Applied Probability Trust (25 February 2011)

TWENTY QUESTIONS WITH NOISE: BAYES OPTIMAL POLICIES FOR ENTROPY LOSS

BRUNO JEDYNAK,* *Johns Hopkins University*

PETER I. FRAZIER,** *Cornell University*

RAPHAEL SZNITMAN,*** *Johns Hopkins University*

Abstract

We consider the problem of 20 questions with noisy answers, in which we seek to find a target by repeatedly choosing a set, asking an oracle whether the target lies in this set, and obtaining an answer corrupted by noise. Starting with a prior distribution on the target's location, we seek to minimize the expected entropy of the posterior distribution. We formulate this problem as a dynamic program and show that any policy optimizing the one-step expected reduction in entropy is also optimal over the full horizon. Two such Bayes-optimal policies are presented: one generalizes the probabilistic bisection policy due to Horstein and the other asks a deterministic set of questions. We study the structural properties of the latter, and illustrate its use in a computer vision application.

Keywords: twenty questions; dynamic programing; bisection; search; object detection; entropy loss; sequential experimental design; Bayesian experimental design

2010 Mathematics Subject Classification: Primary 60J20

Secondary 62C10;90B40;90C39

1. Introduction

In this article, we consider the problem of finding a target $X^* \in \mathbb{R}^d$ by asking a knowledgeable oracle questions. Each question consists in choosing a set $A \subseteq \mathbb{R}^d$,

* Postal address: Whitehead 208-B, 3400 North Charles Street, Baltimore, Maryland, 21218

** Postal address: 232 Rhodes Hall, Cornell University, Ithaca, NY 14853

*** Postal address: Hackerman Hall, 3400 North Charles Street, Baltimore, Maryland, 21218

querying the oracle whether X^* lies in this set, and observing the associated response. While this is closely related to the popular game of “twenty questions”, we consider here the case where answers from the oracle are corrupted with noise from a known model. This game appears naturally in a number of problems in stochastic search, stochastic optimization, and stochastic root finding. In this paper we present an illustrative application in computer vision.

We consider a Bayesian formulation of this problem using entropy loss. In $d = 1$ dimension, we seek to minimize the expected entropy of the posterior after a fixed number of questions and provide two Bayes optimal policies for this problem. The first policy poses questions about intervals, $A = [-\infty, x]$, while the second poses questions about more general sets. In $d = 2$ dimensions, we seek to minimize the maximum expected entropy of the posterior in each dimension, and provide an asymptotically Bayes optimal procedure.

When the noise corrupting the oracle's responses is of a special form, that of a symmetric channel, and when the questions A are restricted to be intervals, the Bayes optimal policy for $d = 1$ takes a particularly natural form: choose $A = [-\infty, x]$ where x is the median of the posterior distribution. This policy, called the probabilistic bisection strategy, was first proposed in [13] (later republished in [14]). This policy was recently shown to be optimal in the binary symmetric case by one of the authors in [10]. [4] introduces a similar procedure that measures on either side of the median of the posterior over a discrete set of points, and shows that its error probability decays at an asymptotically optimal rate. For a review of these two procedures, see [5]. [15, 1] also both consider a noisy binary search problem with constant error probability over a discrete set of points, and give optimality results for policies similar to measuring at the median of the posterior. In [15], this is part of a larger analysis in which the error probability may vary. In addition to this type of stochastic noise, a parallel line of research has considered the case when the oracle is adversarial. For a survey, see [20].

When the questions are restricted to be about intervals, the problem that we consider is similar to the stochastic root-finding problem considered by the seminal paper [23] and generalized to multiple dimensions by [3]. In the stochastic root-finding problem, one chooses a sequence of points $x_1, x_2, \ldots$ to query, and observes the corresponding values $f(x_1), f(x_2), \ldots$ of some decreasing function f at x, obscured by noise. The goal in this problem is to find the root of f. Procedures include the stochastic approximation methods of [23, 3], as well as Polyak-Ruppert averaging introduced independently by [24, 21]. Asymptotic rates of convergence of these procedures are well understood — see [16]. Our problem and the stochastic root-finding problem are similar because, if X^* is the root of f, then querying whether X^* is in $(-\infty, x]$ can be recast as querying whether $f(x) < 0$. The problems differ because the noise in observing whether $f(x) < 0$ depends upon x and is generally larger when $f(x)$ is closer to 0, while in our formulation we assume that the distribution of the oracle's response depends only on whether X^* is in the queried subset or not.

Both our problem and stochastic root-finding lie within the larger class of problems in sequential experimental design, in which we choose at each point in time which experiment to perform in order to optimize some overall value of the information obtained. The study of this area began with [22], who introduced the multi-armed bandit problem later studied by [17, 2, 12, 30, 31] and others. For a self-contained discussion of sequential experimental design in a Bayesian context, see [7].

After formulating the problem in Sec. 2, we provide a general characterization of the optimal policy in Sec. 3. Then, Sec. 4 provides two specific optimal policies for $d = 1$, one using intervals as questions that generalizes the probabilistic bisection policy, and the other using more general subsets as questions that we call the dyadic policy. It also provides further analysis of the dyadic policy: a law of large numbers and a central limit theorem for the posterior entropy; and an explicit characterization of the expected number of size-limited noise-free questions required to find the target after noisy questioning ceases. Sec. 5 considers a modified version of the entropic loss in $d = 2$ dimensions, and shows that a simple modification of the dyadic policy is asymptotically Bayes optimal for this loss function. Sec. 6 provides a detailed illustrative application in computer vision, and Sec. 7 concludes the paper.

2. Formulation of the problem

Nature chooses a continuous random variable X^* with density p_0 with respect to the Lebesgues measure over $\mathbb{R}^d$. The fact that X^* is continuous will turn out to be important and the arguments presented below do not generalize easily to the case where X^* is a discrete random variable.

To discover X^*, we can sequentially ask N questions. Asking the n^{th} question, $0 \leq n \leq N-1$, involves choosing a Lebesgues measurable set $A_n \subset \mathbb{R}^d$ and evaluating: "Is $X^* \in A_n$?". To avoid technical issues below, we require that A_n is the union of at most J_n intervals. The answer, denoted Z_n, is the indicator function of the event $\{X^* \in A_n\}$. However, Z_n is not openly communicated to us. Instead, Z_n is the input of a memoryless noisy transmission channel from which we observe the output Y_{n+1}. Y_{n+1} is a random variable which can be discrete or continuous, univariate or multivariate. The memoryless property of the channel expresses the fact that Y_{n+1} depends on Z_n, but not on previous questions or answers. As a consequence, repeatedly answering the same question may not provide the same answer each time. Moreover, we assume that the distribution of Y_{n+1} given Z_n does not depend on n. Finally, the probability distribution of Y_{n+1} given Z_n is as follows:

$$P\left(Y_{n+1}=y|Z_n=z\right)=\begin{cases} f_1(y) & \text{if } z=1 \\ f_0(y) & \text{if } z=0 \end{cases} \tag{1}$$

Here, we lump together the cases where Y_{n+1} is a discrete and continuous random variable (or vector). In both cases the analysis is very similar: in the former, f_0 and f_1 are point mass functions, while in the latter case, f_0 and f_1 are densities and the expression on the left hand side of the equal sign is the value at y of the conditional density of Y_{n+1} given $Z_n = z$ (with a slight abuse of notation). We require that the Shannon entropy of f_0 and f_1 be finite. At any time step n, we may characterize what we know about X^* by recalling the history of previous measurements $(A_m, Y_{m+1})_{m=0}^{n-1}$ or equivalently by computing the posterior density p_n of X^* given these measurements. The study of the stochastic sequences of densities p_n, under different policies, constitutes the main mathematical contribution of this paper. For an event A, we will use the notation

$$p_n(A)=\int_A p_n(x)dx.$$

The posterior density p_{n+1} of X^* after observing $(Y_m)_{m=1}^{n+1}$ is elegantly described as a function of p_n, f_0, f_1, the n^{th} question $A_n = A$ and the answer to this question $Y_{n+1} = y$.

Lemma 1. *The posterior density on X^* is given by*

$$p_{n+1}(u|A_n = A, Y_{n+1} = y, p_n) = \frac{1}{\mathcal{Z}} \left(f_1(y) \mathbb{1}_{u \in A} + f_0(y) \mathbb{1}_{u \notin A} \right) p_n(u),$$

where

$$\mathcal{Z} = P(Y_{n+1} = y|p_n, A_n = A) = f_1(y)p_n(A) + f_0(y)(1 - p_n(A)) \tag{2}$$

Proof. The result follows from the straightforward use of the Bayes formula. The posterior $p_{n+1}(u|A_n = A, Y_{n+1} = y, p_n)$ can be written as

$$\begin{aligned} &\frac{1}{\mathcal{Z}} P(Y_{n+1} = y|p_n, A_n = A, X^* = u) P(X^* = u|p_n, A_n = A) \\ =&\frac{1}{\mathcal{Z}} \left(f_1(y) \mathbb{1}_{u \in A} + f_0(y) \mathbb{1}_{u \notin A} \right) p_n(u) \end{aligned}$$

where $\mathcal{Z} = \int_u \left(f_1(y) \mathbb{1}_{u \in A} + f_0(y) \mathbb{1}_{u \notin A} \right) p_n(u) du$.

We will measure the quality of the information gained about X^* from these N questions using the Shannon differential entropy. The Shannon differential entropy (see [6] Chapter 9), or simply "the entropy" of p_n, $H(p_n)$, is defined as

$$H(p_n) = -\int_{-\infty}^{+\infty} p_n(x) \log p_n(x) dx$$

where log is the logarithm in base 2. In particular, we consider the problem of finding a sequence of N questions such that the expected entropy of X^* after observing the N^{th} answer is minimized.

