11},\beta_{12}}(\theta_1 | H = j) \; f_{\beta_{21},\beta_{22}}(\theta_2 | H = j) \qquad (19) \\ &\quad := \frac{\beta_{12}^{\beta_{11}} \theta_1^{\beta_{11}-1} e^{-\beta_{12}\theta_1}}{\Gamma(\beta_{11})} \; \frac{\beta_{22}^{\beta_{21}} \theta_2^{\beta_{21}-1} e^{-\beta_{22}\theta_2}}{\Gamma(\beta_{21})} \qquad (20) \end{aligned}$$

denote the product gamma densities on the parameters θ_1 and θ_2. We will use $f_{1,1,1,1}(\Psi = (j, \theta_1, \theta_2) | H = j)$ as an *artificial prior* on the parameter space $\Theta = \{(\theta_1, \theta_2)\}$ in our proposed algorithm. While any positive $(\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22})$ would suffice, $(\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}) = (1, 1, 1, 1)$ makes the calculations and the presentation simpler. θ_1 and θ_2 then, initially, have the exponential distribution with mean 1.

Let $\hat{\theta}_j^n = (\hat{\theta}_{j,1}^n, \hat{\theta}_{j,2}^n)$ denote the maximum likelihood estimates of the odd and non-odd rates at time n conditioned on $H = j$, i.e.,

$$\hat{\theta}_{j,1}^n = \frac{Y_j^n}{N_j^n} \text{ and } \hat{\theta}_{j,2}^n = \frac{(Y^n - Y_j^n)}{(n - N_j^n)}. \tag{21}$$

We now substitute this into the likelihood function and let

$$\hat{f}\left(X^n, A^n | H = j\right) := \max_{\Psi: H=j} f\left(X^n, A^n | \Psi\right) \tag{22}$$

$$= f\left(X^n, A^n | \Psi = (j, \hat{\theta}_{j,1}^n, \hat{\theta}_{j,2}^n)\right) \tag{23}$$

$$= \frac{1}{\prod_{t=1}^n (X_t!)} \left(\frac{Y_j^n}{N_j^n}\right)^{Y_j^n} e^{-Y_j^n} \times \left(\frac{Y^n - Y_j^n}{n - N_j^n}\right)^{(Y^n - Y_j^n)} e^{-(Y^n - Y_j^n)} \tag{24}$$

denote the maximum likelihood of the observations and actions till time n conditioned on $H = j$. The maximum is taken over all possible odd and non-odd rates, and the best choices are those in (21). Let the averaged likelihood function at time n, averaged according to the artificial prior $f_{1,1,1,1}$ over all configurations Ψ given $H = i$, be

$$f(X^n, A^n | H = i) := \int f(X^n, A^n | \Psi = (i, \theta_1, \theta_2)) \times f_{1,1,1,1}((i, \theta_1, \theta_2) | H = i) d\theta_1 d\theta_2 \tag{25}$$

$$= \frac{1}{\prod_{t=1}^n (X_t!)} \int \theta_1^{Y_i^n} e^{-N_i^n \theta_1} \theta_2^{(Y^n - Y_i^n)} e^{-(n - N_i^n)\theta_2} \times e^{-\theta_1} e^{-\theta_2} d\theta_1 d\theta_2 \tag{26}$$

$$= \frac{1}{\prod_{t=1}^n (X_t!)} \frac{\Gamma(Y_i^n + 1)}{(N_i^n + 1)^{(Y_i^n + 1)}} \times \frac{\Gamma(Y^n - Y_i^n + 1)}{(n - N_i^n + 1)^{(Y^n - Y_i^n + 1)}}, \tag{27}$$

where the second equality follows after substitution of (18) and (20), and the last equality follows by recognising the presence of Gamma$(Y_i^n + 1, N_i^n + 1)$ and Gamma$(Y^n - Y_i^n + 1, n - N_i^n + 1)$ densities without scale factors in (27). The modified GLR is defined as

$$Z_{ij}(n) := \log\left(\frac{f(X^n, A^n | H = i)}{\hat{f}(X^n, A^n | H = j)}\right) \tag{28}$$

$$= \log\left(\frac{\Gamma(Y_i^n + 1)}{(N_i^n + 1)^{(Y_i^n + 1)}} \cdot \frac{\Gamma(Y^n - Y_i^n + 1)}{(n - N_i^n + 1)^{(Y^n - Y_i^n + 1)}}\right) - Y_j^n\left(\log\left(\frac{Y_j^n}{N_j^n}\right) - 1\right) - (Y^n - Y_j^n)\left(\log\left(\frac{(Y^n - Y_j^n)}{(n - N_j^n)}\right) - 1\right). \tag{29}$$

Note that the numerator is an averaged likelihood under $H = i$, averaged with respect to an artificial prior, and denominator is a maximum likelihood under $H = j$. This modification of the GLR will prove to be crucial in demonstrating that the following policy that uses this statistic meets the error tolerance criterion. Let

$$Z_i(n) := \min_{j \neq i} Z_{ij}(n) \tag{30}$$

denote the modified GLR of i against its nearest alternate.

We now describe our proposed policy.

Policy: Modified GLRT ($\pi_M(L)$)

Fix $L \geq 1$. At time n (end of slot n):

- Let $i^*(n) = \arg\max_i Z_i(n)$, the index with the largest modified GLR after n time slots. Ties are resolved uniformly at random.
- If $Z_{i^*(n)}(n) < \log((K - 1)L)$ then A_{n+1} is chosen according to $\lambda^*(i^*(n), \hat{\theta}_{i^*(n)1}^n, \hat{\theta}_{i^*(n)2}^n)$, i.e.,

$$\Pr(A_{n+1} = j | X^n, A^n) = \lambda^*(i^*(n), \hat{\theta}_{i^*(n)1}^n, \hat{\theta}_{i^*(n)2}^n)(j). \tag{31}$$

- If $Z_{i^*(n)}(n) \geq \log((K - 1)L)$ then the test retires and declares $i^*(n)$ as the oddball location.

We now make some remarks on the complexity of this policy. The maximum likelihood estimates (21) of the odd and non-odd rates must be computed for every hypothesis. These being averages can be done in an iterated fashion via $\overline{\xi}_{n+1} := (n+1)^{-1} \sum_{i=1}^{n+1} \xi_i = (1-(n+1)^{-1})\overline{\xi}_n + (n+1)^{-1}\xi_{n+1}$ in $O(K)$ steps. The computation of $Z_{ij}(n)$ in (29) can also be done iteratively and requires $O(K^2)$ steps. The computation of the modified GLR requires another $O(K^2)$ since computation of (30) requires K comparisons, and this must be done for each hypothesis. The index of current best hypothesis requires $O(K)$ comparisons. The check for stopping requires $O(1)$ steps, and if the process continues, the identification of the best sampling distribution is a one-dimensional bounded convex optimisation and can be done easily via a line search in $O(1)$ steps. Finally, the sampling itself requires $O(K)$ steps. Adding these up we see that the complexity at every slot is $O(K^2)$ steps.

As done in [5], we also consider two variants of $\pi_M(L)$ which are useful in the analysis.

- *Policy* $\pi_M^i(L)$: This is the same as $\pi_M(L)$, but stops only at decision i when $Z_i(n) \geq \log((K - 1)L)$.
- *Policy* $\tilde{\pi}_M$: This is the same as $\pi_M(L)$, but never stops, and hence L is irrelevant.

Under a fixed hypothesis $H = i$, and the triplet of policies $(\pi_M(L), \pi_M^i(L), \tilde{\pi}_M(L))$, 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. The coupling is as follows. They all act on the same infinite sequence of samples derived from the randomised sampling policy (31). Policy $\pi_M(L)$ stops first. If it decides i at stoppage, then $\pi_M^i(L)$ also stops; otherwise $\pi_M^i(L)$ continues until it reaches an epoch where it decides i. The policy $\tilde{\pi}_M(L)$ continues to sample forever.

We denote the stopping times by $\tau(\pi_M(L))$ and $\tau(\pi_M^i(L))$, respectively. Under the above coupling, the following conclusions are obvious:

$$\tau(\pi_M^i(L)) \geq \tau(\pi_M(L)),$$
$$\{\tau(\pi_M(L)) > n\} \subseteq \{\tau(\pi_M^i(L)) > n\} \subseteq \{Z_i(n) < \log((K-1)L)\}.$$

We now explore the characteristics of the proposed policy $\pi_M(L)$.

Proposition 4: Fix $L > 1$. Policy $\pi_M(L)$ stops in finite time with probability 1, that is, $P(\tau(\pi_M(L)) < \infty) = 1$.

Proof: See Appendix B-A. ■

The main idea of the proof is as follows. We argue that, when the odd process has index i, i.e., $H = i$, the test statistic $Z_i(n)$ has a strictly positive drift and hence will cross the threshold $\log((K-1)L)$ in finite time almost surely. The details are given in Appendix B-A.

We next show that for any α, the policy $\pi_M(L)$, with L chosen to meet the smallest constraint, belongs to $\Pi(\alpha)$, and so $\pi_M(L)$ satisfies the constraint on the probability of false detection.

Proposition 5: Fix $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_K)$. Let $L = 1/\min_k \alpha_k$. We then have $\pi_M(L) \in \Pi(\alpha)$.

Proof: From the choice of L, we have $1/L \leq \alpha_k$ for all $k \in \{1, 2, \ldots, K\}$. This implies $\Pi((1/L, 1/L, \ldots, 1/L)) \subseteq \Pi(\alpha)$. Hence, it suffices to show that $\pi_M(L) \in \Pi((1/L, 1/L, \ldots, 1/L))$.

Fix $\Psi = (i, R_1, R_2)$. Let $\Delta_j^n = \{\omega : \tau(\pi_M(L))(\omega) = n, \delta(\omega) = j\}$ denote the sample paths for which the decision maker stops sampling after n time slots and decides in favour of $H = j$. The decision region in favour of j is denoted $\Delta_j := \cup_{n\geq 1} \Delta_j^n$. Note that

$$\Delta_j^n \cap \Delta_j^m = \emptyset \quad \text{for all } m \neq n. \tag{32}$$

We now use a standard change of measure argument to bound the conditional probability of false detection as follows, with P in place of P^{π_M}:

$$\begin{aligned}
&P(\delta \neq i | \Psi = (i, R_1, R_2)) \\
&= \sum_{j\neq i} P(\delta = j | \Psi = (i, R_1, R_2)) \\
&\quad + P(\tau(\pi_M(L)) = \infty | \Psi = (i, R_1, R_2)) \\
&= \sum_{j\neq i} \sum_{n\geq 1} \int_{\omega\in\Delta_j^n} dP(\omega | \Psi = (i, R_1, R_2)) + 0 \qquad (33) \\
&= \sum_{j\neq i} \sum_{n\geq 1} \int_{\omega\in\Delta_j^n} f(x^n, a^n | \Psi = (i, R_1, R_2))) d(x^n, a^n) \\
&\leq \sum_{j\neq i} \sum_{n\geq 1} \int_{\omega\in\Delta_j^n} \left(\frac{\hat{f}(x^n, a^n | H = i)}{f(x^n, a^n | H = j)} \right) \\
&\quad \times f(x^n, a^n | H = j) d(x^n, a^n) \qquad (34) \\
&\leq \sum_{j\neq i} \frac{1}{L(K-1)} \sum_{n\geq 1} \int_{\omega\in\Delta_j^n} f(x^n, a^n | H = j) d(x^n, a^n) \qquad (35) \\
&\leq \frac{1}{L}. \qquad (36)
\end{aligned}$$

The equality in (33) follows from (32) and from Proposition (4). The inequality in (34) follows because the maximum likelihood function satisfies $\hat{f}(x^n, a^n | H = i) \geq f(x^n, a^n | \Psi = (i, R_1, R_2))$ for all Ψ such that $H = i$. The inequality in (35) follows because $\omega \in \Delta_j^n$ implies $Z_{ji} \geq \log((K-1)L)$, which in turn implies that the term within parenthesis is upper bounded by $1/((K-1)L)$, a consequence of (28). Inequality in (36) follows because the inner summation in (35) is a sum of probabilities of disjoint events, and hence is upper bounded by one. ■

Observe that we chose the modified GLR instead of GLR precisely because we want to recognise the inner summation in (35) as a probability of an event and upper bounded by 1. If we use the GLR, the integrand would have been a maximum likelihood which after summation and integration may not even be finite.

We now move on to show that π_M is asymptotically optimal. We first assert that the process $(Z_i(n))_{n\geq 1}$ has an asymptotic drift equal to $D^*(i, R_1, R_2)$.