We will write this problem more formally as the infimum over policies of the expectation of the posterior entropy, but before doing so we must formally define a policy. Informally, a policy is a method for choosing the questions A_n as a function of the observations available at time n. The technical assumption that each question A_n is a union of only finitely many intervals ensures the Borel-measurability of $H(p_N)$ under each policy.

First, A_n is the union of at most J_n half-open intervals, and so may be written

$$A_n = \bigcup_{j=1}^{J_n} [a_{n,j}, b_{n,j}),$$

where $a_{n,j} \leq b_{n,j}$ are elements of $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. If $a_{n,j} = -\infty$ then the corresponding interval is understood to be open on the left. If A_n comprises strictly less than J_n intervals then we may take $a_{n,j} = b_{n,j}$ for some j. When A_n is written in this way, the space in which A_n takes values may be identified with the space $\mathbb{A}_n = \{(a_j, b_j) : j = 1, \ldots, J_n,\ a_j \leq b_j\}$, which is a closed subset of $\overline{\mathbb{R}}^{2J_n}$.

Then, with p_0 fixed, p_n may be identified with the sequence $((a_{m,j}, b_{m,j})_{j=1}^{J_m}, Y_{m+1})_{m=0}^{n-1}$, which takes values in the space $\mathbb{S}_n = (\mathbb{A}_0 \times \cdots \times \mathbb{A}_{n-1}) \times \mathbb{R}^n$. Furthermore, the function $p_n \mapsto H(p_n)$ may be written as a measurable function from $\mathbb{S}_n$ to $\mathbb{R}$.

Thus, after having identified possible values for A_n with points in $\mathbb{A}_n$ and possible values for p_n with points in $\mathbb{S}_n$, we define a policy π to be a sequence of functions $\pi = (\pi_0, \pi_1, \ldots)$, where $\pi_n : \mathbb{S}_n \mapsto \mathbb{A}_n$ is a measurable function. We let Π be the space of all such policies. Any such policy π induces a probability measure on $((a_{n,j}, b_{n,j})_{j=1}^{J_n}, Y_{n+1})_{n=0}^{N-1}$. We let E^π indicate the expectation with respect to this probability measure. In a slight abuse of notation, we will sometimes talk of $p \in \mathbb{S}_n$ and $A \in \mathbb{A}_n$, by which we mean the density p assocatied with a vector in $\mathbb{S}_n$, or the set A associated with a vector in $\mathbb{A}_n$.

With this definition of a policy π, the associated measure E^π, and the space of all policies Π, the problem under consideration may be written,

$$\inf_{\pi \in \Pi} E^\pi[H(p_N)] \tag{3}$$

Any policy attaining the infimum is called optimal. We consider this problem for the general case in Sec. 3, and for the specific cases of $d = 1$ and $d = 2$ in Sec. 4 and Sec. 5 respectively. In Sec. 5, we also consider a modification of this objective function that separately considers the entropy of the marginal posterior distribution, and ensures that both entropies are small. This prevents a policy from obtaining optimality by learning one coordinate of X^* without learning the other.

3. Entropy Loss and Channel Capacity

In this section we consider the problem (3) of minimizing the expected entropy of the posterior over $\mathbb{R}^d$. We present general results characterizing optimal policies, which will be used to create specific policies in Sec. 4 and Sec. 5.

We first present some notation that will be used within our results. Let φ be the function with domain $[0, 1]$ defined by

$$\varphi(u) = H(uf_1 + (1-u)f_0) - uH(f_1) - (1-u)H(f_0).$$

The channel capacity, denoted C, is the supremum of φ,

$$C = \sup_{u \in [0,1]} \varphi(u).$$

Below, in Theorem 1, we show that this maximum is attained in $(0, 1)$. Let $u^* \in (0, 1)$ be a point attaining this maximum, so $\varphi(u^*) = C$.

We show that an optimal policy consists of choosing each A_n so that $p_n(A_n) = u^*$. When the A_n are chosen in this way, the expected entropy decreases arithmetically by the constant C at each step. Moreover, if the communication channel is symmetric in the sense that $\varphi(1-u) = \varphi(u)$, $\forall 0 \le u \le 1$, then $u^* = \frac{1}{2}$. Optimal policies constructed by choosing $p_n(A_n) = u^*$ are greedy policies (or "knowledge-gradient" policies as defined in [9]), since they make decisions that would be optimal if only one measurement remained, i.e., if N were equal to $n+1$. Such greedy policies are usually used only as heuristics, and so it is interesting that they are optimal in this problem.

Our analysis relies on dynamic programming. To support this analysis, we define the value function,

$$V(p, n) = \inf_{\pi \in \Pi} E^\pi[H(p_N) | p_n = p], \quad p \in \mathbb{S}_n,\ n = 0, \ldots, N.$$

Standard results from controlled Markov processes show that this value function satisfies Bellman's recursion (Section 3.7 of [8]),

$$V(p, n) = \inf_{A \in \mathbb{A}_n} E[V(p_{n+1}, n+1) | A_n = A, p_n = p], \quad p \in \mathbb{S}_n,\ n < N, \tag{4}$$

and any policy attaining the minimum of (4) is optimal (Section 2.3 of [8]). In general, the results of [8] for general Borel models imply only that $V(\cdot, n) : \mathbb{S}_n \mapsto \mathbb{R}$ is universally measurable, and do not imply Borel-measurability. However, we show below in Theorem 2 that, in our case, $V(\cdot, n) : \mathbb{S}_n \mapsto \mathbb{R}$ is a Borel-measurable function.

As a preliminary step toward solving Bellman's recursion, we present the following theorem, which shows that minimizing the expected entropy of the posterior one step into the future can be accomplished by choosing A_n as described above. Furthermore, it shows that the expected reduction in entropy is the channel capacity C.

Theorem 1.

$$\inf_{A\in\mathbb{A}_n} E[H(p_{n+1})|A_n = A; p_n] = H(p_n) - C. \tag{5}$$

Moreover, there exists a point $u^* \in (0,1)$ such that $\varphi(u^*) = C$, and the minimum in (5) *is attained by choosing A such that $p_n(A) = u^*$.*

Proof. We first rewrite the expected entropy as

$$E[H(p_{n+1}) \mid A_n = A, p_n] = H(p_n) - I(X^*, Y_{n+1} \mid A_n = A, p_n),$$

where $I(X^*, Y_{n+1} \mid A_n = A, p_n)$ is the mutual information between the conditional distributions of X^* and Y_{n+1} (see [6] Chapter 2), and we have noted that the entropy of X^* given $A_n = A$ and p_n is exactly $H(p_n)$. This leads to

$$\inf_{A\in\mathbb{A}_n} E[H(p_{n+1}) \mid A_n = A, p_n] = H(p_n) - \sup_{A\in\mathbb{A}_n} I(X^*, Y_{n+1} \mid A_n = A, p_n). \tag{6}$$

Temporarily fixing A, we expand the mutual information as

$$I(X^*, Y_{n+1} \mid A_n = A, p_n) = H(Y_{n+1} \mid A_n = A, p_n) - E[H(Y_{n+1}) \mid X^*, A_n = A, p_n].$$

Using (2),

$$H(Y_{n+1} \mid A_n = A, p_n) = H(p_n(A) f_1 + (1 - p_n(A)) f_0). \tag{7}$$

Also,

$$\begin{aligned} E[H(Y_{n+1}) \mid X^*, A_n = A, p_n] &= \int_u p_n(u) H(Y_{n+1} \mid X^* = u, A_n = A, p_n) du \\ &= \int_{u\in A} p_n(u) H(f_1) du + \int_{u\notin A} p_n(u) H(f_0) du \\ &= H(f_1) p_n(A) + H(f_0)(1 - p_n(A)). \end{aligned} \tag{8}$$

The difference between (7) and (8) is $\varphi(p_n(A))$, and so $I(X^*, Y_{n+1} \mid A_n = A, p_n) = \varphi(p_n(A))$. This and (6) together show that

$$\sup_{A\in\mathbb{A}_n} I(X^*, Y_{n+1} \mid A_n = A, p_n) = \sup_{A\in\mathbb{A}_n} \varphi(p_n(A)) = \sup_{u\in[0,1]} \varphi(u) = C.$$

This shows (5), and that the infimum in (5) is attained by any set A with $\varphi(p_n(A)) = C$. It remains only to show the existence of a point $u^* \in (0,1)$, with $\varphi(u^*) = C$.

First, φ is a continuous function, so its maximum over the compact interval $[0,1]$ is attained. If the maximum is attained in $(0,1)$, then we simply choose u^* to be this point. Now consider the case when the maximum is attained at $u \in \{0,1\}$. Because φ is a mutual information, it is non-negative. Also, $\varphi(0) = \varphi(1) = 0$. Thus, if the maximum is attained at $u \in \{0,1\}$, then $\varphi(u) = 0$ for all u, and one can choose u^* in the open interval $(0,1)$.

We are ready now to present the main result of this section, which gives a simple characterization of optimal policies.

Theorem 2. *Any policy that chooses each A_n to satisfy*

$$p_n(A_n) = u^* \in \arg\max_{u\in[0,1]} \varphi(u) \tag{9}$$

is optimal. In addition, for each n, the value function $V(\cdot,n) : \mathbb{S}_n \mapsto \mathbb{R}$ is Borel-measurable and is given by

$$V(p_n, n) = H(p_n) - (N-n)C. \tag{10}$$

Proof. It is enough to show for each $n = 0, 1, \ldots, N$ that the value function is given by (10), and that the described policy achieves the minimum in Bellman's recursion (4). Measurability of $V(\cdot,n) : \mathbb{S}_n \mapsto \mathbb{R}$ then follows from the fact that $p_n \mapsto H(p_n)$ is Borel-measurable when written as a function from $\mathbb{S}_n$ to $\mathbb{R}$. We proceed by backward induction on n. The value function clearly has the claimed form at the final time $n = N$. Now, fix any $n < N$ and assume that the value function is of the form claimed for $n+1$. Then, Bellman's recursion and the induction hypothesis show,

$$\begin{aligned}
V(p_n, n) &= \inf_{A\in\mathbb{A}_n} E[V(p_{n+1}, n+1) \mid A_n = A, p_n] \\
&= \inf_{A\in\mathbb{A}_n} E[H(p_{n+1}) - (N-n-1)C \mid A_n = A, p_n] \\
&= \inf_{A\in\mathbb{A}_n} E[H(p_{n+1}) \mid A_n = A, p_n] - (N-n-1)C \qquad (11) \\
&= H(p_n) - C - (N-n-1)C \\
&= H(p_n) - (N-n)C
\end{aligned}$$

where we have used Theorem 1 in rewriting (11) in the next line. Theorem 1 also shows that the infimum in (11) is attained when A satisfies $p_n(A) = u^*$, and so the described policy achieves the minimum in Bellman's recursion.

We offer the following interpretation of the optimal reduction in entropy shown in Theorem 2. First, the entropy of a random variable uniformly distributed over $[a, b]$ is $\log(b-a)$. The quantity $2^{H(X)}$ for a continuous random variable X can then be interpreted as the length of the support of a uniform random variable with the same entropy as X. We refer to this quantity more simply as the "length of X." If the prior distribution of X^* is uniform over $[0,1]$, then the length of X^* under p_0 is 1 and Theorem 2 shows that the expected length of X^* under p_N is no less than 2^{-CN}, where this bound on the expected length can be achieved using an optimal policy.

We conclude this section by discussing u^* and C in a few specific cases. In general, there is no simple expression for u^* and for C. However, in certain symmetric cases the following proposition shows that $u^* = \frac{1}{2}$.

Proposition 1. *If the channel has the following symmetry*

$$\varphi(u) = \varphi(1-u), \forall 0 \leq u \leq 1$$

then $\frac{1}{2} \in \arg\max_{u \in [0,1]} \varphi(u)$ *and we may take* $u^* = \frac{1}{2}$.

Proof. Let u' be a maximizer of $\varphi(u)$. It might be equal to u^*, or if there is more than one maximizer, it might differ. Note that $\frac{1}{2} = \frac{1}{2}u' + \frac{1}{2}(1-u')$. φ is concave ([6] Chapter 2, Theorem 2.7.4), implying $\varphi(\frac{1}{2}) \geq \frac{1}{2}\varphi(u') + \frac{1}{2}\varphi(1-u')$. Now, using $\varphi(u') = \varphi(1-u')$, we obtain $\varphi(\frac{1}{2}) \geq \varphi(u')$, which concludes the proof.

A few simple channels with expressions for u^* and C are presented in Table 1. In the multivariate normal case, one can directly check that $\varphi(1-u) = \varphi(u)$, $0 \leq u \leq 1$ and conclude that $u^* = \frac{1}{2}$ using Proposition 1.

4. 1-Dimensional Optimal Policies

We now present two specific policies in $d = 1$ dimension that satisfy the sufficient conditions for optimality given in Theorem 2: the probabilistic bisection policy, and the dyadic policy. After defining these two policies in Sections 4.1 and 4.2, we study the sequence of entropies $(H(p_n) : n \geq 1)$ that they generate, focusing on the dyadic policy. In addition to Theorem 2, which shows that $E^\pi[H(p_n)] = H(p_0) - nC$ for any optimal policy π, the analysis of the dyadic policy in Sec. 4.2 provides a strong law of

Channel	Model	Channel Capacity	u^*
Binary Symmetric	$\begin{array}{c\Vert cc} & 0 & 1 \\ \hline f_0 & 1-\epsilon & \epsilon \\ f_1 & \epsilon & 1-\epsilon \end{array}$	$1-h(\epsilon)$	$\frac{1}{2}$
Binary Eraser	$\begin{array}{c\Vert ccc} & 0 & 1 & e \\ \hline f_0 & 1-\epsilon & 0 & \epsilon \\ f_1 & 0 & 1-\epsilon & \epsilon \end{array}$	$1-\epsilon$	$\frac{1}{2}$
Z	$\begin{array}{c\Vert cc} & 0 & 1 \\ \hline f_0 & 1 & 0 \\ f_1 & \epsilon & 1-\epsilon \end{array}$	$h(u^*(1-\epsilon)) - u^*h(\epsilon)$	$\frac{1/(1-\epsilon)}{1+e^{h(\epsilon)/(1-\epsilon)}}$
Multivariate Normal Symmetric	$f_0 \sim N(m_0, \Sigma)$ $f_1 \sim N(m_1, \Sigma)$	Not analytical	$\frac{1}{2}$

Table 1: Channel capacity, and the value u^* at which the channel capacity is achieved

large numbers and a central limit theorem for $H(p_n)$. In further analysis of the dyadic policy, Sec. 4.3 analyzes the number of size-limited noise-free questions required to find X^* after noisy questioning with the dyadic policy ceases, which is a metric important in the application discussed in Sec. 6.

To support the analysis in Sections 4.1 and 4.2, we first give here a general expression for the one-step change in entropy, $H(p_{n+1}) - H(p_n)$, under any policy π satisfying $p_n(A_n) = u^*$. First, we define two densities:

$$p_n^+(x) = \begin{cases} \frac{p_n(x)}{u^*}, \text{ if } x \in A_n, \\ 0, \text{ if } x \in \bar{A}_n, \end{cases} \qquad p_n^-(x) = \begin{cases} \frac{p_n(x)}{1-u^*}, \text{ if } x \in \bar{A}_n, \\ 0, \text{ if } x \in A_n, \end{cases}$$

where $\bar{A}_n$ is the complement of A_n. Their entropies are respectively,

$$H(p_n^+) = \log u^* - \frac{1}{u^*}\int_{A_n} p_n(x)\log p_n(x)dx,$$
$$H(p_n^-) = \log(1-u^*) - \frac{1}{1-u^*}\int_{\bar{A}_n} p_n(x)\log p_n(x)dx,$$

and $H(p_n) = u^* H(p_n^+) + (1-u^*)H(p_n^-) + h(u^*)$, where $h(u^*)$ is the entropy of a Bernouilli random variable with parameter u^*, denoted $B(u^*)$. Using Lemma 1, for a given observation $Y_{n+1} = y$, we have

$$\begin{aligned}
H(p_{n+1}) =& \log Z - p_{n+1}(A_n)\log f_1(y) - p_{n+1}(\bar{A}_n)\log f_0(y) \\
& - \frac{1}{Z} f_1(y)\int_{A_n} p_n(x)\log p_n(x)dx - \frac{1}{Z} f_0(y)\int_{\bar{A}_n} p_n(x)\log p_n(x)dx \\
=& \log Z - \frac{1}{Z} u f_1(y)\log f_1(y) - \frac{1}{Z}(1-u)f_0(y)\log f_0(y) \\
& - \frac{1}{Z}u^* f_1(y)(\log u^* - H(p_n^+)) - \frac{1}{Z}(1-u^*)f_0(y)(\log(1-u^*) - H(p_n^-)).
\end{aligned}$$

Expanding and rearranging, we obtain,

$$\begin{aligned}
H(p_{n+1}) - H(p_n) =& -D\left(B\left(\frac{u^* f_1(y)}{Z}\right), B(u^*)\right) \\
& + \frac{u(1-u^*)}{Z}(f_1(y) - f_0(y))(H(p_n^+) - \log u^* - H(p_n^-) + \log(1-u^*)), \quad (12)
\end{aligned}$$

where D is the Kullback-Leibler divergence.

Under an optimal policy, the density of Y_{n+1} is the mixture of densities $u^* f_1 + (1-u^*)f_0$ according to Lemma 1, and the random variables $Y_1, Y_2, \ldots$ are i.i.d.

4.1. Probabilistic Bisection Policy

We first consider the case when questions are limited to intervals $A = (-\infty, a)$, $a \in \mathbb{R}$. This limitation appears naturally in applications such as stochastic root-finding [23] and signal estimation [5]. In this case, an optimal policy consists of choosing a_n such that $\int_{-\infty}^{a_n} p_n(x)dx = u^*$. Such an a_n always exists but is not necessarily unique.

When the model is symmetric, $u^* = \frac{1}{2}$, and a_n is the median of p_n. This policy of measuring at the median of the posterior is the probabilistic bisection policy introduced by [13]. Thus, the optimal policy with interval questions and general channels is a generalization of the probabilistic bisection policy, and we continue to refer to it as the probabilistic bisection policy even when $u^* \neq \frac{1}{2}$.

We briefly consider the behavior of $(H(p_n) : n \geq 1)$ under the probabilistic bisection policy. We assume a binary symmetric channel with noise parameter ϵ. Recall that $u^* = \frac{1}{2}$ in this case, and

$$D\left(B\left(\frac{f_1(Y_{n+1})}{2Z}\right), B\left(\frac{1}{2}\right)\right) = 1 - h(\epsilon).$$

Moreover,

$$H(p_{n+1}) - H(p_n) = h(\epsilon) - 1 + \left(\frac{1}{2} - \epsilon\right) W_{n+1}(H(p_n^+) - H(p_n^-)),$$

where the W_n are i.i.d Rademacher random variables. In this situation, even when p_0 is the density of the uniform distribution over the interval $[0, 1]$, the behavior of the process $H(p_n)$ can be complicated. A simulation of $H(p_n)$ is presented in Fig. 1. The variance of $H(p_n)$ increases with n at a high rate. This high degree of variation may be disadvantageous in some applications, and we do not pursue the probabilistic bisection policy further in this paper.