Proposition 6: Consider the non-stopping policy $\tilde{\pi}_M$. Let $\Psi = (i, R_1, R_2)$ be the true configuration. Then,

$$\lim_{n\to\infty} \frac{Z_i(n)}{n} = D^*(i, R_1, R_2) \text{ almost surely.} \tag{37}$$

Proof: See Appendix B-B. ■

With the above ingredients in place, we now assert that our proposed policy $\pi_M(L)$ has a conditional expected stopping time upper bounded by the desired quantity.

Proposition 7: Consider the policy $\pi_M(L)$. Let $\Psi = (i, R_1, R_2)$ be the true configuration. Then

$$\limsup_{L\to\infty} \frac{\tau(\pi_M(L))}{\log(L)} \leq \frac{1}{D^*(i, R_1, R_2)} \text{ almost surely,} \tag{38}$$

and further,

$$\limsup_{L\to\infty} \frac{E[\tau(\pi_M(L)) | \Psi]}{\log(L)} \leq \frac{1}{D^*(i, R_1, R_2)}. \tag{39}$$

Proof: See Appendix B-C. ■

We now state the main theorem that combines the lower bound in Proposition 1 and the upper bound in Proposition 7 to show that our proposed policy $\pi_M(L)$ is asymptotically optimal.

Theorem 8: Consider K homogeneous Poisson point processes with configuration $\Psi = (i, R_1, R_2)$. Let $(\alpha^{(n)})_{n\geq 1}$ be a sequence of vectors, where $\alpha^{(n)}$ is the nth tolerance vector, such that $\lim_{n\to\infty} \|\alpha^{(n)}\| = 0$ and

$$\limsup_{n\to\infty} \frac{\|\alpha^{(n)}\|}{\min_k \alpha_k^{(n)}} \leq B \text{ for some } B. \tag{40}$$

Then, for each n, the policy $\pi_M(L_n)$ with $L_n = 1/\min_k \alpha_k^{(n)}$ belongs to $\Pi(\alpha^{(n)})$. Furthermore,

$$\liminf_{n\to\infty} \inf_{\pi\in\Pi(\alpha^{(n)})} \frac{E[\tau(\pi)) | \Psi]}{\log(L_n)} = \limsup_{n\to\infty} \frac{E[\tau(\pi_M(L_n))) | \Psi]}{\log(L_n)} \tag{41}$$

$$= \frac{1}{D^*(i, R_1, R_2)}. \tag{42}$$

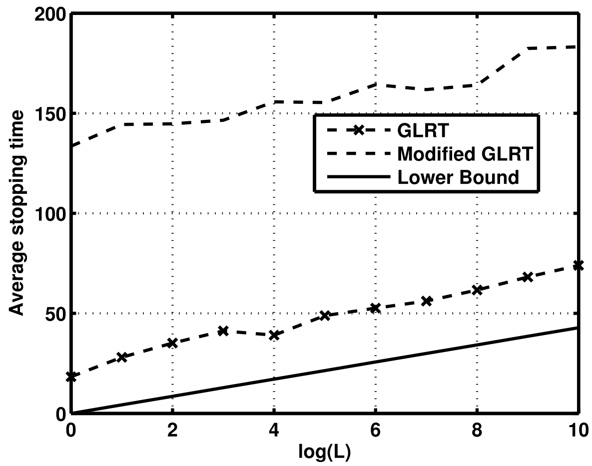

Fig. 2. Performance of $\pi_M(L)$ against GLRT. $R_1 = 10, R_2 = 15, K = 8. D^* = 0.2355$.

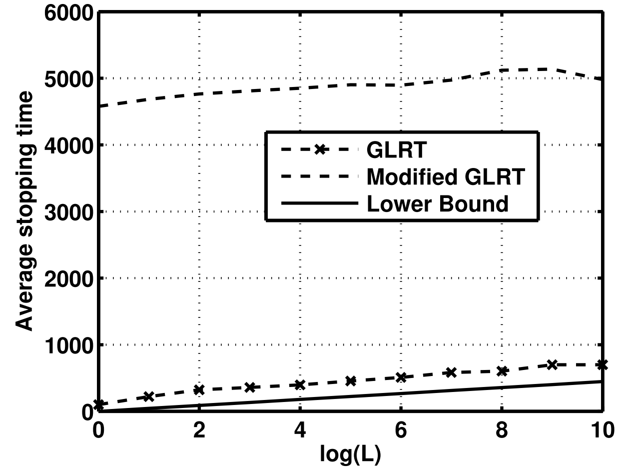

Fig. 3. Performance of $\pi_M(L)$ against GLRT. $R_1 = 45, R_2 = 48, K = 8. D^* = 0.02238$.

Proof: The fact that $\pi_M(L_n) \in \Pi(\alpha^{(n)})$ follows from Proposition 5. We then have the following inequalities:

$$\frac{1}{D^*(i, R_1, R_2)} \leq \liminf_{n\to\infty} \inf_{\pi\in\Pi(\alpha^{(n)})} \frac{E\left[\tau(\pi)\right)|\Psi]}{-\log(\|\alpha^{(n)}\|)} \tag{43}$$

$$= \liminf_{n\to\infty} \inf_{\pi\in\Pi(\alpha^{(n)})} \frac{E\left[\tau(\pi)\right)|\Psi]}{\log(L_n)} \tag{44}$$

$$\leq \limsup_{n\to\infty} \frac{E\left[\tau(\pi_M(L_n))\right)|\Psi]}{\log(L_n)} \tag{45}$$

$$\leq \frac{1}{D^*(i, R_1, R_2)}. \tag{46}$$

Inequality (43) follows from Proposition 1. Equality (44) follows from the choice of L_n and from assumption (40). Inequality (45) follows because $\pi_M(L_n)$ belongs to $\Pi(\alpha^{(n)})$. Inequality (46) follows from Proposition 7. ■

V. Numerical Simulations

In this section we study the performance of our modified GLR based algorithm $\pi_M(L)$ via numerical simulations. Fig. 2 and Fig. 3 show the empirical average stopping time of our modified GLR based $\pi_M(L)$ (modified GLRT) and the standard GLRT (with the same threshold as $\pi_M(L)$), averaged across 100 independent runs of the algorithms, and plotted against $\log(L)$. Note that $\pi_M(L)$ ensures that the probability of error is upperbounded by $1/L$, though such a guarantee is not available for GLRT. We also plot the lower bound on the expected stopping time as obtained in Proposition 12. We observe that the expected stopping time for $\pi_M(L)$ is greater than the lower bound by a constant additive factor, thus validating the asymptotic optimality of $\pi_M(L)$. (Asymptotic optimality is a statement about the slopes.)

We now make some interesting observations for which we do not, as yet, have a sound theoretical basis. For large sample size, the modified GLRT appears to converge to the true likelihood ratio minus the log of the density of the fictitious prior. For our specific choice of the gamma prior, the modified GLRT then lags behind the likelihood ratio by an additive term $(R_1 + R_2)$. This leads to an average increase in delay of $(R_1 + R_2)/D^*$ for our modified GLRT. Further, since our threshold increases with the number of locations K by $\log(K - 1)$, this appears to lead to an extra delay of $\log(K-1)/D^*$. For the specific example given in Fig. 2 below, where $R_1 = 10$ and $R_2 = 15$, the above mentioned delays add up to $(R_1 + R_2)/D^* + \log(K - 1)/D^* = 25/0.2355 + \log(7)/0.2355 \approx 114$. Similarly for Fig. 3, where $R_1 = 45$ and $R_2 = 48$, the above calculation suggests an average increase in stopping time of $(R_1+R_2)/D^*+\log(K-1)/D^* = 93/0.02238 + \log(7)/0.02238 \approx 4242$. Both the values are close to the difference in the empirical average stopping time of the modified GLRT $\pi_M(L)$ and the lower bound, as seen in Fig. 2 and Fig. 3.

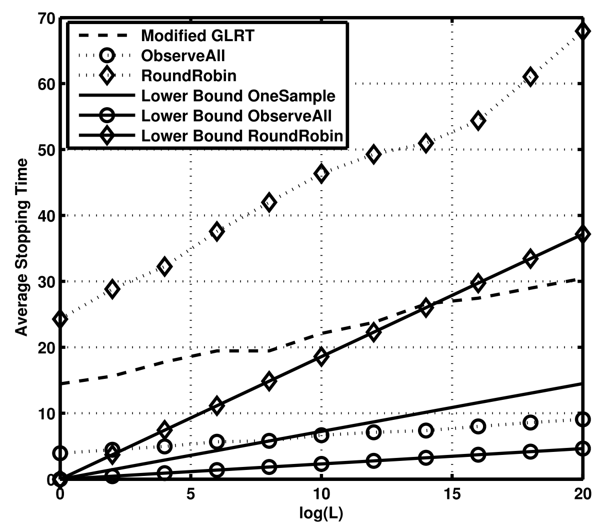

Fig. 4. Performance of $\pi_M(L)$ compared with (a) the setting when all processes are observed, and (b) a simple round-robin search. $R_1 = 2, R_2 = 10, K = 8. D^* = 1.3791, \overline{D} = 4.3055, \overline{D}_{RR} = 0.5382$.

In Fig. 4 we compare the performance of $\pi_M(L)$ against the following scenarios: (a) the decision maker can observe all the processes at every time slot, and (b) the decision maker samples the locations in a round-robin sequence. In the former case, the only control is on when to stop. In the latter case the decision maker does not optimally choose the location to sample. The stopping criterion is the same for the three sampling schemes – observe-all, round-robin, and $\pi_M(L)$. We also plot the lower bounds for the three policies.

Let us first identify the lower bound for observe-all. This follows easily from Proposition 2. Indeed, for the observe-all case, following the optimisation steps in the proof of Proposition 2, the optimum drift can be calculated to be

$$\begin{aligned}\overline{D}(i, R_1, R_2) &:= \min_{R_1', R_2', j: j \neq i} \Big[D(R_1 || R_2') \\ &\qquad + D(R_2 || R_1') + (K-2) D(R_2 || R_2') \Big] \\ &= D(R_1 || R_{ave}) + (K-2) D(R_2 || R_{ave}),\end{aligned}$$

where

$$R_{ave} = \frac{R_1 + (K-2)R_2}{(K-1)}.$$

The drift for the round-robin case is simply $\overline{D}(i, R_1, R_2)/K$. This too follows from Proposition 2 by observing that the sampling distribution is uniform on the K locations.

As expected, the slopes of the three policies match with their lower bounds corroborating our theory. Further, $\pi_M(L)$ performs better than round-robin, justifying the need for choosing the location to sample carefully, but worse than observe-all since $\pi_M(L)$ gets very limited information per sample.

VI. Application to Visual Search

In this section we apply our results to the visual search experiments of Sripati and Olson [4]. These experiments were motivated by Sripati's and Olson's desire to quantify the notion of the so-called "perceptual distance", a problem set in the context of their larger goal of understanding how objects are represented in our brains. Two images (say two faces) may be very close to each other in their pixel representation. Yet they may be easily recognised by a human subject as two distinct faces. The objects then have small "pixel distance", but are far apart in "perceptual space" and are separated by a large perceptual distance.

How do we measure perceptual distances between objects? Neuroscientists have proposed an ingenious method to quantify this distance between pairs of images, say I_i and I_j. For concreteness, consider two images I_1 and I_2. A subject is shown a picture with K images of which $K-1$ are identically I_1 (distracters) and the oddball image is I_2 (target). The subject does not know the location of the oddball image. The subject is then asked to identify, as quickly as possible and without guessing, the location of the oddball image. The time taken to identify the location of the oddball, from the onset of the K images, is measured across several trials and subjects, and then averaged. The inverse of this average time is taken to be an estimate of the perceptual distance separating I_2 from I_1. Often, the distance estimate is symmetrised by repeating trials with the role of I_1 and I_2 reversed, and by averaging the time to decision across all trials.

A decision-theoretic perspective to this problem was provided by Vaidhiyan *et al.* [5]. Assume that the images I_1 and I_2 are known, but which of the two is the oddball target is not specified. The location of the oddball is unknown. In the visual search model of Vaidhiyan *et al.* [5], it was assumed that subjects can focus at any location of their choice (a control). Given a chosen focus location, the image at that location elicited, in a population of neurons, a spiking pattern according to a multi-dimensional Poisson point process. Also, given the firing rates, the processes were assumed to be independent of each other. The goal of the human subject was then to identify the oddball as quickly as possible. The human subject would make this decision through a series of decisions on whether sufficient confidence has already been gained to stop and make a decision on the oddball location, or if not, where to focus next.

It is well-established (see Sripati and Olson [4] and references therein) that neurons that respond to gross-level object attributes and features, in macaque monkey brains, reside in the region of the brain called the inferotemporal cortex. Sripati and Olson [4] made extensive measurements in this region on how neurons responded to the images, say I_i and I_j. They then proposed the L_1 distance between the vector of observed average neuronal firing rates (in macaque monkeys), a neuronal dissimilarity index, as an alternative measure of perceptual distance between I_i and I_j. They then demonstrated that this neuronal dissimilarity index, obtained from measurements on macaque monkeys, correlated exceedingly well with the inverse of the symmetrised and averaged search times that human subjects took to tell apart I_i and I_j in the visual search experiments, across several pairs of images I_i and I_j.