4.2. Dyadic Policy

Consider now the situation where all sets in $\mathbb{A}_n$ are available as questions, and p_0 is piecewise constant with finite support. Let $I = \{I_k : k = 0, \ldots, K-1\}$ be a finite partition of the support of p_0 into intervals such that p_0 is constant and strictly positive in each of these intervals. We assume that each interval I_k is closed on the left and open on the right, so $I_k = [a_k, b_k)$ with $a_k \in \mathbb{R}$ and $b_k \in \mathbb{R}$. This assumption is without loss of generality, because if it is not met, we can alter the prior density p_0 on a set of Lebesgue measure 0 (which does not change the corresponding prior probability measure) to meet it. We also assume that the constants J_n used to construct $\mathbb{A}_n$ satisfy $J_n \geq 2^{n+1}K$. If this restriction is not met, then we are free to increase J_n in most applications.

For each $k = 0, \ldots, K-1$ we partition I_k into two intervals, $A_{0,2k}$ and $A_{0,2k+1}$, as follows:

$$A_{0,2k} = [a_{0,2k}, b_{0,2k}) = [a_k, a_k + u^*(b_k - a_k)),$$
$$A_{0,2k+1} = [a_{0,2k+1}, b_{0,2k+1}) = [a_k + u^*(b_k - a_k), b_k).$$

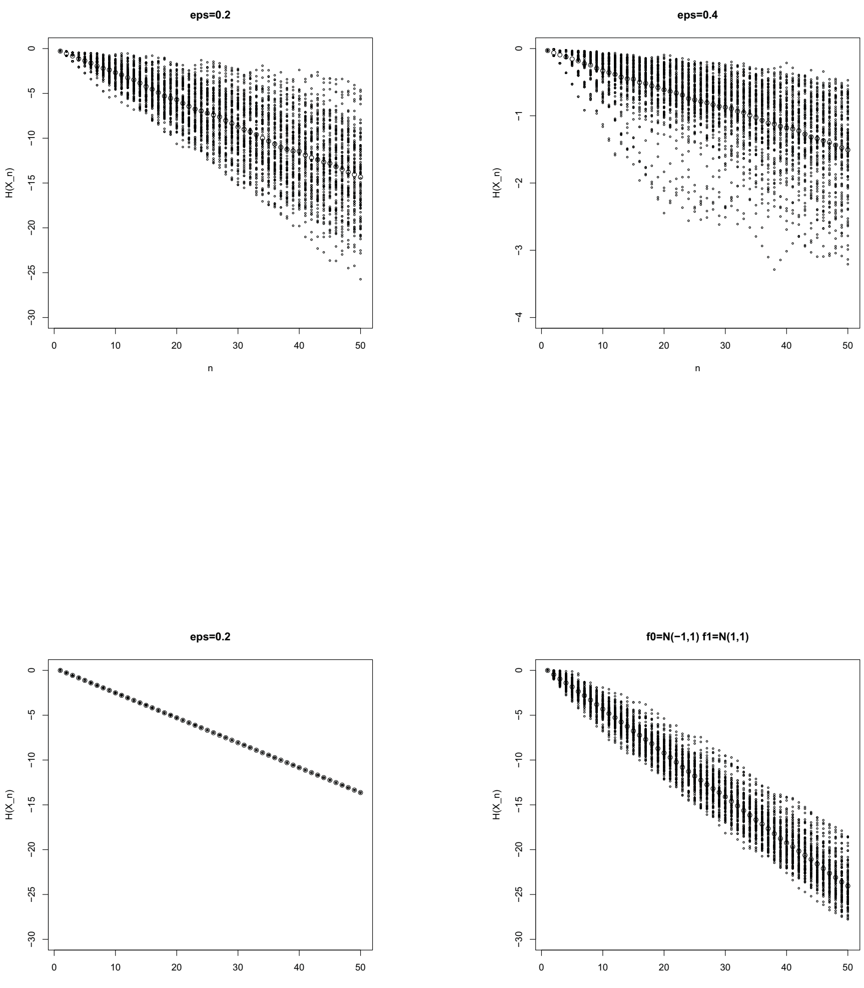


FIGURE 1: The process $H(p_n)$ for the binary symmetric channel. **Top:** The questions are the intervals $(-\infty, \text{Median}(p_n)]$. p_0 is Uniform([0, 1]). **Top Left:** $\epsilon = 0.2$, $C = 0.28$. **Top Right:** $\epsilon = 0.4$, $C = 0.03$. **Bottom:** The questions are chosen according to the dyadic policy. **Bottom Left:** binary symmetric channel $\epsilon = 0.2$ **Bottom Right:** Normal channel. $f_0 \sim N(-1, 1)$, $f_1 \sim N(1, 1)$. $C = 0.47$.

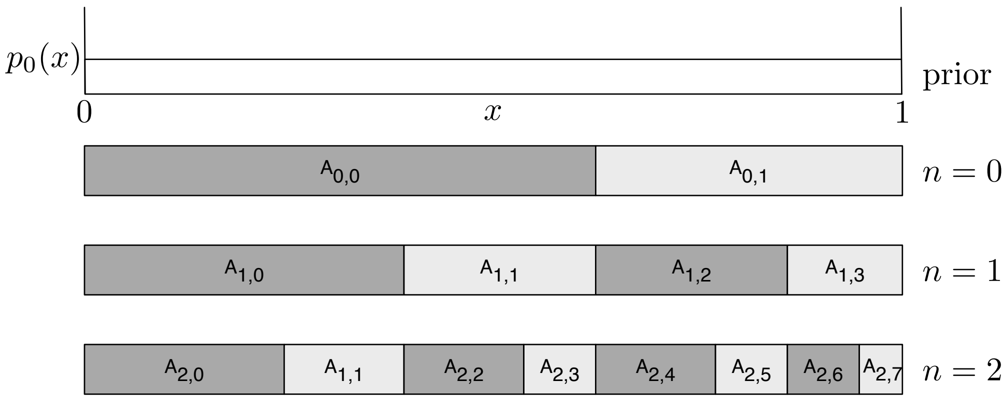


FIGURE 2: Illustration of the dyadic policy when p_0 is uniform on $[0, 1]$ and $u^* = 5/8$. The prior is displayed on top. Below, the sets $A_{n,k}$ are illustrated for $n = 0, 1, 2$. Each question A_n is the union of the dark grey subsets $A_{n,k}$ for that value of n.

With this partition, the mass $p_0(A_{0,2k}) = u^*\, p_0(I_k)$. The question asked at time 0 is

$$A_0 = \bigcup_{k=0}^{K-1} A_{0,2k},$$

and $p_0(A_0) = u^*$.

We use a similar procedure recursively for each $n = 0, 1, \ldots$ to partition each $A_{n,k}$ into two intervals, $A_{n+1,2k}$ and $A_{n+1,2k+1}$, and then construct the question A_{n+1} from these partitions. Let $K_n = 2^{n+1}K$ and for $k = 0, \ldots, K_n - 1$ define

$$\begin{aligned} A_{n+1,2k} &= [a_{n+1,2k}, b_{n+1,2k}) = [a_{n,k}, a_{n,k} + u^*(b_{n,k} - a_{n,k})), \\ A_{n+1,2k+1} &= [a_{n+1,2k+1}, b_{n+1,2k+1}) = [a_{n,k} + u^*(b_{n,k} - a_{n,k}), b_{n,k}). \end{aligned}$$

Then from these, we define the question to be asked at time $n + 1$,

$$A_{n+1} = \bigcup_{k=0}^{K_n-1} A_{n+1,2k}.$$

This construction is illustrated in Fig. 2.

Observe $p_{n+1}(A_{n+1,2k}) = u^*\, p_{n+1}(A_{n,k})$ implies $p_{n+1}(A_{n+1}) = \sum_{k=0}^{K_n-1} u^*\, p_{n+1}(A_{n,k}) = u^*$ because $\{A_{n,k} : k = 0, \ldots, K_n - 1\}$ is a partition of the support of p_0. Thus, this construction satisfies $p_n(A_{n+1}) = u^*$, and is optimal. In addition, the sets $A_0, \ldots, A_{n-1}$ are constructed without knowledge of the responses, and thus this policy is non-adaptive. This is useful in applications allowing multiple questions to be asked simultaneously.

We call this policy the dyadic policy because each question is constructed by dividing the previous question's intervals into two pieces.

We now provide an analysis that leads to a law of large numbers and a central limit theorem for $H(p_n)$ under this policy when n is large. Under the dyadic policy, we have

$$H(p_n^+) = H(p_n) + \log u^* \text{ and } H(p_n^-) = H(p_n) + \log(1-u^*),$$

which implies, using (12), that

$$H(p_{n+1}) - H(p_n) = -D\left(B\left(\frac{u^* f_1(Y_{n+1})}{u^* f_1(Y_{n+1}) + (1-u^*) f_0(Y_{n+1})}\right), B(u^*)\right), \tag{13}$$

where Y_n is, as already stated, a sequence of i.i.d random variables with density the mixture $u^* f_1 + (1-u^*) f_0$. We read from (13) that $H(p_n)$ is, in this case, a sum of i.i.d random variables. Moreover, each one is bounded above and below. Indeed,

$$\begin{aligned} 0 &\le D\left(B\left(\frac{u^* f_1(Y_{n+1})}{u^* f_1(Y_{n+1}) + (1-u^*) f_0(Y_{n+1})}\right), B(u^*)\right) \\ &\le \max(D(B(0), B(u^*)), D(B(1), B(u^*))), \end{aligned}$$

implying the bound

$$\min(\log(u^*), \log(1-u^*)) \le H(p_{n+1}) - H(p_n) \le 0. \tag{14}$$

This proves the following theorem.