The goal of Vaidhiyan *et al.* [5] was to argue that the L_1 distance between the average firing rates of neurons does not have a sound decision theoretic basis as a notion of perceptual distance. Instead, following an approach similar to this paper, they proposed a relative entropy based dissimilarity index, denoted $\tilde{D}_{ij}$. This was the inverse of the constant to which $E[\tau(i, j)/\log(L)]$, conditioned on I_i being the oddball, converges as $L \to \infty$, where $1/L$ is the constraint on the probability of false detection and $\tau(i, j)$ is the stopping time of the optimal policy to locate oddball I_i among distracters I_j. It was assumed that the firing rate of the neurons, under the two images, were known. The $\tilde{D}_{ij}$ dissimilarity index had the added feature that it was asymmetric, a property that the L_1 dissimilarity index does not have. Object search is known to be asymmetric. For example, identifying a 'Q' in a sea of 'O's is easier that identifying an 'O' in a sea of 'Q's. The choice of $\tilde{D}_{ij}$ as a neuronal dissimilarity index naturally brings out the asymmetry, and has a decision theoretic basis.

We discussed two dissimilarity indices: for a pair of images I_i and I_j, the dissimilarity indices are $\tilde{D}_{ij}$ and the L_1 distance between the associated neuronal average firing rates. There are also several other natural ones, for example, the relative entropy and the Chernoff entropy between the two multi-dimensional Poisson point processes associated with average firing rate vectors. Which of these best explains the observations in the experiments with human subjects? This too was investigated in [5]. An ideal neuronal dissimilarity index

TABLE I
CORRELATION WITH DIFFERENT INFORMATION DISSIMILARITY INDICES

Neuronal information dissimilarity indices	Correlation - Average decision delay vs. (neuronal information dissimilarity index)$^{-1}$	p-value
$\tilde{D}$ [5]	0.89	4.3×10^{-09}
KL	0.90	3.1×10^{-09}
Chernoff	0.88	2.1×10^{-08}
L^1	0.88	1.1×10^{-08}
D^* (this paper)	0.89	8.7×10^{-09}

TABLE II
EQUALITY OF MEANS TEST USING VARIOUS TEST STATISTICS

Neuronal information dissimilarity indices	ANOVA statistic	ANOVA p-values	log(AM/GM)
$\tilde{D}$ [5]	06.30	9.35×10^{-19}	0.0200
KL	06.68	2.88×10^{-20}	0.0211
Chernoff	06.74	1.61×10^{-20}	0.0252
L^1	24.00	3.42×10^{-87}	0.0678
D^* (this paper)	06.34	6.93×10^{-19}	0.0233

between a pair of images I_i and I_j, say $\mathsf{diff}(i, j)$, would satisfy $E[\tau(i, j)]\mathsf{diff}(i, j) = \text{constant}$, for any image pair (i, j). The constant should be independent of the image pair. Vaidhiyan *et al.* [5] proposed tests of equality of means to measure the *dispersion* of $E[\tau(i, j)]\mathsf{diff}(i, j)$, across image pairs, about a common mean (the constant). A natural statistic to test the dispersion of group means about a common mean is the ratio of arithmetic mean (AM) to geometric mean (GM) of the group means. It turns out that (AM/GM) is the statistic for a GLRT based equality of means test for gamma distributed random variables under a fixed shape parameter assumption (see [5]). The test for equality of means across groups for Gaussian random variables is the well-known one-way ANOVA test. ANOVA is also widely used for non-Gaussian random variables also because of its robustness. On these dispersion measures, the relative entropy based dissimilarity index $\tilde{D}_{ij}$ outperformed the L_1 distance and other natural metrics. These results from [5] are highlighted in Tables I and II, and will be discussed shortly.

Often, in experiments with human subjects, the oddball and distracter images are not disclosed up front. They are unknown and have to be learnt along the way by the subject. This latter problem then falls within the framework of this paper, and was our main motivation for studying the problem formulated in this paper. This decision theoretic problem has a similar limiting $E[\tau(i, j)/\log(L)]$, and the corresponding neuronal dissimilarity index is the $D^*(i, R_1, R_2)$ of this paper (Theorem 8), where i is the oddball image, R_1 is the firing rate of a single neuron responding to image i, and R_2 is the firing rate of the distracter.

The model described in this paper was the simplified one-dimensional Poisson point process. This is equivalent to making decisions based on observations from a single neuron that responds differently to two images. However, all our results on the one-dimensional Poisson point process (one neuron case) extend naturally to multi-dimensional Poisson point processes (multiple neurons firing independently). Hence, the extension of $D^*(i, R_1, R_2)$ to $D^*(i, \underline{R}_1, \underline{R}_2)$ for vectors of rates $\underline{R}_1, \underline{R}_2$ is straightforward – formula (11) continues to hold with R_1, R_2, $\tilde{R}$ replaced by vectors $\underline{R}_1$, $\underline{R}_2$, $\underline{\tilde{R}}$ respectively, and $D(R_i, \tilde{R})$ replaced by $D(\underline{R}_i, \underline{\tilde{R}}) = \sum_d D(\underline{R}_i(d) \| \underline{\tilde{R}}(d))$, where the summation is over neurons indexed by d.

Table I shows the correlation values for different information dissimilarity indices (from Vaidhiyan et al. [5]). We call these neuronal dissimilarity indices because these information dissimilarity indices are computed based on firing rates of neurons gathered in the measurement experiments of Sripati and Olson [4]. We see that the inverse of the proposed D^*, as are the inverses of other indices, is strongly correlated with the average decision delay.

Table II shows the statistics related to ANOVA and (AM/GM) tests. As with other indices, D^* fails the equality of means tests (indicated by the p-values for ANOVA in the second column; similarly for log(AM/GM) tests). When the statistics are used to rank order the indices, from the ANOVA statistic (smaller the better), we see that D^* is ranked below $\tilde{D}$, but above the other indices. From the log(AM/GM) statistics we see that D^* is ranked below $\tilde{D}$ and the KL indices, but above Chernoff and L_1.

The slight degradation in performance of D^* with respect to $\tilde{D}$ may be attributed to the particular experimental setup of Sripati and Olson [4]. The search tasks associated with a given image pair belonged to the same block of trials, and hence were contiguous. This may have possibly cued the human subject about the upcoming image pair, and the subject may have already 'learnt' the firing rates. This violates our assumption on the lack of prior knowledge of the image pairs to the decision maker at the beginning of each experiment. Of course, a more thorough experimentation with a wide variety of image pairs and few repetitions is required for a good evaluation of the performance of D^*. But the fact that D^* is close to the top is encouraging. We believe that this paper may provide the necessary theoretical basis to carry out similar analysis of other experimental data.

Our paper focussed on the regime of large samples although our motivation came from a neuroscience application where it is generally difficult to have a large number of samples. We feel that the regime of large samples may yet be applicable. Typical error rates observed by Sripati and Olson [4] in their experiments are $1/L = 0.04$ to 0.05. This yields $L = 20$ to 25 or $\log(L) \approx 3$ to 3.2. Observe from Fig. 2 and Fig. 3 that asymptotic linearity is already apparent at these error rates.

VII. Conclusion

We studied the problem of detecting an odd Poisson point process having a rate different from the common rate of others. We developed a lower bound on the conditional expected stopping time for any policy that satisfies the given constraint on the probability of false detection. We proposed a modified GLRT based algorithm, that we called π_M and showed that it satisfies the given constraint on the probability of false detection, and that it is asymptotically optimal with respect to the conditional expected stopping time. The proposed algorithm employs a simple threshold criterion for stopping. Interestingly, we also showed that, independent of the configuration, the sampling probability for each process is strictly above a positive constant.

We applied our results to the visual search experiments of Sripati and Olson [4]. We proposed D^* as a candidate neuronal dissimilarity index. D^* correlated strongly with the behavioral data. The performance of D^* was only marginally inferior to the neuronal dissimilarity index proposed by Vaidhiyan *et al.* [5], and this is quite encouraging.

This work was restricted to Poisson processes. Extension to other class of distributions, especially exponential family is under consideration. Extension to general class of distributions will be an interesting extension.

Appendix A
Proof of Proposition 3

Let us rewrite (11) as

$$\lambda^*(k,\theta_1,\theta_2)(k) = \arg\max_{0\le\lambda\le 1}\left[\lambda D(\theta_1\|\tilde{\theta}) + (1-\lambda)\frac{(K-2)}{(K-1)}D(\theta_2\|\tilde{\theta})\right],$$

where $\tilde{\theta}$, as in (12), is given by

$$\tilde{\theta} = \frac{\lambda\theta_1 + (1-\lambda)\frac{(K-2)}{(K-1)}\theta_2}{\lambda + (1-\lambda)\frac{(K-2)}{(K-1)}}. \quad (47)$$

We have abused notation and have used λ to denote the scalar $\lambda(k)$ of (11). We first show that the second derivative of the objective function in the above optimisation is negative for all λ to establish concavity. Define the objective function as

$$f(\lambda) := \lambda D(\theta_1\|\tilde{\theta}) + (1-\lambda)\frac{(K-2)}{(K-1)}D(\theta_2\|\tilde{\theta}),$$

where $\tilde{\theta}$, a function of λ, is as in (47). We then have

$$\frac{df}{d\lambda} = D(\theta_1\|\tilde{\theta}) - \frac{(K-2)}{(K-1)}D(\theta_2\|\tilde{\theta}) \quad (48)$$
$$+\left(\lambda D'(\theta_1\|\tilde{\theta}) + (1-\lambda)\frac{(K-2)}{(K-1)}D'(\theta_2\|\tilde{\theta})\right)\frac{d\tilde{\theta}}{d\lambda} \quad (49)$$
$$= D(\theta_1\|\tilde{\theta}) - \frac{(K-2)}{(K-1)}D(\theta_2\|\tilde{\theta}), \quad (50)$$

where, we recall, $D'(x\|y)$ is the derivative of $D(x\|y)$ with respect to the second argument y, which turns out to be $1-x/y$. Equality (50) follows from (16), which ensures that the term within the parenthesis is identically zero. Differentiating once again,

$$\begin{aligned}\frac{d^2 f}{d\lambda^2} &= \left(D'(\theta_1\|\tilde{\theta}) - \frac{(K-2)}{(K-1)}D'(\theta_2\|\tilde{\theta})\right)\frac{d\tilde{\theta}}{d\lambda}\\ &= \left(\left(1-\frac{\theta_1}{\tilde{\theta}}\right) - \frac{(K-2)}{(K-1)}\left(1-\frac{\theta_2}{\tilde{\theta}}\right)\right)\frac{d\tilde{\theta}}{d\lambda}\\ &= -\frac{\tilde{\theta}}{\lambda+(1-\lambda)\frac{(K-2)}{(K-1)}}\\ &\quad\times\left(\left(1-\frac{\theta_1}{\tilde{\theta}}\right) - \frac{(K-2)}{(K-1)}\left(1-\frac{\theta_2}{\tilde{\theta}}\right)\right)^2\\ &\le 0, \end{aligned} \quad (51)$$

where we have used the fact that

$$\frac{d\tilde{\theta}}{d\lambda} = -\frac{\tilde{\theta}}{\lambda+(1-\lambda)\frac{(K-2)}{(K-1)}} \times\left(\left(1-\frac{\theta_1}{\tilde{\theta}}\right) - \frac{(K-2)}{(K-1)}\left(1-\frac{\theta_2}{\tilde{\theta}}\right)\right).$$

Since $f(\lambda)$ is concave in λ, and since $f(0) = f(1) = 0$, and $f'(0) > 0$ and $f'(1) < 0$, the maximiser λ^* satisfies

$$D(\theta_1\|\tilde{\theta}) - \frac{(K-2)}{(K-1)}D(\theta_2\|\tilde{\theta}) = 0. \quad (52)$$

We do not know of a closed form expression for λ^* from (52).

Let $\hat{\lambda}$ denote a parametrisation of λ of the form

$$\hat{\lambda} := \frac{\lambda}{\lambda + (1-\lambda)\frac{(K-2)}{(K-1)}}, \quad (53)$$

so that $\tilde{\theta} = \hat{\lambda}\theta_1 + (1-\hat{\lambda})\theta_2$. We recognise that $\hat{\lambda}$ is increasing in λ. Let $\hat{\lambda}^*$ denote the re-parametrisation for λ^* according to (53). Hence, to show that λ^* is bounded away from 0 and 1, it suffices to show that $\hat{\lambda}^*$ is bounded away from 0 and 1. Let us first consider the case when $\theta_1 < \theta_2$. The case when $\theta_1 > \theta_2$ has similar arguments. Let us consider a new parametrisation of (52). Let υ denote

$$\upsilon = \frac{\theta_2}{\theta_2 - \theta_1}, \quad (54)$$

so that

$$\upsilon - 1 = \frac{\theta_1}{\theta_2-\theta_1}, \quad (55)$$

and

$$\frac{\tilde{\theta}}{\theta_2-\theta_1} = \frac{\hat{\lambda}\theta_1 + (1-\hat{\lambda})\theta_2}{\theta_2-\theta_1} \quad (56)$$
$$= \upsilon - \hat{\lambda}. \quad (57)$$

The left-hand side of (52) can now be written in terms of υ and $\hat{\lambda}$ as

$$D(\upsilon-1\|\upsilon-\hat{\lambda}) - \frac{(K-2)}{(K-1)}D(\upsilon\|\upsilon-\hat{\lambda}). \quad (58)$$

Let $\hat{\lambda}_r(\upsilon)$ denote the solution to

$$D(\upsilon-1\|\upsilon-\hat{\lambda}) - rD(\upsilon\|\upsilon-\hat{\lambda}) = 0. \quad (59)$$

Fig. 5 gives a geometric interpretation of $\hat{\lambda}^*$. Note that $\hat{\lambda}^*(\upsilon) = \hat{\lambda}_r(\upsilon)$ for $r = (K-2)/(K-1)$. For each $\upsilon \ge 1$, we also have

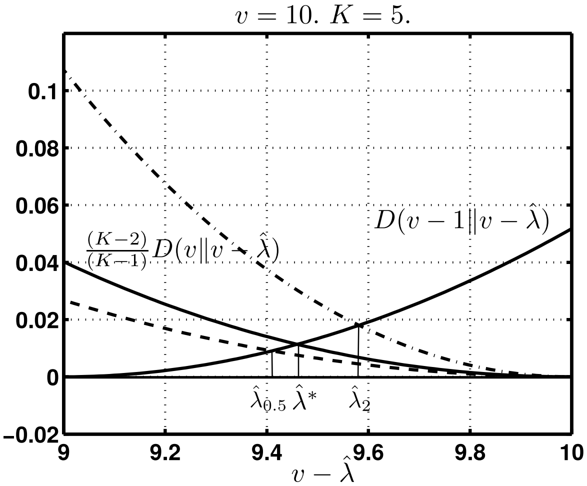

Fig. 5. Geometric interpretation of $\hat{\lambda}^*$.

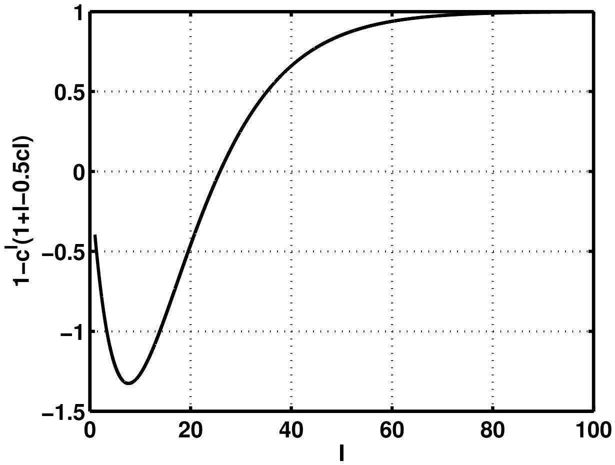

Fig. 6. Variation of $(1 - c^l(1 + l - 0.5cl))$ with l.

that $\hat{\lambda}_r(\upsilon)$ decreases with r. Furthermore, $0.5 \le (K-2)/(K-1) \le 2$. We then have $\hat{\lambda}_2(\upsilon) < \hat{\lambda}^*(\upsilon) < \hat{\lambda}_{0.5}(\upsilon)$. Hence, to show that $\hat{\lambda}^*(\upsilon)$ is bounded away from 0 and 1 for all υ, it suffices to show that $\sup_{\upsilon \ge 1} \hat{\lambda}_{0.5}(\upsilon) < 1$, and that $\inf_{\upsilon \ge 1} \hat{\lambda}_2(\upsilon) > 0$.

We now obtain a Taylor's series based alternate expression for $D(\upsilon - a \| \upsilon - b)$ when $\upsilon \ge 1$ and $0 \le a, b \le 1$. The alternate expression replaces the log terms in (58) with infinite sums and enables easier bounding of (58).

Lemma 9: Let $\upsilon \ge 1$. Let $0 \le a, b \le 1$. Let $D(x \| y) = x \log(x/y) - y + x$ denote the relative entropy between two Poisson random variables with means x and y. Then,

$$D(\upsilon - a \| \upsilon - b) = (\upsilon - a) \log\left(\frac{\upsilon - a}{\upsilon - b}\right) - (\upsilon - a) + (\upsilon - b)$$
$$= \sum_{l \ge 1} \frac{1}{\upsilon^l l(l+1)} \left(a^{l+1} - b^l(a + (a-b)l)\right).$$

Proof: Case 1: Let $\upsilon > 1$. Let $0 \le a, b \le 1$.

Using the Taylor's series expansion for $-\log(1-x) = \sum_{l \ge 1} \frac{x^l}{l}$, when $|x| < 1$, we get

$$\begin{aligned}
&D(\upsilon - a \| \upsilon - b) \\
&= (\upsilon - a) \log\left(\frac{1 - a/\upsilon}{1 - b/\upsilon}\right) - (\upsilon - a) + (\upsilon - b) \\
&= -(\upsilon - a) \sum_{l \ge 1} \frac{a^l}{\upsilon^l l} + (\upsilon - a) \sum_{l \ge 1} \frac{b^l}{\upsilon^l l} + (a - b) \\
&= (-a + b) + \sum_{l \ge 2} \frac{1}{\upsilon^{l-1} l} \left(b^l - a^l\right) \\
&\quad + \sum_{l \ge 1} \frac{1}{\upsilon^l l} \left(a^{l+1} - ab^l\right) + (a - b) \\
&= \sum_{l \ge 1} \frac{1}{\upsilon^l (l+1)} \left(b^{l+1} - a^{l+1}\right) + \sum_{l \ge 1} \frac{1}{\upsilon^l l} \left(a^{l+1} - ab^l\right) \\
&= \sum_{l \ge 1} \frac{1}{\upsilon^l l(l+1)} \left(a^{l+1} - b^l(a + (a-b)l)\right).
\end{aligned}$$

Case 2: Let $\upsilon = 1$. Let $0 < a, b < 1$. The same arguments as above holds.

Case 3: Let $\upsilon = 1$. Let $a = 1$, $b < 1$. Then,

$$\begin{aligned}
&\sum_{l \ge 1} \frac{1}{l(l+1)} \left(1 - b^l(1 + (1-b)l)\right) \\
&\quad = \sum_{l \ge 1} \left[\frac{1}{l} - \frac{1}{(l+1)} - \frac{b^l}{l} + \frac{b^{l+1}}{l+1}\right] \\
&\quad = (1 - b) \\
&\quad = D(0 \| 1 - b).
\end{aligned}$$

Case 4: Let $\upsilon = 1$. Let $a < 1$, $b = 1$. Then, both $D(\upsilon - a \| \upsilon - b)$ and the infinite sum are infinity.

Case 5: Let $\upsilon = 1$. Let $a = 1, b = 1$. Then both $D(\upsilon - a \| \upsilon - b)$ and the infinite sum are zero. ■

We now show that $\hat{\lambda}_{0.5}(\upsilon) < 0.9$ for all $\upsilon \ge 1$. For this, it suffices to show that for $c = 0.9$, $D(\upsilon - 1 \| \upsilon - c) - 0.5\, D(\upsilon \| \upsilon - c) < 0$ for all $\upsilon \ge 1$.

$$\begin{aligned}
&D(\upsilon - 1 \| \upsilon - c) - 0.5 D(\upsilon \| \upsilon - c) \\
&\quad = \sum_{l \ge 1} \frac{1}{\upsilon^l l(l+1)} \left(1 - c^l(1 + (1-c)l) - 0.5c^{l+1} l\right) \\
&\quad = \sum_{l \ge 1} \frac{1}{\upsilon^l l(l+1)} \left(1 - c^l(l+1) + c^{l+1} l - 0.5c^{l+1} l\right) \\
&\quad = \sum_{l \ge 1} \frac{1}{\upsilon^l l(l+1)} \left(1 - c^l(l + 1 - 0.5cl)\right).
\end{aligned}$$

Let us first consider the case when $\upsilon = 1$ and $c = 0.9$. We then have

$$\begin{aligned}
&D(\upsilon - 1 \| \upsilon - c) - 0.5 D(\upsilon \| \upsilon - c) \\
&\quad = D(0 \| 0.1) - 0.5 D(1 \| 0.1) \\
&\quad = 0.1 - 0.5(\log(10) - 0.9) \\
&\quad < 0.
\end{aligned} \tag{60}$$

Thus, $\hat{\lambda}_{0.5}(1) < 0.9$. For $\upsilon > 1$ and $c = 0.9$, we observe that $(1 - c^l(1 + l - 0.5cl))$ is initially negative and then becomes positive in l (See Fig. 6). Thus, there exists $M > 1$ such that

$$(1 - c^l(1 + l - 0.5cl)) \begin{cases} \le 0 & \forall\, l < M \\ \ge 0 & \forall\, l \ge M. \end{cases} \tag{61}$$

Then, for $c = 0.9$, we have

$$\begin{aligned}
&D(\upsilon - 1\|\upsilon - c) - 0.5D(\upsilon\|\upsilon - c)\\
&= \sum_{l\geq 1} \frac{1}{\upsilon^l l(l+1)}\left(1 - c^l((l+1) - 0.5cl\right)\\
&\leq \sum_{1\leq l<M} \frac{1}{\upsilon^M l(l+1)}\left(1 - c^l((l+1) - 0.5cl\right)\\
&\quad + \sum_{l\geq M} \frac{1}{\upsilon^M l(l+1)}\left(1 - c^l((l+1) - 0.5cl\right) \qquad (62)\\
&= \frac{1}{\upsilon^M}\sum_{l\geq 1} \frac{1}{l(l+1)}\left(1 - c^l((l+1) - 0.5cl\right)\\
&= \frac{1}{\upsilon^M}\left(D(0\|1-c) - 0.5D(1\|1-c)\right)\\
&= \frac{1}{\upsilon^M}\left(D(0\|0.1) - 0.5D(1\|0.1)\right)\\
&< 0. \qquad (63)
\end{aligned}$$

Inequality (62) is obtained by upperbounding 1) the initial negative terms, till $l < M$, by replacing υ^l by a larger υ^M, and 2) the later non-negative terms, for $l \geq M$, by replacing υ^l by a smaller υ^M. Inequality (63) follows from (60). Thus, we have shown that $\hat{\lambda}_{0.5}(\upsilon) < 0.9$ for all $\upsilon \geq 1$.

We now show the second part of the proof, i.e., $\hat{\lambda}_2(\upsilon) > 0.1 > 0$. For this, it suffices to show that for $c = 0.1$, $D(\upsilon-1\|\upsilon-c)-2D(\upsilon\|\upsilon-c) > 0$ for all $\upsilon \geq 1$. For $c = 0.1$, we have

$$\begin{aligned}
&D(\upsilon - 1\|\upsilon - c) - 2D(\upsilon\|\upsilon - c)\\
&= \sum_{l\geq 1} \frac{1}{\upsilon^l l(l+1)}\left(1 - c^l(1 + (1-c)l) - 2c^{l+1}l\right)\\
&= \sum_{l\geq 1} \frac{1}{\upsilon^l l(l+1)}\left(1 - c^l(1 + l + cl)\right)\\
&= \sum_{l\geq 1} \frac{1}{\upsilon^l l(l+1)}\left(1 - (0.1)^l(1 + l + (0.1)l)\right) \qquad (64)\\
&> 0, \qquad (65)
\end{aligned}$$

where (65) follows as each term inside the summation in (64) is positive. Thus, when $\theta_1 < \theta_2$ and for all $\upsilon \geq 1$, we have shown that

$$0.1 \leq \hat{\lambda}_2(\upsilon) \leq \hat{\lambda}^*(\upsilon) \leq \hat{\lambda}_{0.5}(\upsilon) < 0.9.$$

We now consider the case when $\theta_1 > \theta_2$. Let

$$\upsilon' = \frac{\theta_1}{\theta_1 - \theta_2},$$

so that

$$\upsilon' - 1 = \frac{\theta_2}{\theta_1 - \theta_2},$$

and

$$\begin{aligned}
\frac{\tilde{\theta}}{\theta_1 - \theta_2} &= \frac{\hat{\lambda}\theta_1 + (1-\hat{\lambda})\theta_2}{\theta_1 - \theta_2}\\
&= \upsilon' - 1 + \hat{\lambda}.
\end{aligned}$$

Equation (52) can now be written in terms of υ' and $\hat{\lambda}$ as

$$D(\upsilon'\|\upsilon' - 1 + \hat{\lambda}) - \frac{(K-2)}{(K-1)}D(\upsilon' - 1\|\upsilon' - 1 + \hat{\lambda}) = 0. \qquad (66)$$

Let $\hat{\lambda}^*(\upsilon')$ be the solution to (66). Recognise that (66) has the same form as in the previous case for $\theta_1 < \theta_2$, with only the multiplicative constant being different. From arguments similar to the ones used in the previous case of $\theta_1 < \theta_2$, we can show that

$$0.1 < (1 - \hat{\lambda}^*(\upsilon')) < 0.9,$$

or equivalently, $0.1 < \hat{\lambda}^*(\upsilon') < 0.9$.