Theorem 3. *For any piecewise constant p_0, using the dyadic policy,*

$$\lim_{n\to\infty} \frac{H(p_n)}{n} = -C \ \ a.s. \tag{15}$$

and

$$\lim_{n\to\infty} \frac{H(p_n) + nC}{\sqrt{n}} \stackrel{\mathrm{D}}{=} N(0, \sigma^2), \tag{16}$$

where σ^2 is the variance of the increment $H(p_{n+1}) - H(p_n)$, and can be computed from the distribution given in (13)*. A degenerate situation occurs for the binary symmetric channel with noise ϵ. In this case, the sequence $H(p_n) = H(p_0) - nC$ is constant.*

The dyadic policy is illustrated in the bottom graphs of Fig 1. $H(p_n)$ is plotted as a function of n. The binary symmetric channel model with $\epsilon = 0.2$ is shown on the bottom left. The sequence $H(p_n)$ is constant, in sharp contrast with the behavior of $H(p_n)$ for the same model under the probabilistic bisection policy, shown on the top left of the same figure. Finally, a Normal channel is presented on the bottom right.

4.3. Expected number of noise-free questions

Throughout this article we have proposed optimal policies in order to reduce the expected entropy of the posterior distribution over X^*. It remains to measure how this will permit us to better estimate X^*, which often is the practical goal.

In this section, we investigate a setting addressing this issue. Suppose that, in addition to the noisy questions previously discussed, we also have the ability to ask a noise-free oracle whether X^* lies in a given set, where the sets about which we can ask noise-free questions come from some restricted class, e.g., their size is below a threshold. This situation occurs in the example considered in Sec. 6, where the sets about which we can ask noise-free questions correspond to pixels in an image. We suppose that after a fixed number N of noisy questions, we query sets using the noise-free questions until we find X^*. The loss function that arises naturally in this situation is the expected number of noise-free questions until X^* is found.

Given a posterior p_N that results from the first stage of noisy questions, the optimal way in which to ask the noise-free questions is to first sort the available sets about which noise-free questions can be asked, in decreasing order of their probability of containing X^* under p_N. Then, these sets should be queried in this order until X^* is found. Note that observing that X^* is not in a particular set alters the probability of the other sets, but does not change the order of these probabilities. Thus it is sufficient to ask the noise-free questions in an order that depends only upon p_N, and no subsequent information.

We consider the binary symmetric channel with a uniform p_0, and give an explicit expression for the expected number of noise-free questions required after the dyadic policy completes. We assume that the sets about which we can ask noise-free questions evenly subdivide each interval $A_{N-1,k}$, for $k = 0, \ldots, 2^N - 1$. That is, each interval $A_{N-1,k}$ has some fixed number ℓ of equally sized sets about which we can ask noise-free questions. We refer here to $A_{N-1,k}$ more simply as B_k.

Each B_k has some corresponding number M_k of questions to which the oracle has responded that $X^* \in B_k$, either because $B_k \subseteq A_n$ and $Y_n = 1$, or because $B_k \subseteq I \setminus A_n$ and $Y_n = 0$.

Each time the oracle indicates that $X^* \in B_k$, we multiply the posterior density on

B_k by $2(1-\epsilon)$, and each time the oracle indicates that $X^* \notin B_k$ we multiply by 2ϵ. Since the prior density was $|I|^{-1}$, the posterior density on B_k after all N measurements is

$$|I|^{-1} 2^N (1-\epsilon)^{M_k} \epsilon^{N-M_k}.$$

Since M_k depends on the oracle's responses, it is random. However, for each $m \in \{0, \ldots, N\}$, the number of k with $M_k = m$ is deterministic and is equal to $\binom{N}{m}$. This is shown in the following proposition.

Proposition 2. *For fixed N and any $m \in \{0, \ldots, N\}$, the number of sets B_k with $M_k = m$ is $\binom{N}{m}$.*

Proof. Fix N. For each $k \in \{0, \ldots, 2^N - 1\}$, let $b_{kn} = \mathbb{I}\{B_k \subseteq A_n\}$, and define the binary sequence $b_k = (b_{k1}, \ldots, b_{kN})$. By construction of the sets B_k, each b_k is unique. Since there are 2^N possible binary sequences of N bits, and 2^N sets B_k, the mapping between B_k and b_k is a bijection.

Consider a sequence of responses from the oracle, $Y_1, \ldots, Y_N$. For each b_k define a subset $D_k = \{n \in \{1, \ldots, N\} : b_{kn} = Y_n\}$. Each b_k defines a unique subset D_k. Since there are 2^N subsets and 2^N sequences b_k, each subset $D \subseteq \{1, \ldots, N\}$ is equal to some D_k. Thus, the mapping between b_k and D_k is a bijection.

Because $M_k = |D_k|$, the number of k with $M_k = m$ is equal to the number of subsets $D \subseteq \{1, \ldots, N\}$ of size m. This number is exactly $\binom{N}{m}$.

Thus, the number of B_k with any given posterior density $|I|^{-1} 2^N (1-\epsilon)^m \epsilon^{N-m}$ is deterministic. Because the expected number of noise-free questions required to find X^* depends only upon the posterior probability density after sorting, this quantity is also deterministic. Fig. 3(a) shows this sorted probability distribution for particular values of N and ϵ.

The expected number of questions before finding X^* in this procedure can be calculated as follows. We first observe that, if we have a collection of disjoint subsets $C_1, \ldots, C_K$, each with equal probability $1/K$ of containing X^*, and we query each subset in order of increasing index until we find X^*, then we ask k questions when $X^* \in C_k$ and the expected number of questions asked is

$$\sum_{k=1}^{K} k P\{X^* \in C_k\} = \sum_{k=1}^{K} k/K = (K+1)/2.$$

We now observe that X^* has probability

$$\binom{N}{m}(1-\epsilon)^m \epsilon^{N-m}$$

of being in a subset B_k with $M_k = m$ (this is because there are $\binom{N}{m}$ such intervals, and each has size $2^{-N}|I|$). Recalling that the number of available questions within a set B_k is notated ℓ, the expected number of noise-free questions, conditioned on X^* being in a subset B_k with $M_k = m$, is

$$\frac{\binom{N}{m}\ell + 1}{2} + \sum_{m'=m+1}^{N} \binom{N}{m'} \ell,$$

where the first term is the number of questions asked in subsets with $M_k = m$, and the second term is the number asked in subsets with $M_k > m$ (these subsets had a strictly higher density $p_N(x)$, and were queried earlier).

Thus, the expected total number of noise-free questions is

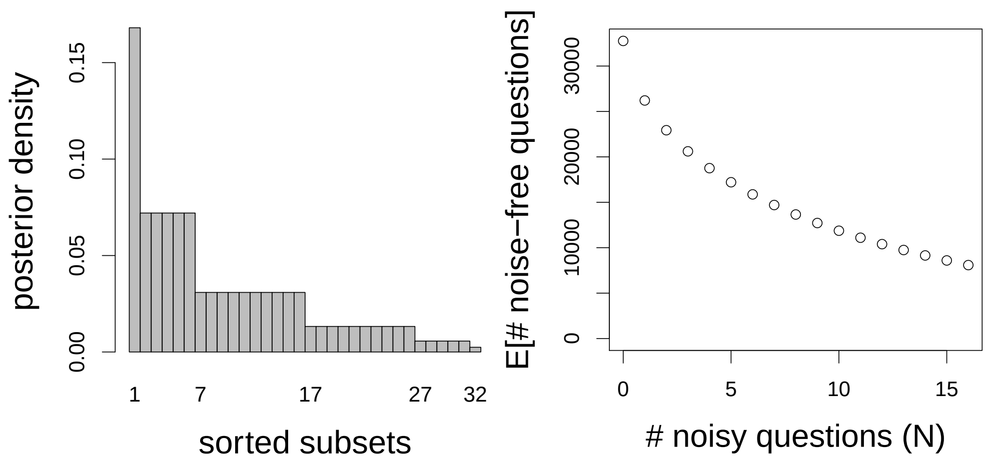
$$\sum_{m=0}^{N} \binom{N}{m}(1-\epsilon)^m \epsilon^{N-m} \left[\frac{\binom{N}{m} + 1/\ell}{2} + \sum_{m'=m+1}^{N} \binom{N}{m'} \right] \ell. \tag{17}$$

Using this expression, we may consider the effect of varying N. Suppose one has a fixed collections of sets about which noise-free questions may be asked, as in the example in Sec. 6 where these sets correspond to pixels in an image. Take $I = [0,1]$ and suppose each pixel is of size 2^{-L} and occupies a region $[k2^{-L}, (k+1)2^{-L}]$ for some $k = 0, \ldots, 2^L$. If sets B_k must contain integer numbers of pixels, then we may naturally consider any N between 0 and L. For any such N, the number of pixels ℓ in a subset B_k is $\ell = 2^{L-N}$.

From the expression (17) one can then compute the expected number of noise-free questions that will need to be queried as a function of N. This is shown in Fig. 3(b) for $L = 16$ and $\epsilon = 0.3$. The figure shows a dramatic decrease in the expected number of noise-free questions as the number of noisy questions increases.

5. Optimal Policies in 2 Dimensions with Entropy Loss

We now consider the case $d = 2$, in which X^* is a two-dimensional random variable, $X^* = (X_1^*, X_2^*)$, with joint density p_0. To minimize the expected entropy $E[H(p_N)]$ of



(a) (b)

FIGURE 3: (a) The posterior density p_N for the binary symmetric channel with the dyadic policy, with subsets $A_{k,N}$ sorted in order of decreasing posterior density $p_N(x)$, and $N = 5$. (b) The expected number of noise-free questions as a function of N, for a fixed collection of 2^{16} subsets about which noise-free questions may be asked. In both panels, $\epsilon = 0.3$.

the two-dimensional posterior distribution on X^* at time N, Theorem 2 from Section 3 shows it is optimal to use any policy satisfying $p_n(A_n) = u^*$.

While the objective function $E[H(p_N)]$ is natural in $d = 1$ dimension, it has a drawback in $d = 2$ and higher dimensions. This is well illustrated using an example. Assume that X_1^* and X_2^* are independent and uniformly distributed over intervals of lengths s_1 and s_2 respectively. Then $H(p) = \log(s_1 s_2)$. In this case, $H(p)$ can be arbitrarily small even if the entropy of one of the marginal densities remains large, *e.g.* $s_2 = 1$.

This leads us to consider objective functions without this drawback. For example, we might wish to solve $\inf_\pi E^\pi\left[\max\left(H_1(p_N), H_2(p_N)\right)\right]$ where $H_1(p_N) = H(\int p_N(\,\cdot\,, u_2)\, du_2)$ and $H_2(p_N) = H(\int p_N(u_1, \cdot\,)\, du_1)$ are the entropies of the marginals. However, solving this problem directly seems out of reach. Instead, we focus on reducing $E^\pi[\max\left(H_1(p_N), H_2(p_N)\right)]$ at an asymptotically optimal rate by solving

$$V(p) = \inf_\pi \liminf_{N\to\infty} \frac{1}{N} E^\pi[\max\left(H_1(p_N), H_2(p_N)\right)|p_0 = p]. \tag{18}$$

We use the lim inf to include policies for which the limit might not exist. Throughout this section, we assume that both $H_1(p_0)$ and $H_2(p_0)$ are finite.

For further simplification, we assume that questions concern only one coordinate. That is, the sets queried are either of type 1, $A_n = B \times \mathbb{R}$ where B is a finite union of intervals of $\mathbb{R}$, or alternatively of type 2, $A_n = \mathbb{R} \times B$. In each case, we assume that the response passes through a memoryless noisy channel with densities $f_0^{(1)}$ and $f_1^{(1)}$ for questions of type 1, and $f_0^{(2)}$ and $f_1^{(2)}$ for questions of type 2. We also assume that p_0 is a product of its marginals. This guarantees that p_n for all $n > 0$ remains a product of its marginals and that only one marginal distribution is modified at each point in time. This is shown by the following lemma.

Lemma 2. *Assume $p_n(u_1, u_2) = p_n^{(1)}(u_1)\, p_n^{(2)}(u_2)$ and we choose a question of type 1 with $A_n = B \times \mathbb{R}$. Then, given $Y_{n+1} = y$,*

$$p_{n+1}(u_1, u_2) = \frac{1}{\mathcal{Z}_1} \left(f_1^{(1)}(y) \mathbb{1}_{\{u_1 \in B\}} + f_0^{(1)}(y) \mathbb{1}_{\{u_1 \notin B\}} \right) p_n^{(1)}(u_1)\, p_n^{(2)}(u_2),$$

where $\mathcal{Z}_1 = P\left(Y_{n+1} = y \mid p_n, A_n = B \times \mathbb{R}\right) = f_1^{(1)}(y)\, p_n^{(1)}(B) + f_0^{(1)}(y)(1 - p_n^{(1)}(B))$. Similarly, if we choose a question of type 2 with $A_n = \mathbb{R} \times B$ then

$$p_{n+1}(u_1, u_2) = \frac{1}{\mathcal{Z}_2} \left(f_1^{(2)}(y) \mathbb{1}_{\{u_2 \in B\}} + f_0^{(2)}(y) \mathbb{1}_{\{u_2 \notin B\}} \right) p_n^{(2)}(u_2)\, p_n^{(1)}(u_1)$$

where $\mathcal{Z}_2 = P\left(Y_{n+1} = y \mid p_n, A_n = B \times \mathbb{R}\right) = f_1^{(2)}(y)\, p_n^{(2)}(B) + f_0^{(2)}(y)(1 - p_n^{(2)}(B))$.

Proof. The proof is straightforward using Bayes formula, and is similar to the proof of Lemma 1 from the 1-dimensional case.

In the 2-dimensional setting, any policy can be understood as making two decisions at each time n. The first decision is which coordinate to query, that is, whether to ask a question of type 1 or type 2. Given this choice, the second decision is which question of this type to ask, which corresponds to a finite union of intervals of $\mathbb{R}$. As before, these decisions may depend only upon the information gathered by time n, for which the corresponding sigma-algebra is $\mathcal{F}_n$. For $N > 0$, let S_N be the number of questions of type 1 answered by time N. That is, S_N is the number of $n \in \{0, \ldots, N-1\}$ such that A_n is of the form $A_n = B \times \mathbb{R}$. We take $S_0 = 0$.

We first present a lower bound on the expected decrease in the entropy of each marginal posterior distribution.

Lemma 3. *Under any valid policy* π*,*

$$E^\pi[H_1(p_n)] \geq H_1(p_0) - C_1 E^\pi[S_n],$$
$$E^\pi[H_2(p_n)] \geq H_2(p_0) - C_2(n - E^\pi[S_n]).$$

Proof. Define $M_n^{(1)} = H_1(p_n) + C_1 S_n$ and $M_n^{(2)} = H_2(p_n) + C_2(n - S_n)$. We will show that $M^{(1)}$ and $M^{(2)}$ are sub-martingales. Focusing first on $M^{(1)}$, we calculate,

$$E^\pi[M_{n+1}^{(1)} \mid \mathcal{F}_n] = E^\pi[H_1(p_{n+1}) \mid \mathcal{F}_n] + C_1 S_{n+1}$$

since S_{n+1} is $\mathcal{F}_n$-measurable. We consider two cases. First, if $S_{n+1} = S_n$ (which occurs if A_n is of type 2) then $H_1(p_{n+1}) = H_1(p_n)$ and the $\mathcal{F}_n$-measurability of $H_1(p_n)$ implies $E^\pi[M_{n+1}^{(1)} \mid \mathcal{F}_n] = M_n^{(1)}$. Second, if $S_n^{n+1} = S_n + 1$ (which occurs if A_n is of type 1), then Theorem 2 implies

$$E^\pi[H_1(p_{n+1}) \mid \mathcal{F}_n] \geq H_1(p_n) - C_1.$$

Hence,

$$E^\pi[M_{n+1}^{(1)} \mid \mathcal{F}_n] \geq C_1(S_n + 1) + H_1(p_n) - C_1 = M_n^{(1)},$$

which shows that $M_n^{(1)}$ is a sub-martingale. The proof is similar for $M_n^{(2)}$.

Now, because $M_n^{(1)}$ is a sub-martingale, $E^\pi[M_n^{(1)}] \geq M_0^{(1)}$, which implies $E^\pi[H_1(p_n)] \geq H_1(p_0) - C_1 E^\pi[S_n]$. Proceeding simlarly for $M_n^{(2)}$ concludes the proof.

Consider the following policy, notated π^*. At step n, choose the type of question at random, choosing type 1 with probability $\frac{C_2}{C_1+C_2}$ and type 2 with probability $\frac{C_1}{C_1+C_2}$. Then, in the dimension chosen, choose the subset to be queried according to the 1-dimensional dyadic policy.

We show below in Theorem 4 that π^* is optimal for the objective function (18). Before presenting this result, which is the main result of this section, we present an intermediate result concerning the limiting behavior of π^*. This intermediate result is essentially a strong law of large numbers for the objective function (18).

Lemma 4. *Let*

$$T_N = \frac{1}{N} \max(H_1(p_N), H_2(p_N)),$$

Under π^**, as* $N \to \infty$*,*

$$T_N \to -\frac{C_1 C_2}{C_1 + C_2} \ a.s. \tag{19}$$

Moreover there is a constant K such that $|T_N| < K$ for all N.

Proof. Recall that S_N is the number of questions of type 1 answered by time N, so $S_N/N \to C_2/(C_1 + C_2)$ a.s. The law of large numbers established in (15) for the one-dimensional posterior shows $H_1(p_N)/S_N \to -C_1$ a.s. Combining these two facts shows $H_1(p_N)/N \to -C_1C_2/(C_1 + C_2)$ a.s. By a similar argument, $H_2(p_N)/N \to -C_1C_2/(C_1 + C_2)$ a.s., which shows (19).

We now show the bound on $|T_N|$. Using π^*, according to (14),

$$H_1(p_N) = H_1(p_0) + \sum_{n=1}^{N} Z_n,$$

where Z_n are independent bounded random variables and $|Z_n| \leq |\min(\log(u), \log(1 - u))| = \beta$. As a consequence, for any $N \geq 1$,

$$\left| \frac{H_1(p_N)}{N} \right| \leq |H_1(p_0)| + \beta.$$

The same is true for $H_2(p_N)$, which proves there is a constant K such that $|T_N| < K$.

We now present the main result of this section.

Theorem 4. *The policy π^* is optimal with respect to* (18)*. Moreover, the optimal value is, for any p_0, with $H(p_0) < \infty$,*

$$V(p_0) = -\frac{C_1C_2}{C_1 + C_2} \tag{20}$$

Proof. First we show that the value in (20) constitutes a lower bound for $V(p_0)$. Second, we show (20) is an upper bound on $V(p_0)$ using the properties of the policy π^* presented in the Lemma 4.

$$\begin{aligned}
V(p_0) &\geq \inf_{\pi} \liminf_{N\to\infty} \frac{1}{N} \max(E^{\pi}[H_1(p_N)], (E^{\pi}[H_1(p_N)]) \\
&\geq \inf_{\pi} \liminf_{N\to\infty} \frac{1}{N} \max(H_1(p_0) - E[S_N]C_1, H_2(p_0) - (N - E[S_N])C_2) \\
&= \inf_{0\leq a\leq 1} \max(-aC_1, -(1-a)C_2) \\
&= -\frac{C_1C_2}{C_1 + C_2}
\end{aligned}$$

We obtain the first line using Jensen inequality, the second line using Lemma 3, the third line by choosing $a = \liminf_{n\to\infty} E[S_N]/N$ and the fourth line by recalling that $C_1 > 0$ and $C_2 > 0$.