Thus, we have shown that $\hat{\lambda}^*$ is bounded away from 0.1 and 0.9 for all θ_1 and θ_2. ■

Appendix B

We stated the main properties of the proposed policy π_M in Section IV. We prove them in this Appendix.

The organisation of this Appendix is as follows. This also outlines the structure of the proof of achievability given in Proposition 7. First, in Appendix B-A, we prove Proposition 4 which is that the policy $\pi_M(L)$ stops with probability 1. Next, in Appendix B-B, we prove Proposition 6 which is that the log-likelihood ratio under the policy $\tilde{\pi}_M$ grows at the correct rate, i.e., the normalised log-likelihood ratio converges to $D^*(i, R_1, R_2)$ almost surely. Finally, in Appendix B-C, we prove Proposition 7 which establishes the desired upper bound on $\tau(\pi_M(L))/\log L$ almost surely and in expectation. The beginning of each subsection outlines the plan for that subsection.

A. Proof of Proposition 4 and Associated Ingredients

Before we prove Proposition 4, which is that policy $\pi_M(L)$ stops with probability 1, we develop some convergence results for $\pi_M(L)$.

In Proposition 10, we show that under the non-stopping policy $\tilde{\pi}_M$, the empirical rate associated with a process converges to the true rate of that process. The results are akin to convergence results for independent random variables, but applied to the dependent random variables in our setting with the dependency being induced by the policy. This is then crucially used in establishing that the log-likelihood ratios of the correct hypothesis against each incorrect hypothesis, $Z_{ij}(n)$ has a strictly positive drift under the non-stopping policy $\tilde{\pi}_M$ (Lemma 11). These are then quickly put together to establish Proposition 4. We then end this subsection with assertions that the maximum likelihood estimate of the odd-ball location and all relevant estimates of the rates of point processes for each location converge to the correct values (Proposition 12). This result will then be used in the next subsection.

We now begin with the convergence of empirical rate convergences.

Proposition 10: Fix $K \geq 3$. Let $\Psi = (i, R_1, R_2)$ be the true configuration. Consider the non stopping policy $\tilde{\pi}_M$. As $n \to \infty$ the following convergences hold almost surely,

$$\begin{aligned}
\frac{Y_j^n}{N_j^n} &\to \begin{cases} R_1 & \text{if } j = i,\\ R_2 & \text{if } j \neq i,\end{cases}\\
\frac{Y^n - Y_i^n}{n - N_i^n} &\to R_2, \qquad (67)
\end{aligned}$$

and

$$\begin{aligned} R'_{min} &\leq \liminf_{n\to\infty} \frac{Y^n - Y_j^n}{n - N_j^n} \\ &\leq \limsup_{n\to\infty} \frac{Y^n - Y_j^n}{n - N_j^n} \\ &\leq R'_{max} \text{ for all } j \neq i, \end{aligned} \tag{68}$$

where

$$R'_{min} = (1 - c_K)\min\{R_1, R_2\} + c_K \max\{R_1, R_2\} \tag{69}$$

and

$$R'_{max} = c_K \min\{R_1, R_2\} + (1 - c_K)\max\{R_1, R_2\}, \tag{70}$$

and c_K is as in Proposition 3.

Proof: Let $\mathcal{F}_{l-1}$ denote the σ-field generated by (X^{l-1}, A^{l-1}). Consider the martingale difference sequence

$$S_i^n = Y_i^n - N_i^n R_1 = \sum_{l=1}^{n} (X_l - R_1) 1_{\{A_l = i\}}.$$

Given the Poisson assumption on X_l, we have $E\left[(X_l - R_1)^2 1_{\{A_l=i\}} | \mathcal{F}_{l-1}\right] < \infty$ for all l. Then, by the convergence result for martingales, see De la Pena [20, Th. 1.2A], for any $\epsilon > 0$, there exists $c_\epsilon > 0$ such that

$$P(S_i^n > n\epsilon) \leq e^{-c_\epsilon n}, \tag{71}$$

which in turn, by the Borel-Cantelli Lemma [21, Sec. 4.2], implies

$$\frac{S_i^n}{n} \to 0 \text{ almost surely.} \tag{72}$$

Similarly arguing, we conclude that convergence result holds for other S_j^n/n, for $j = 1, 2, \ldots, K$. Further, from Proposition 3, we have

$$\liminf_{n\to\infty} \frac{N_i^n}{n} > c_K > 0 \text{ almost surely.} \tag{73}$$

Combining (72) and (73), we have,

$$\frac{S_i^n}{N_i^n} \to 0 \text{ almost surely,}$$

or equivalently,

$$\frac{Y_i^n}{N_i^n} \to R_1 \text{ almost surely.}$$

Similar result hold for other j, with R_1 replaced by R_2, and we have established (67). Furthermore, these results imply that

$$\frac{(Y^n - Y_j^n) - \sum_{k\neq j} N_k^n R_k}{(n - N_j^n)} \to 0 \text{almost surely.} \tag{74}$$

Consequently, we get

$$\frac{(Y^n - Y_i^n)}{(n - N_i^n)} \to R_2 \text{ almost surely.} \tag{75}$$

Fix $j \neq i$, we then have

$$\frac{\sum_{k\neq j} N_k^n R_k}{(n - N_j^n)} = \frac{N_i^n}{n - N_j^n} R_1 + \frac{\sum_{k\neq j,i} N_k^n}{n - N_j^n} R_2.$$

We do not yet have a convergence result for N_k^n/n for any k. Proposition (3) only says that at every slot and for each process, the probability of choosing that process is greater than c_K. Thus, we are not in a position to say, as $n \to \infty$, whether

$$\frac{Y^n - Y_j^n}{n - N_j^n} \to constant.$$

However, from Proposition 3, we get the following bound

$$\begin{aligned} &(1 - c_K)\min\{R_1, R_2\} + c_K \max\{R_1, R_2\} \\ &\quad\leq \liminf_{n\to\infty} \frac{\sum_{k\neq j} N_k^n R_k}{(n - N_j^n)} \\ &\quad\leq \limsup_{n\to\infty} \frac{\sum_{k\neq j} N_k^n R_k}{(n - N_j^n)} \\ &\quad\leq c_K \min\{R_1, R_2\} + (1 - c_K)\max\{R_1, R_2\} \text{ "a.s.".} \end{aligned} \tag{76}$$

Thus, (74) combined with (76) yields (68). ∎

We now state a lemma that asserts that, under the non-stopping policy $\tilde{\pi}_M$, $Z_i(n)$, the test statistic associated with the index of the odd process, drifts off to infinity.

Lemma 11: Fix $K \geq 3$. Let $\Psi = (i, R_1, R_2)$ be the true configuration. Consider the non-stopping policy $\tilde{\pi}_M$. Then for all $j \neq i$, we have

$$\liminf_{n\to\infty} \frac{Z_{ij}(n)}{n} > 0 \text{ almost surely.} \tag{77}$$

Proof: Without loss of generality assume $R_1 < R_2$. Observe that we have $R_1 < R'_{min} < R'_{max} < R_2$. Recall that

$$D(x\|y) := x \log(x/y) - x + y,$$

the relative entropy between two Poisson distributions with means x and y. We can lower bound (29) by (79), as shown at the top of the next page, where the inequality (78), as shown at the top of the next page, follows from the lower bound for the gamma function $\Gamma(x+1) = x! \geq x^x e^{-x}\sqrt{2\pi x}$ [22, p.54], and the equality (79) follows from the use of the formula for $D(x\|y)$ and some rearrangement of terms.

We now study the convergence of each of the terms in (79). All convergence statements are in the almost sure sense. Consider the first term in (79). From Proposition 10 and Proposition 3, as $n \to \infty$, we have

$$\frac{Y_i^n}{N_i^n + 1} \to R_1, \quad \liminf_{n\to\infty} \frac{Y^n - Y_j^n}{n - N_j^n} \geq R'_{min},$$

and

$$\liminf_{n\to\infty} \frac{N_i^n + 1}{n} \geq c_K.$$

Consequently, and using the fact that $D(x\|y)$ is monotone increasing in y, for $y > x$, we have

$$\liminf_{n\to\infty} \frac{N_i^n + 1}{n} D\left(\frac{Y_i^n}{N_i^n + 1} \| \frac{Y^n - Y_j^n}{n - N_j^n}\right) \geq c_K D\left(R_1 \| R'_{min}\right) > 0.$$

$$
\begin{aligned}
Z_{ij}(n) &= \log\left(\frac{\Gamma(Y_i^n+1)}{(N_i^n+1)^{(Y_i^n+1)}}\frac{\Gamma(Y^n-Y_i^n+1)}{(n-N_i^n+1)^{(Y^n-Y_i^n+1)}}\right) - Y_j^n\left(\log\left(\frac{Y_j^n}{N_j^n}\right)-1\right) - (Y^n-Y_j^n)\left(\log\left(\frac{Y^n-Y_j^n}{n-N_j^n}\right)-1\right)\\
&\geq Y_i^n \log\left(\frac{Y_i^n}{N_i^n+1}\right) - Y_i^n + \log\left(\frac{\sqrt{2\pi Y_i^n}}{N_i^n+1}\right) + (Y^n-Y_i^n)\log\left(\frac{Y^n-Y_i^n}{n-N_i^n+1}\right) - (Y^n-Y_i^n)\\
&\quad + \log\left(\frac{\sqrt{2\pi(Y^n-Y_i^n)}}{n-N_i^n+1}\right) - \left[Y_j^n\log\left(\frac{Y_j^n}{N_j^n}\right) - Y_j^n + (Y^n-Y_j^n)\log\left(\frac{Y^n-Y_j^n}{n-N_j^n}\right) - (Y^n-Y_j^n)\right] \qquad (78)\\
&= (N_i^n+1)D\left(\frac{Y_i^n}{N_i^n+1}\Big\|\frac{Y^n-Y_j^n}{n-N_j^n}\right) + (n-N_i^n-N_j^n)D\left(\frac{Y^n-Y_i^n-Y_j^n}{n-N_i^n-N_j^n}\Big\|\frac{Y^n-Y_j^n}{n-N_j^n}\right)\\
&\quad - N_j^n D\left(\frac{Y_j^n}{N_j^n}\Big\|\frac{Y^n-Y_i^n}{n-N_i^n+1}\right) - (n-N_i^n-N_j^n)D\left(\frac{Y^n-Y_i^n-Y_j^n}{n-N_i^n-N_j^n}\Big\|\frac{Y^n-Y_i^n}{n-N_i^n+1}\right)\\
&\quad - \frac{Y^n-Y_i^n}{n-N_i^n+1} - \frac{Y^n-Y_j^n}{n-N_j^n} + \log\left(\frac{\sqrt{2\pi Y_i^n}}{N_i^n+1}\right) + \log\left(\frac{\sqrt{2\pi(Y^n-Y_i^n)}}{n-N_i^n+1}\right). \qquad (79)
\end{aligned}
$$

Similarly, for the the second term in (79), we have

$$\frac{Y^n-Y_i^n-Y_j^n}{n-N_i^n-N_j^n} \to R_2, \quad \limsup_{n\to\infty}\frac{Y^n-Y_j^n}{n-N_j^n} \leq R'_{max},$$

and

$$\liminf_{n\to\infty}\frac{n-N_i^n-N_j^n}{n} \geq (K-2)c_K \geq c_K.$$

Consequently, and using the fact that $D(x\|y)$ is monotone decreasing in y, for $y < x$, we have

$$\liminf_{n\to\infty}\frac{n-N_i^n-N_j^n}{n}D\left(\frac{Y^n-Y_i^n-Y_j^n}{n-N_i^n-N_j^n}\Big\|\frac{Y^n-Y_j^n}{n-N_j^n}\right) \geq c_K D(R_2\|R'_{max}) > 0.$$

Consider the third term in (79). From Proposition 10, as $n \to \infty$, we have

$$\frac{Y_j^n}{N_j^n} \to R_2 \text{ and } \frac{Y^n-Y_i^n}{n-N_i^n+1} \to R_2.$$