Now, the other equality,

$$\begin{aligned} V(p_0) &\leq \liminf_{N\to\infty} E^{\pi^*}\left[\max\left(\frac{H_1(p_N)}{N}, \frac{H_2(p_N)}{N}\right)\right] \\ &= E^{\pi^*}\left[\max\left(\liminf_{N\to\infty}\frac{H_1(p_N)}{N}, \liminf_{N\to\infty}\frac{H_2(p_N)}{N}\right)\right] = -\frac{C_1C_2}{C_1+C_2} \end{aligned}$$

The uniform bound on T_N from Lemma 4 is sufficient to justify the exchange between the limit and the expected value in going from the first to the second line.

We remark as an aside that in the case where $C_1 = C_2$, this policy is also optimal for the value function (3) since it verifies (10).

We conclude this section by providing a central limit theorem for the objective under this policy π^*.

Theorem 5. *Under* π^*,

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\left[\max(H_1(p_n), H_2(p_n)) + \frac{C_1C_2}{C_1+C_2}n\right] \overset{\mathrm{D}}{=} \frac{\max\left(\sigma_1\sqrt{C_2}Z_1, \sigma_2\sqrt{C_1}Z_2\right)}{\sqrt{C_1+C_2}}. \tag{21}$$

Here, Z_1 *and* Z_2 *are independent standard normal random variables, and* σ_i^2 *is the variance of the increment of* $H_i(p_{n+1}) - H_i(p_n)$ *when measuring type* i*, whose distribution is given by* (13).

Proof. For $i = 1, 2$, let $S_{n,i}$ be the number of questions of type i answered by time n, so $S_{n,1} = S_n$ and $S_{n,2} = n - S_n$. Let $t_{s,i} = \inf\{n : S_{n,i} = s\}$ for $s = 0, 1, \ldots$. Then $t_{0,i} = 0$ and $\{t_{s,i} : s = 1, 2, \ldots\}$ are the times when questions of type i are answered. Thus, each stochastic process $\{H_i(p_{t_{s,i}}) : s = 0, 1, \ldots\}$ for $i = 1, 2$ has a distribution identical to that of the entropy of the one-dimensional posterior under the dyadic policy. In addition, the two stochastic processes are independent.

The central limit theorem established in (16) shows

$$\lim_{s\to\infty}\frac{H_i(p_{t_{s,i}}) + sC_i}{\sqrt{s}} \overset{\mathrm{D}}{=} \sigma_i Z_i,$$

where each Z_i is a standard normal random variable and Z_1 is independent of Z_2.

From the definition of $t_{s,i}$,

$$\lim_{s\to\infty}\frac{H_i(p_{t_{s,i}}) + sC_i}{\sqrt{s}} \overset{\mathrm{D}}{=} \lim_{n\to\infty}\frac{H_i(p_n) + S_{n,i}C_i}{\sqrt{S_{n,i}}}$$

Let $j = 1$ when $i = 2$, and $j = 2$ when $i = 1$. Then $\lim_{n\to\infty} S_{n,i}/n = C_j/(C_1 + C_2)$ a.s. and

$$\lim_{n\to\infty} \frac{H_i(p_n) + S_{n,i}C_i}{\sqrt{S_{n,i}}} \overset{\mathrm{D}}{=} \lim_{n\to\infty} \frac{H_i(p_n) + n\frac{C_1C_2}{C_1+C_2}}{\sqrt{n}} \sqrt{\frac{C_1 + C_2}{C_j}}.$$

These three facts imply,

$$\lim_{n\to\infty} \frac{H_i(p_n) + n\frac{C_1C_2}{C_1+C_2}}{\sqrt{n}} \overset{\mathrm{D}}{=} \sqrt{\frac{C_j}{C_1 + C_2}}\sigma_i Z_i.$$

This shows the expression (21) for the limit.

6. LaTeX Character Localization

In this section we present an application of the dyadic policy to a well-established problem in computer vision: object localization. While the probabilistic bisection policy has already been applied in computer vision, see [11, 26], the dyadic policy has not, and we feel that it offers considerable promise in this application area.

In the object localization problem, we are given an image and a known object, and must output parameters that describe the *pose* of the object in the image. In the simplest case, the pose is defined by a single pixel, but more complex cases can include, *e.g.* a rotation angle, a scale factor or a bounding box. Machine learning techniques have led to the development of classifiers that, given a specific pose, provide accurate answers to the binary question "Is the object in this pose?" Classifiers such as Support Vector Machines [27] and boosting [25] are combined with discriminant features, *e.g.* [19], to provide the most accurate algorithms, [29, 28]. To find the object's pose within an image, classifiers are evaluated at nearly every possible pose, which is computationally costly. We demonstrate that using the dyadic policy rather than this brute force approach considerably reduces this computational cost. Although a detailed comparison would be beyond the scope of the illustrative example we present here, the branch and bound algorithm used in [18] is an alternative methodology for reducing computational cost in object localization.

6.1. LaTeX Character Images, Noisy Queries, and Model Estimation

The task we consider is localizing a specific LaTeX character in a binary image. In this setting, an image is a binary matrix $I \in \{0, 1\}^{m\times m}$, where the image has m rows

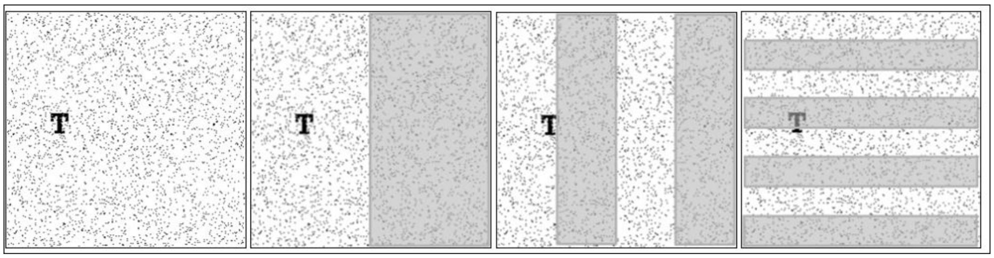


FIGURE 4: From left to right: Example of an image containing the character "T". Examples of subset-based questions. In each image, we show the queried region by the gray area, respectively A_2^1, A_2^2 and A_1^3.

and m columns. A LaTeX character is another smaller binary image $J \in \{0,1\}^{j \times j}$, where $j < m$. We present experiments where the character of interest, or pattern, is the letter "T". We assume that the pattern is always present in the image, and fully visible (*i.e.* not occluded by other objects or only partially visible in the image).The goal is to find the location $X^* = (X_1^*, X_2^*)$ of the pixel at the upper left corner of the pattern within the image.

We generated 1000 images, each of size 256×256 pixels. Each image has a black background (*i.e.* pixel values of zero), and contains a single fully visible "T" at a random location in the image. This "T" is a binary image of size 32×32 pixels (see Fig. 4(a)). Noise is added to the image by flipping each pixel value independently with probability 0.1. We then randomly assign each image into one of two sets of approximately equal size: one for training and the other for testing. The training set is used to learn the noise model as described below, and the testing set is used to evaluate the performance of the algorithm.

In this task, querying a set A corresponds to asking whether the upper left corner of the "T" resides in this set. We use a simple image-processing technique to provide a noisy answer to this question. The technique we use is chosen for its simplicity, and other more complex image-processing techniques might produce more informative responses, improving the overall performance of the algorithm.

In describing this technique, we first observe that all the images are of size 256×256 pixels and so any pixel coordinate can be represented in base 2 using two 8-bit strings, or octets. For example, the pixel with column-row location (32,14) is represented by

(00100000, 00001110). We define 16 sets of pixels. Let A_1^i, $i = 1\ldots, 8$ be the set of pixels whose column pixel coordinate has a 1 for its ith bit. Similarly, let A_2^i, $i = 1, \ldots, 8$ be the set of pixels whose row pixel coordinate has a 1 for its ith bit. Fig. 4 (b-d) show the sets A_1^1, A_1^2 and A_2^3, respectively. For any given image I and set A_j^i, we define the response

$$y(A_j^i) = \sum_{x \in A_j^i} I(x) - \sum_{x \notin A_j^i} I(x) \tag{22}$$

where $I(x) \in \{0, 1\}$ is the binary image's value at pixel x. The motivation for using the response defined by (22) is that $y(A_j^i)$ is more likely to be large when A_j^i contains the "T".

Although the response $y(A_j^i)$ is entirely determined by the image I and the location of the "T" within it, our algorithm models the response using a noise model of the form (1). For simplicity, we assume that both the density f_1 of $y(A)$ when A contains the "T", and the density f_0 of $y(A)$ when A does not contain the "T", are normal with respective distributions $N(\mu, \sigma^2)$ and $N(-\mu, \sigma^2)$. The training set is used to estimate these parameters, leading to $\mu = 64.76$ and $\sigma = 105.7$. Because the model is symmetric, $u^* = 0.5$. The channel capacity is estimated with Monte Carlo integration to be $C = 0.23$.