Consequently,

$$D\left(\frac{Y^n-Y_i^n-Y_j^n}{n-N_i^n-N_j^n}\Big\|\frac{Y^n-Y_j^n}{n-N_j^n}\right) \to D(R_2\|R_2) = 0.$$

Similarly, for the fourth term in (79) we get

$$\frac{Y^n-Y_i^n-Y_j^n}{n-N_i^n-N_j^n} \to R_2 \text{ and } \frac{Y^n-Y_i^n}{n-N_i^n+1} \to R_2.$$

Consequently,

$$D\left(\frac{Y^n-Y_i^n-Y_j^n}{n-N_i^n-N_j^n}\Big\|\frac{Y^n-Y_i^n}{n-N_i^n+1}\right) \to D(R_2\|R_2) = 0.$$

Consider the fifth and sixth terms in (79). From Proposition 10, we have

$$\frac{Y^n-Y_i^n}{n-N_i^n+1} \to R_2 \text{ and } \limsup_{n\to\infty}\frac{Y^n-Y_j^n}{n-N_j^n} \leq R'_{max}.$$

Consequently, when divided by n and as $n \to \infty$, both the terms go to zero, i.e.,

$$\frac{1}{n}\frac{Y^n-Y_i^n}{n-N_i^n+1} \to 0 \text{ and } \limsup_{n\to\infty}\frac{1}{n}\frac{Y^n-Y_j^n}{n-N_j^n} = 0.$$

Consider the seventh and eight terms in (79). Both the terms go to negative infinity, but only logarithmically in n, and hence when divided by n and as $n \to \infty$, we get

$$\frac{1}{n}\log\left(\frac{\sqrt{2\pi Y_i^n}}{N_i^n+1}\right) \to 0 \text{ and } \frac{1}{n}\log\left(\frac{\sqrt{2\pi(Y^n-Y_i^n)}}{n-N_i^n+1}\right) \to 0.$$

Thus, we have

$$
\begin{aligned}
&\liminf_{n\to\infty}\frac{Z_{ij}(n)}{n}\\
&\geq \liminf_{n\to\infty}\left[\frac{(N_i^n+1)}{n}D\left(\frac{Y_i^n}{N_i^n+1}\Big\|\frac{Y^n-Y_j^n}{n-N_j^n}\right) + \frac{(n-N_i^n-N_j^n)}{n}D\left(\frac{Y^n-Y_i^n-Y_j^n}{n-N_i^n-N_j^n}\Big\|\frac{Y^n-Y_j^n}{n-N_j^n}\right)\right] \qquad (80)\\
&\geq c_K D(R_1\|R'_{min}) + c_K D(R_2\|R'_{max}) \qquad (81)\\
&> 0.
\end{aligned}
$$

This completes the proof of Lemma 11. ■

Proof of Proposition 4: We now have the ingredients to prove Proposition 4. The following inequalities hold almost surely,

$$
\begin{aligned}
\tau(\pi_M(L)) &\leq \tau(\pi_M^i(L))\\
&= \inf\{n \geq 1 | Z_i(n) > \log((K-1)L)\}\\
&\leq \inf\{n \geq 1 | Z_{ij}(n') > \log((K-1)L)\\
&\qquad \text{for all } n' \geq n \text{ and for all } j \neq i\}\\
&< \infty, \qquad (82)
\end{aligned}
$$

where the last inequality follows from Lemma 11. ■

While in Proposition 10 we established that under the non-stopping policy $\tilde{\pi}_M$ $Y_k^n/N_k^n \to R_k$ almost surely, the question of convergence of N_k^n/n to some real constant under the $\tilde{\pi}_M$ policy remained to be established. We now show that under the $\tilde{\pi}_M$ policy it does converge to a real constant. Furthermore, we show that $(Y^n - Y_j^n)/(n - N_j^n)$ also converges to a constant.

Proposition 12: Fix $K \geq 3$. Let $\Psi = (i, R_1, R_2)$ be the true configuration. Consider the non-stopping policy $\tilde{\pi}_M$. Then as $n \to \infty$, the following convergences hold almost surely,

(i)
$$i^*(n) \to i, \tag{83}$$
(ii)
$$\hat{\theta}^n_{i^*(n),1} \to R_1, \tag{84}$$
(iii)
$$\hat{\theta}^n_{i^*(n),2} \to R_2, \tag{85}$$
(iv)
$$\lambda^*(i^*(n), \hat{\theta}^n_{i^*(n),1}, \hat{\theta}^n_{i^*(n),2}) \to \lambda^*(i, R_1, R_2), \tag{86}$$
(v)
$$\frac{N_j^n}{n} \to \lambda^*(i, R_1, R_2)(j) \quad \text{for all } j = 1, 2, \ldots, K, \tag{87}$$
(vi)
$$\frac{Y^n - Y_j^n}{n - N_j^n} \to \tilde{R}(\lambda^*(i, R_1, R_2)(i)) \quad \text{for all } j \neq i, \tag{88}$$

where $\tilde{R}$ is as in (12).

Proof: From Lemma 11 we have
$$\liminf_{n\to\infty} Z_i(n) = \liminf_{n\to\infty} \min_{j\neq i} Z_{ij}(n) > 0 \text{ almost surely.} \tag{89}$$

Fix $j \neq i$. Then, the following inequalities hold almost surely,
$$\begin{aligned}\limsup_{n\to\infty} Z_j(n) &= \limsup_{n\to\infty} \min_{k\neq j} Z_{jk}(n)\\ &\leq \limsup_{n\to\infty} Z_{ji}(n)\\ &\leq \limsup_{n\to\infty} -Z_{ij}(n)\\ &\leq -\liminf_{n\to\infty} \min_{k\neq i} Z_{ik}(n)\\ &= -\liminf_{n\to\infty} Z_i(n)\\ &< 0.\end{aligned}$$

It further implies, $i^*(n) = \max_k Z_k(n) = i$ almost surely. This proves (i).

All convergence statements are in the almost sure sense. From (i) and Proposition 10 we get
$$\hat{\theta}^n_{i^*(n),1} = \frac{Y^n_{i^*(n)}}{N^n_{i^*(n)}} \to \frac{Y_i^n}{N_i^n} \to R_1,$$
and similarly we get,
$$\hat{\theta}_{i^*(n),2} = \frac{Y^n - Y^n_{i^*(n)}}{n - N^n_{i^*(n)}} \to \frac{Y^n - Y_i^n}{n - N_i^n} \to R_2.$$

This proves (ii) and (iii).

From (i), (ii) and (iii) we have
$$\begin{aligned}\lambda^*(i^*(n), \hat{\theta}^n_{i^*(n),1}, \hat{\theta}^n_{i^*(n),2}) &\to \lambda^*(i, \hat{\theta}^n_{i,1}, \hat{\theta}^n_{i,2})\\ &\to \lambda^*(i, R_1, R_2),\end{aligned}$$

where we have used that fact that $\lambda^*(i, x, y)$ is jointly continuous in (x, y), a fact that follows from Berge's Maximum Theorem [23].

Consider the martingale sequence $N_j^n - \sum_{l=1}^n \lambda^*(i^*(n), \hat{\theta}^n_{i^*(n),1}, \hat{\theta}^n_{i^*(n),2})(j)$. From (iv) and martingale convergence arguments, as used in (72), we get
$$\begin{aligned}\frac{N_j^n}{n} &\to \frac{1}{n}\sum_{l=1}^n \lambda^*(i^*(n), \hat{\theta}^n_{i^*(n),1}, \hat{\theta}^n_{i^*(n),2})(j)\\ &\to \lambda^*(i, R_1, R_2)(j).\end{aligned}$$

For ease of notation, let $\lambda^*(i)$ denote $\lambda^*(i, R_1, R_2)(i)$. We can rewrite $(Y^n - Y_j^n)/(n - N_j^n)$ as
$$\begin{aligned}&\frac{Y^n - Y_j^n}{n - N_j^n}\\ &\quad= \frac{Y_i^n + Y^n - Y_i^n - Y_j^n}{n - N_j^n}\\ &\quad= \left[\frac{N_i^n}{n}\frac{Y_i^n}{N_i^n} + \frac{n - N_i^n - N_j^n}{n}\frac{Y^n - Y_i^n - Y_j^n}{n - N_i^n - N_j^n}\right]\frac{n}{n - N_j^n}.\end{aligned}$$

Then, from (v) we have the following convergence in almost sure sense,
$$\begin{aligned}\frac{Y^n - Y_j^n}{n - N_j^n} &\to \frac{\lambda^*(i)R_1 + (1 - \lambda^*(i))\frac{(K-2)}{(K-1)}R_2}{\lambda^*(i) + (1 - \lambda^*(i))\frac{(K-2)}{(K-1)}}\\ &= \tilde{R}(\lambda^*(i)).\end{aligned}$$

This completes the proof of the Proposition. ∎

B. Proof of Proposition 6 and Some of Its Consequences

This subsection is organised as follows. We first establish Proposition 6 which is that log-likelihood ratio associated with the oddball index, under the non-stopping policy, has the correct drift. Lemmas 13 and 14 are then some immediate consequences on the corresponding stopping policy $\pi_M(L)$.

Proof of Proposition 6: We already established (79). Using Proposition 12, we now recognise that all the fractions converge to their respective quantities. Hence,
$$\begin{aligned}&\liminf_{n\to\infty} \frac{Z_{ij}(n)}{n}\\ &\quad\geq \liminf_{n\to\infty} \Bigg[\frac{(N_i^n + 1)}{n} D\left(\frac{Y_i^n}{N_i^n + 1}\Big\|\frac{Y^n - Y_j^n}{n - N_j^n}\right)\\ &\qquad + \frac{(n - N_i^n - N_j^n)}{n} D\left(\frac{Y^n - Y_i^n - Y_j^n}{n - N_i^n - N_j^n}\Big\|\frac{Y^n - Y_j^n}{n - N_j^n}\right)\Bigg]\end{aligned} \tag{90}$$
$$\begin{aligned}&\quad= (\lambda^*(i, R_1, R_2)(i))D(R_1\|\tilde{R})\\ &\qquad + (1 - (\lambda^*(i, R_1, R_2)(i)))\frac{(K-2)}{(K-1)}D(R_2\|\tilde{R})\end{aligned} \tag{91}$$
$$\quad= D^*(i, R_1, R_2) \quad \text{almost surely.} \tag{92}$$

Similarly, by using $\Gamma(x + 1) = x! \leq x^x e^{-x+1}\sqrt{2\pi x}$, and following the steps leading to (79) with limsup instead of liminf, it can be shown that $\limsup_{n\to\infty} \frac{Z_{ij}(n)}{n} \leq D^*(i, R_1, R_2)$

almost surely. It follows that

$$\lim_{n\to\infty} \frac{Z_i(n)}{n} = D^*(i, R_1, R_2) \text{ almost surely,}$$

which establishes Proposition 6. ∎

From Proposition 1 we know that the expected stopping time, $E[\tau(\pi_M(L))]$, grows to infinity as $L \to \infty$, but we now show that $\tau(\pi_M(L))$ grows to infinity in almost sure sense also.

Lemma 13: Fix $K \geq 3$. Let $\Psi = (i, R_1, R_2)$ be the true configuration. Consider the policy $\pi_M(L)$. Then,

$$\liminf_{L\to\infty} \tau(\pi_M(L)) \to \infty \text{ almost surely.} \tag{93}$$

Proof: It is evident that the sequence of random variables $\tau(\pi_M(L))$, indexed by L, is non-decreasing in L. Hence, it suffices to show that, as $L \to \infty$,

$$P\left(\tau(\pi_M(L)) < n\right) \to 0 \text{ for all } n. \tag{94}$$

To see this, observe that

$$\begin{aligned}
&\limsup_{L\to\infty} P\left(\tau(\pi_M(L)) < n\right) \\
&= \limsup_{L\to\infty} P\left(\max_{1\leq l\leq n} Z_j(l) > \log((K-1)L) \text{ for some } j\right) \\
&\leq \limsup_{L\to\infty} \sum_{j=1}^{K}\sum_{l=1}^{n} P\left(Z_j(l) > \log((K-1)L)\right) \qquad (95)\\
&\leq \limsup_{L\to\infty} \frac{1}{\log((K-1)L)} \sum_{j=1}^{K}\sum_{l=1}^{n} E\left[l + 2(Y^l)^2\right] \qquad (96)\\
&\leq \limsup_{L\to\infty} \frac{1}{\log((K-1)L)} \sum_{j=1}^{K}\sum_{l=1}^{n} \Big[l + 2l^2(\max\{R_1, R_2\} \\
&\qquad + (\max\{R_1, R_2\})^2)\Big] \\
&= 0. \qquad (97)
\end{aligned}$$

Inequality (95) follows from union bound. In inequality (97) we have used the convexity of x^2 to bound $E[(\sum_{k=1}^{l} X_k)^2] < l^2 E[(X_k)^2]$, and also that for Poisson random variables $E[X^2] = E[X] + E[X]^2$. Inequality (96) is obtained by using the Markov inequality and by bounding $Z_j(l)$ as follows:

$$\begin{aligned}
Z_j(l) &= \log\left(\frac{f(X^l, A^l|H=j)}{\max_{k\neq j} \hat{f}(X^l, A^l|H=i)}\right) \\
&\leq \log\left(\frac{\hat{f}(X^l, A^l|H=j)}{\hat{f}(X^l, A^l|H=k)}\right) \text{ for some } k \neq j \qquad (98)\\
&= Y_j^l \log\left(\frac{Y_j^l}{N_j^l}\right) - Y_j^l \\
&\quad + (Y^l - Y_j^l)\log\left(\frac{Y^l - Y_j^l}{l - N_j^l}\right) - (Y^l - Y_j^l) \\
&\quad - \Bigg[Y_k^l \log\left(\frac{Y_k^l}{N_k^l}\right) - Y_k^l \\
&\quad + (Y^l - Y_k^l)\log\left(\frac{Y^l - Y_k^l}{l - N_k^l}\right) - (Y^l - Y_k^l)\Bigg] \qquad (99)\\
&\leq (Y_j^l)^2 + (Y^l - Y_j^l)^2 + l \\
&\quad - \left[N_k^l D\left(\frac{Y_k^l}{N_k^l}\|1\right) + (l - N_k^l) D\left(\frac{Y^l - Y_k^l}{l - N_k^l}\|1\right)\right] \qquad (100)\\
&\leq (Y_j^l)^2 + (Y^l - Y_j^l)^2 + l \qquad (101)\\
&\leq 2(Y^l)^2 + l. \qquad (102)
\end{aligned}$$

Inequality (98) follows by upper bounding the numerator in by the maximum likelihood function and lower bounding the denominator by choosing the maximum likelihood function with respect to an arbitrary $k \neq j$ instead of the maximiser. Inequality (100) follows by recognising that the terms inside square brackets in (99) can be written as a sum of relative entropy terms minus an l. Also, we upper bound $x\log(x/N) - x$ by x^2. Inequality (101) follows by ignoring the negative terms. Inequality (102) follows by upper bounding Y_j^l and $Y^l - Y_j^l$ by Y^l. ∎

In Proposition 6 we showed that, as $n \to \infty$ and under the non-stopping policy $\tilde{\pi}_M$, $Z_i(n)/n$ converges to $D^*(i, R_1, R_2)$. We now show that, as $L \to \infty$, $Z_i(\tau(\pi_M(L)))/\tau(\pi_M(L)) \to D^*(i, R_1, R_2)$.

Lemma 14: Fix $K \geq 3$. Let $\Psi = (i, R_1, R_2)$ be the true configuration. Consider the policy $\pi_M(L)$. We then have

$$\lim_{L\to\infty} \frac{Z_i(\tau(\pi_M(L)))}{\tau(\pi_M(L))} = D^*(i, R_1, R_2) \text{ almost surely.} \tag{103}$$

Proof: It follows from Proposition 6 and Lemma 13. ∎

C. Proof of Proposition 7

We now have all the ingredients to prove the main achievability result of Proposition 7. By the definition of $\tau(\pi_M(L))$, we have that $Z_i(\tau(\pi_M(L)) - 1) < \log((K-1)L)$ at the previous slot. Using this we get

$$\limsup_{L\to\infty} \frac{Z_i(\tau(\pi_M(L)) - 1)}{\log(L)} \leq \limsup_{L\to\infty} \frac{\log((K-1)L)}{\log L}. \tag{104}$$

Substituting (103) in (104), we get

$$\begin{aligned}
\limsup_{L\to\infty} \frac{\tau(\pi_M(L))}{\log(L)} &= \limsup_{L\to\infty} \frac{\tau(\pi_M(L)) - 1}{\log(L)} \qquad (105)\\
&\leq \frac{1}{D^*(i, R_1, R_2)}. \qquad (106)
\end{aligned}$$

A sufficient condition to establish convergence of the expected stopping time is to show that

$$\limsup_{L\to\infty} E\left[\exp\left(\frac{\tau(\pi_M(L))}{\log(L)}\right)\right] < \infty.$$

Without loss of generality assume $R_1 < R_2$, such that $R_1 < R'_{min} < R'_{max} < R_2$, where R'_{min} and R'_{max} are as defined in (69) and (70), respectively. Let c_K be as in Proposition 3. We then have

$$\begin{aligned}
&\limsup_{L\to\infty} E\left[e^{\frac{\tau(\pi_M(L))}{\log(L)}}\right] \\
&= \limsup_{L\to\infty} \int_{x\geq 0} P\left(\frac{\tau(\pi_M(L))}{\log(L)} > \log(x)\right) dx \qquad (107)\\
&\leq \limsup_{L\to\infty} \int_{x\geq 0} P\left(\tau^i(\pi_M(L)) > \lfloor \log(x)\log(L) \rfloor\right) dx. \qquad (108)
\end{aligned}$$

Let us now define

$$u(L) := \exp\left(\frac{3\log((K-1)L)}{c_K D(R_1\|R'_{min})\log(L)} + \frac{1}{\log(L)}\right). \quad (109)$$

For $x < u(L)$ let us upper bound the probability by 1. We then get the right-hand side of (108) to be

$$\begin{aligned}\limsup_{L\to\infty} \int_{x\geq 0} & P\left(\tau^i(\pi_M(L)) > \lfloor \log(x)\log(L)\rfloor\right) dx \\ \leq \limsup_{L\to\infty} & \Bigg[u(L) \\ & + \int_{x\geq u(L)} P\left(\tau^i(\pi_M(L)) > \lfloor \log(x)\log(L)\rfloor\right) dx \Bigg]. \end{aligned} \quad (110)$$

Recognising that $P\left(\tau^i(\pi_M(L)) > \lfloor \log(x)\log(L)\rfloor\right)$ is constant in the interval

$$x \in \left[\exp\left(\frac{n}{\log(L)}\right), \exp\left(\frac{n+1}{\log(L)}\right)\right)$$

and recognising that the interval length is upper bounded by $\exp\left(\frac{n+1}{\log(L)}\right)$, we can further upper bound (110) by (112), as shown at the top of the next page. To show that the right-hand side of (112) is finite, it suffices to show that for all

$$n \geq \lfloor \log(u(L))\log(L)\rfloor \geq \frac{3\log((K-1)L)}{c_K D(R_1\|R'_{min})}$$

and for sufficiently large L, there exist constants $\gamma > 0$ and $0 < B < \infty$ such that

$$P\left(Z_i(n) < \log((K-1)L)\right) < Be^{-\gamma n}. \quad (113)$$

We now show that such an exponential bound does exist.

Lemma 15: Fix $K \geq 3$. Fix $L > 1$. Let $\Psi = (i, R_1, R_2)$ be the true configuration. Let $u(L)$ be as in (109). Then, there exist constants $\gamma > 0$ and $0 < B < \infty$, independent of L, such that for all $n \geq \lfloor \log(u(L))\log(L)\rfloor$, we have

$$P\left(Z_i(n) < \log((K-1)L)\right) \leq Be^{-\gamma n}. \quad (114)$$

Proof: The following upper bounds for $P\big(Z_i(n) < \log((K-1)L)\big)$ is self evident

$$\begin{aligned} P\left(Z_i(n) < \log((K-1)L)\right)) &= P\left(\min_{j\neq i} Z_{ij}(n) < \log((K-1)L)\right) \\ &\leq \sum_{j\neq i} P\left(Z_{ij}(n) < \log((K-1)L)\right). \end{aligned}$$

It now suffices to show that for every $j \neq i$ the probability term in the above expression is exponentially bounded. We upper bound $Z_{ij}(n)$ in the same way as we earlier did in (79).

$$\begin{aligned} P\big(Z_{ij}(n) & \leq \log((K-1)L)\big) \\ \leq P\Bigg(& (N_i^n+1) D\left(\frac{Y_i^n}{N_i^n+1} \Big\| \frac{Y^n - Y_j^n}{n-N_j^n}\right) \\ & + (n - N_i^n - N_j^n) D\left(\frac{Y^n - Y_i^n - Y_j^n}{n-N_i^n-N_j^n} \Big\| \frac{Y^n - Y_j^n}{n-N_j^n}\right) \\ & - N_j^n D\left(\frac{Y_j^n}{N_j^n} \Big\| \frac{Y^n - Y_i^n}{n-N_i^n+1}\right) \\ & - (n - N_i^n - N_j^n) D\left(\frac{Y^n - Y_i^n - Y_j^n}{n-N_i^n-N_j^n} \Big\| \frac{Y^n - Y_i^n}{n-N_i^n+1}\right) \\ & - \frac{Y^n - Y_i^n}{n-N_i^n+1} - \frac{Y^n - Y_j^n}{n-N_j^n} + \log\left(\frac{\sqrt{2\pi Y_i^n}}{N_i^n+1}\right) \\ & + \log\left(\frac{\sqrt{2\pi(Y^n - Y_i^n)}}{n-N_i^n+1}\right) < \log((K-1)L)\Bigg) \end{aligned} \quad (115)$$

Using union bound, we upper bound (115) by a sum of probability terms as given in (116), as shown at the top of the next page. Let us choose $0 < \epsilon'' < c_K/3$, so that

$$\frac{c_K(1-\epsilon'')}{1 - c_K(1-\epsilon'')} > c_K(1+\epsilon''). \quad (117)$$

We then we choose $\epsilon' > 0$ such that

$$3(c_K(1-\epsilon'')D(R_1\|R'_{min}) - 8\epsilon') > c_K D(R_1\|R'_{min}),$$

so that

$$\begin{aligned} P\Bigg(& (N_i^n+1)D(R_1\|R'_{min}) - 8\epsilon' n \\ & < \log((K-1)L), (N_i^n+1) > c_K(1-\epsilon'')n\Bigg) = 0 \end{aligned} \quad (118)$$

for all n under consideration, i.e., for all

$$n \geq \lfloor \log(u(L))\log(L)\rfloor \geq \frac{3\log((K-1)L)}{c_K D(R_1\|R'_{min})}.$$

The last term in (116) can then be upper bounded by

$$\begin{aligned} P & \left((N_i^n+1)D(R_1\|R'_{min}) - 8\epsilon' n < \log((K-1)L)\right) \\ & \leq P\Bigg((N_i^n+1)D(R_1\|R'_{min}) - 8\epsilon' n \\ & \qquad < \log((K-1)L), (N_i^n+1) > c_K(1-\epsilon'')n\Bigg) \\ & \quad + P\left((N_i^n+1) \leq c_K(1-\epsilon'')n\right) \\ & = 0 + P\left((N_i^n+1) \leq c_K(1-\epsilon'')n\right) \quad (119) \\ & \leq \exp(-\frac{\epsilon'' n}{2}). \quad (120) \end{aligned}$$

Equality (119) follows from (118). From Proposition 3, we recognise that $(N_{j'}^n - nc_K)$ is a bounded difference submartingale for all j'. Hence, inequality (120) follows from the Azuma-Hoeffding inequality for bounded difference submartingales. Note that only the last term in (116) is dependent on L. By the choice of ϵ' and for all n under consideration, and from (120), we have shown that it decays exponentially with n, and independent of L.

$$\limsup_{L\to\infty}\int_{x\geq 0} P\left(\tau^i(\pi_M(L)) > \lfloor \log(x)\log(L)\rfloor\right)dx$$
$$\leq \exp\left(\frac{3}{c_K D(R_1\|R'_{min})}\right) + \limsup_{L\to\infty}\sum_{n\geq\lfloor\log(u(L))\log(L)\rfloor}\exp\left(\frac{n+1}{\log(L)}\right)P\left(\tau^i(\pi_M(L)) > n\right)dx \tag{111}$$
$$\leq \exp\left(\frac{3}{c_K D(R_1\|R'_{min})}\right) + \limsup_{L\to\infty}\sum_{n\geq\lfloor\log(u(L))\log(L)\rfloor}\exp\left(\frac{n+1}{\log(L)}\right)P\left(Z_i(n) < \log((K-1)L)\right)dx. \tag{112}$$

$$P\left(Z_{ij}(n) \leq \log((K-1)L)\right)$$
$$\leq P\left((N_i^n+1)\left(D\left(\frac{Y_i^n}{N_i^n+1}\|\frac{Y^n-Y_j^n}{n-N_j^n}\right) - D(R_1\|R'_{min})\right) < -\epsilon' n\right)$$
$$+ P\left((n-N_i^n-N_j^n)D\left(\frac{Y^n-Y_i^n-Y_j^n}{n-N_i^n-N_j^n}\|\frac{Y^n-Y_j^n}{n-N_j^n}\right) < -\epsilon' n\right)$$
$$+ P\left(-N_j^n D\left(\frac{Y_j^n}{N_j^n}\|\frac{Y^n-Y_i^n}{n-N_i^n+1}\right) < -\epsilon' n\right)$$
$$+ P\left(-(n-N_i^n-N_j^n)D\left(\frac{Y^n-Y_i^n-Y_j^n}{n-N_i^n-N_j^n}\|\frac{Y^n-Y_i^n}{n-N_i^n+1}\right) < -\epsilon' n\right)$$
$$+ P\left(-\frac{Y^n-Y_i^n}{n-N_i^n+1} < -\epsilon' n\right) + P\left(-\frac{Y^n-Y_j^n}{n-N_j^n} < -\epsilon' n\right)$$
$$+ P\left(\log\left(\frac{\sqrt{2\pi Y_i^n}}{N_i^n+1}\right) < -\epsilon' n\right) + P\left(\log\left(\frac{\sqrt{2\pi(Y^n-Y_i^n)}}{n-N_i^n+1}\right) < -\epsilon' n\right)$$
$$+ P\left((N_i^n+1)D(R_1\|R'_{min}) - 8\epsilon' n < \log((K-1)L)\right). \tag{116}$$

It now suffices to show that each of the other terms in (116) decays exponentially with n. Let us now look at the first term in (116).

$$P\left((N_i^n+1)\left(D\left(\frac{Y_i^n}{N_i^n+1}\|\frac{Y^n-Y_j^n}{n-N_j^n}\right) - D(R_1\|R'_{min})\right) < -\epsilon' n\right)$$
$$\leq P\left((N_i^n+1)\left(D\left(\frac{Y_i^n}{N_i^n+1}\|\frac{Y^n-Y_j^n}{n-N_j^n}\right) - D(R_1\|R'_{min})\right) < -\epsilon' n,\ N_{j'}^n \geq c_K(1-\epsilon'')n\ \forall j'\right)$$
$$+ \sum_{j'} P\left(N_{j'}^n < c_K(1-\epsilon'')n\right). \tag{121}$$

All the terms inside the summation in (121) have exponential bounds from Proposition 3 and from Azuma-Hoeffding inequality for bounded difference sub-martingales. The first term in (121) can be further upper bounded by (122), as shown at the top of the next page. Inequality (122) follows by replacing $D(R_1\|R'_{min})$ by a larger $D\left(R_1\|\frac{Y^n-Y_j^n}{n-N_j^n}\right)$ using the fact that $D(x\|y)$ is monotonically increasing in y for $y > x$. Let us now consider the first term in (122). Recognise that we have restricted $\frac{Y^n-Y_j^n}{n-N_j^n}$ to lie in a compact interval $[R'_{min}, R_2]$. Further, since $D(x\|y)$ is jointly continuous in (x, y) and since the second argument is restricted to a compact set, we can upper bound the first term in (122), for a suitable δ_ϵ, by (123), as shown at the top of the next page. We recognise that (123) can be expressed as the probability of the deviation of a martingale difference sequence from zero, which we know can be exponentially bounded using the martingale concentration bounds of De la Pena [20, Th. 1.2A], given in (71)

Let us define $R''_{min} := R'_{min} + c_K\epsilon''(R_2 - R_1)$ and $R''_{max} := R'_{max} - c_K\epsilon''(R_2-R_1)$. Let $\epsilon''' > 0$ be such that $R'_{min}+2\epsilon''' < R''_{min}$ and $R''_{max}+2\epsilon''' < R'_{max}$. We then recognise that, given the event $\{N_{j'}^n \geq c_K(1-\epsilon'')n\ \forall j'\}$ and (117), the event

$$\left\{\frac{N_i^n R_1 + (n-N_i^n-N_j^n)R_2}{n-N_j^n} \geq (1-c_K(1+\epsilon''))R_1 + c_K(1+\epsilon'')R_2 = R''_{min}\right\}$$

is also true. Then, the following statements are true

$$\left\{\frac{Y^n-Y_j^n}{n-N_j^n} < R'_{min}\right\}$$
$$\subseteq \left\{\frac{Y^n-Y_j^n}{n-N_j^n} < R''_{min} - \epsilon'''\right\}$$

$$P\left((N_i^n+1)\left(D\left(\frac{Y_i^n}{N_i^n+1}\Big\|\frac{Y^n-Y_j^n}{n-N_j^n}\right)-D(R_1\|R'_{min})\right)<-\epsilon' n,\ N_{j'}^n\geq c_K(1-\epsilon'')n\ \forall j'\right)$$
$$\leq P\left((N_i^n+1)\left(D\left(\frac{Y_i^n}{N_i^n+1}\Big\|\frac{Y^n-Y_j^n}{n-N_j^n}\right)-D\left(R_1\Big\|\frac{Y^n-Y_j^n}{n-N_j^n}\right)\right)<-\epsilon' n,\right.$$
$$\left.N_{j'}^n\geq c_K(1-\epsilon'')n\ \forall j',\ \frac{Y^n-Y_j^n}{n-N_j^n}\geq R'_{min},\ \frac{Y^n-Y_j^n}{n-N_j^n}\leq R_2\right)$$
$$+P\left(\frac{Y^n-Y_j^n}{n-N_j^n}<R'_{min},\ N_{j'}^n\geq c_K(1-\epsilon'')n\ \forall j'\right)+P\left(\frac{Y^n-Y_j^n}{n-N_j^n}>R_2,\ N_{j'}^n\geq c_K(1-\epsilon'')n\ \forall j'\right). \tag{122}$$

$$P\left((N_i^n+1)\left(D\left(\frac{Y_i^n}{N_i^n+1}\Big\|\frac{Y^n-Y_j^n}{n-N_j^n}\right)-D\left(R_1\Big\|\frac{Y^n-Y_j^n}{n-N_j^n}\right)\right)<-\epsilon' n,\right.$$
$$\left.N_{j'}^n\geq c_K(1-\epsilon'')n\ \forall j',\ \frac{Y^n-Y_j^n}{n-N_j^n}\geq R'_{min},\ \frac{Y^n-Y_j^n}{n-N_j^n}\leq R_2\right)$$
$$\leq P\left(\left|\frac{Y_i^n}{N_i^n+1}-R_1\right|>\delta_\epsilon,\ N_{j'}^n\geq c_K(1-\epsilon'')n\ \forall j'\right). \tag{123}$$

$$\subseteq\left\{\frac{Y^n-Y_j^n}{n-N_j^n}<\frac{N_i^nR_1+(n-N_i^n-N_j^n)R_2}{n-N_j^n}-\epsilon'''\right\}$$
$$\subseteq\left\{\left|\frac{Y^n-Y_j^n}{n-N_j^n}-\frac{N_i^nR_1+(n-N_i^n-N_j^n)R_2}{n-N_j^n}\right|>\epsilon'''\right\}. \tag{124}$$

Similarly, given the event $\{N_{j'}^n\geq c_K(1-\epsilon'')n\ \forall j'\}$, we can show that

$$\left\{\frac{Y^n-Y_j^n}{n-N_j^n}>R_2\right\}$$
$$\subseteq\left\{\left|\frac{Y^n-Y_j^n}{n-N_j^n}-\frac{N_i^nR_1+(n-N_i^n-N_j^n)R_2}{n-N_j^n}\right|>\epsilon'''\right\}. \tag{125}$$

From (124) and (125), the second and third term in (122) can then be upper bounded by

$$P\left(\frac{Y^n-Y_j^n}{n-N_j^n}<R'_{min},\ N_{j'}^n\geq c_K(1-\epsilon'')n\ \forall j'\right)$$
$$+P\left(\frac{Y^n-Y_j^n}{n-N_j^n}>R_2,\ N_{j'}^n\geq c_K(1-\epsilon'')n\ \forall j'\right)$$
$$\leq 2P\left(\left|\frac{Y^n-Y_j^n}{n-N_j^n}-\frac{N_i^nR_1+(n-N_i^n-N_j^n)R_2}{n-N_j^n}\right|>\epsilon''',\right.$$
$$\left.N_{j'}^n\geq c_K(1-\epsilon'')n\ \forall j'\right). \tag{126}$$

Again, we recognise that (126) can be expressed as the probability of the deviation of a martingale difference sequence from zero, which we know can be exponentially bounded using the martingale concentration bounds of De la Pena [20, Th. 1.2A], given in (71).

Let us now look at the other terms in (116). The second term is identically zero, as the left-hand side is always positive. Arguments similar to those of the first term hold for the third and fourth terms. For the fifth and sixth terms, the left-hand sides converge to a constant, while the right-hand side goes to negative infinity, and thus its straightforward to obtain exponential bounds for these terms. Similarly, for the seventh and eight terms, the left-hand side goes to negative infinity at a logarithmic rate, while the right-hand side goes to negative infinity at a faster linear rate, and again it is straightforward to obtain exponential bounds for these terms. This completes the proof for Lemma 15. ■

This completes the proof of our main achievability result of Proposition 7. ■

Acknowledgements

We would like to thank Dr. S. P. Arun, from the Centre for Neuroscience, IISc Bangalore, and Prof. Carl R. Olson, from the Center for the Neural Basis of Cognition, Carnegie Mellon University, for the experimental data used in Section VI.

References

[1] H. Chernoff, "Sequential design of experiments," *Ann. Math. Statist.*, vol. 30, no. 3, pp. 755–770, 1959.

[2] A. E. Albert, "The sequential design of experiments for infinitely many states of nature," *Ann. Math. Statist.*, vol. 32, no. 3, pp. 774–799, 1961.

[3] E. Kaufmann and O. Cappé, and A. Garivier, "On the complexity of best-arm identification in multi-armed bandit models," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 1–42, 2016.

[4] A. P. Sripati 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.

[5] N. K. Vaidhiyan, S. P. Arun, and R. Sundaresan. (May 2015). "Neural dissimilarity indices that predict oddball detection in behaviour." [Online]. Available: https://arxiv.org/abs/1505.02362

[6] M. Naghshvar and T. Javidi, "Active *M*-ary sequential hypothesis testing," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 1623–1627.

[7] M. Naghshvar and T. Javidi, "Active sequential hypothesis testing," *Ann. Statist.*, vol. 41, no. 6, pp. 2703–2738, Dec. 2013.
[8] M. Naghshvar and T. Javidi, "Information utility in active sequential hypothesis testing," in *Proc. 48th Annu. Allerton Conf. Commun., Control, Comput.*, Oct. 2010, pp. 123–129.
[9] M. Naghshvar and T. Javidi, "Performance bounds for active sequential hypothesis testing," in *Proc. IEEE Int. Symp. Inf. Theory*, Aug. 2011, pp. 2666–2670.
[10] 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.
[11] 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.
[12] S. Nitinawarat and V. V. Veeravalli, "Controlled sensing for sequential multihypothesis testing with controlled Markovian observations and non-uniform control cost," *Sequential Anal.*, vol. 34, no. 1, pp. 1–24, 2015.
[13] Y. Li, S. Nitinawarat, and V. V. Veeravalli, "Universal outlier hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4066–4082, Jul. 2014.
[14] Y. Li, S. Nitinawarat, and V. V. Veeravalli, "Universal sequential outlier hypothesis testing," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2014, pp. 3205–3209.
[15] S. Nitinawarat and V. V. Veeravalli, "Universal scheme for optimal search and stop," *Bernoulli*, vol. 23, no. 3, pp. 1759–1783, 2017.
[16] K. Cohen and Q. Zhao, "Asymptotically optimal anomaly detection via sequential testing," *IEEE Trans. Signal Process.*, vol. 63, no. 11, pp. 2929–2941, Jun. 2015.
[17] A. Garivier and E. Kaufmann, "Optimal best arm identification with fixed confidence," in *Proc. 29th Conf. Learn. Theory*, 2016, pp. 998–1027.
[18] S. Kullback, *Information and Statistics*. New York, NY, USA: Wiley, 1959.
[19] S. M. Ali and S. D. Silvey, "A general class of coefficients of divergence of one distribution from another," *J. Roy. Statist. Soc. Ser. B, Methodol.*, vol. 28, no. 1, pp. 131–142, 1966.
[20] V. H. de L. Peña, "A general class of exponential inequalities for martingales and ratios," *Ann. Probab.*, vol. 27, no. 1, pp. 537–564, 1999.
[21] K. L. Chung, *A Course in Probability Theory*. San Francisco, CA, USA: Academic, 2001.
[22] W. Feller, *An Introduction to Probability Theory and its Applications*, vol. 1. Hoboken, NJ, USA: Wiley, Jan. 1968.
[23] L. M. Ausubel and R. J. Deneckere, "A generalized theorem of the maximum," *Econ. Theory*, vol. 3, no. 1, pp. 99–107, 1993.

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 has been with Qualcomm India Pvt. Ltd., Bangalore where he is currently a Senior Lead Engineer.

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.