6.2. Prior, Posterior, and Algorithm

We let $X^* = (X_1^*, X_2^*)$, $X_1^* \in [0, 255]$ and $X_2^* \in [0, 255]$, with p_0 uniform over the domain of X^*. Since the sets A_j^i constrain only one coordinate, the posterior over X^* is a product distribution as was discussed in Sec. 5. The posterior for each coordinate $j = 1, 2$ was computed in Lemma 2. We now specialize to the model at hand using the notation $\propto$ to define equality up to a term that does not depend on x_j.

$$p_8^{(j)}(x_j) \propto \prod_{i=1}^{8} (f_1(y_j^i) \mathbb{1}_{x_j \in A_j^i} + f_0(y_j^i) \mathbb{1}_{x_j \notin A_j^i})$$
$$\log p_8^{(j)}(x_j) \propto \sum_{i: x_j \in A_j^i} \log \frac{f_1(y_j^i)}{f_0(y_j^i)} \propto \sum_{i: x_j \in A_j^i} y_j^i$$

The algorithm has two phases: (i) the noisy query phase; and (ii) the noise-free query phase. The noisy query phase comes first, and uses the dyadic policy to obtain a posterior distribution on X^*. The implementation of this noisy query phase is facilitated by the non-adaptive nature of the dyadic policy's questions, which allows us to compute the answers to the questions all at once. The noise-free query phase then uses the posterior resulting from the first phase, together with a sequence of size-limited noise-free questions, to determine the exact location of X^*.

Noisy Query Phase: Given an image I, we begin by computing $y(A_j^i) = y_j^i$, for each $j = 1, 2$, and $i = 1, \ldots, 8$. We then compute $\ell(x)$ for each pixel x, which is proportional to the logarithm of the posterior density at x,

$$\ell(x) = \sum_{i : x \in A_1^i} y_1^i + \sum_{i : x \in A_2^i} y_2^i.$$

Fig. 5(*top*) shows example images from our test set, while (*bottom*) shows the corresponding ℓ-images, in which the value of $\ell(x)$ is plotted for each pixel. Dark regions of the ℓ-image indicate pixels with large $\ell(x)$, which are more likely to contain the "T".

Noise-free Query Phase: We sort the pixels in decreasing order of $\ell(x)$. We then sequentially perform noise-free evaluations at each pixel x in this order until the true pixel location X^* is found. To perform a noise-free evaluation at a given pixel, we compare the "T" pattern with the 32×32 pixel square from the image with upper left corner at x to see if they match. When X^* is found, we stop and record the number of noise-free evaluations performed.

6.3. Results

We validated the algorithm above by evaluating it on the test set described in Sec. 6.1. To do this, we ran the algorithm on each image and recorded the number of noise-free evaluations required to locate the target character. The results described below (i) demonstrate that the dyadic policy significantly reduces the number of noise-free evaluations required to locate the "T" character, and (ii) allows us to visualize the results summarized in (10), (15) and (16) within the context of this application.

Recall that each image has $256 \times 256 = 65{,}536$ pixels. Over 500 test images, the mean, median and standard deviation of the number of noise-free evaluations are 2021.5, 647 and 4066.9, respectively. This corresponds to a speed-up factor of 15 over an exhaustive (and typical) search policy. Fig. 6(a) shows the sample distribution of the number of noise-free evaluations. We also computed the entropy of the posterior

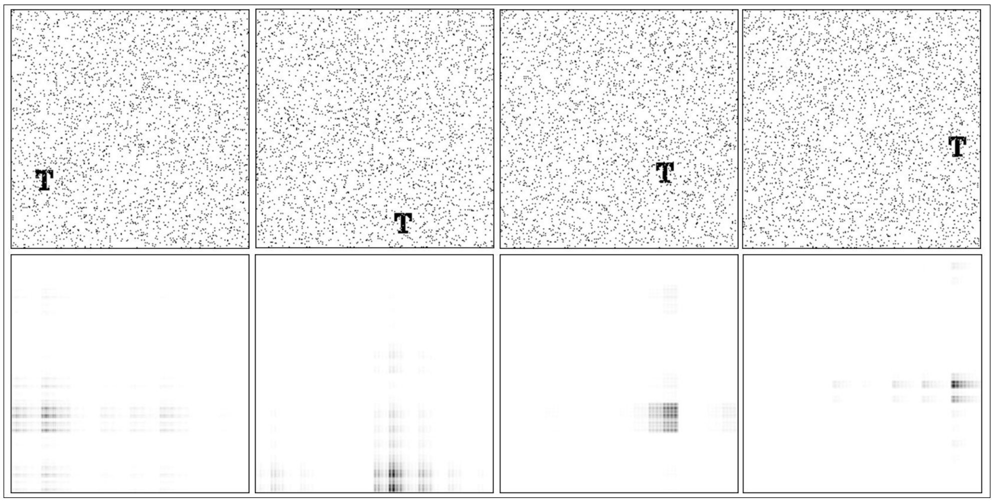
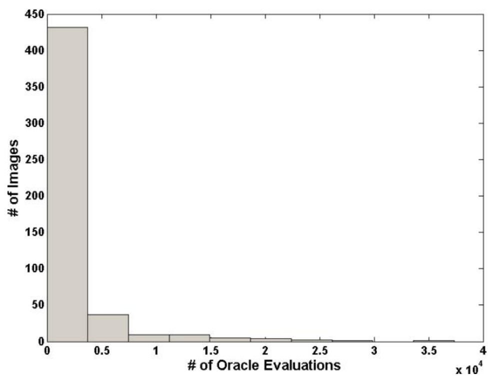


FIGURE 5: Pixel Reordering: (*top*) Example images from the test set. (*bottom*) Corresponding ℓ-images. Dark regions indicate pixels more likely to contain the character, while light regions are less likely.

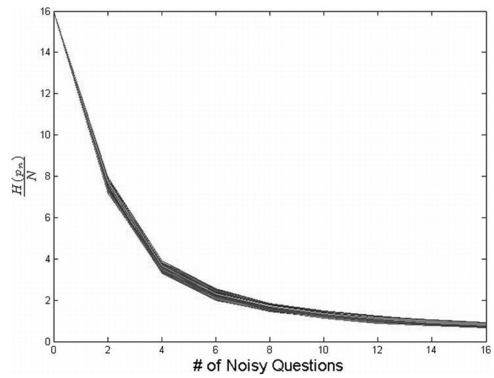
distribution after the 16 noisy questions are answered. According to (10), $E[H(p_{16})] = H(p_0) - 16C = 16 - 16(.23) = 12.32$, which is in agreement with the empirically observed value $E[H(p_{16})] = 12.3$ (with standard deviation 0.716). We also visualized the convergence of the entropy for *each* image, as predicted by the law of large numbers in (15). In Fig. 6(b), we plot $\frac{H(p_n)}{n}, n = 0, \ldots, 16$, for each image in our test set. The empirical variance at $n = 16$ is very small. Finally, according to (16), the distribution of $\frac{H(p_n)-(H(p_0)-nC)}{\sqrt{n}}$ should be approximately normal. Fig. 7(a) shows the histogram and (b) shows a normal Q-Q plot, demonstrating close agreement with the normal distribution.

7. Conclusion

We have considered the problem of 20 questions with noisy responses, which arises in stochastic search, stochastic optimization, computer vision, and other application areas. By considering the entropy as our objective function, we obtained sufficient conditions for Bayes optimality, which we then used to show optimality of two specific policies: probabilistic bisection and the dyadic policy. This probabilistic bisection policy generalizes a previously studied policy, while we believe that the dyadic policy



(a)



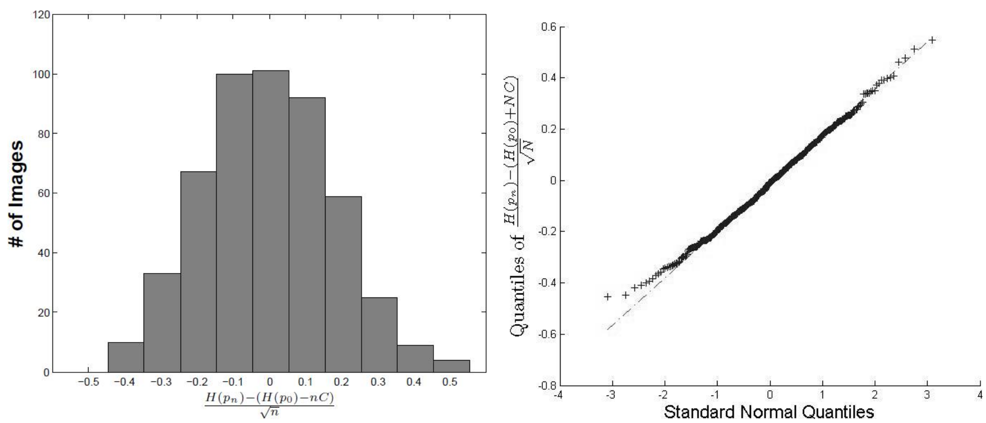
(b)

FIGURE 6: Noise-free evaluations and convergence in entropy. (a) The distribution of number of noise-free evaluations needed to locate the target character. (b) Plot of $H(p_n)/n$ as a function of n. Each line corresponds to one image, with $H(p_n)/n$ plotted over $n = 1, \ldots, 16$. $H(p_n)/n$ converges to 1-C.

has not been previously considered.

The dyadic policy asks a deterministic set of question, despite being optimal among fully sequential policies. This lends it to applications that allow multiple questions to be asked simultaneously. The structure of this policy also lends itself to further analysis. We provided a law of large numbers, a central limit theorem, and an analysis of the number of noise-free questions required after noisy questioning ceases. We also showed that a generalized version of the dyadic policy is asymptotically optimal in two dimensions for a more robust version of the entropy loss function. We then demonstrated the use of this policy on an example problem from computer vision.

A number of interesting and practically important questions present themselves for future work. First, our optimality results assume the entropy as the objective, but in many applications other objectives are more natural, e.g., the expected number of noise-free questions as in Section 4.3, or mean-squared error. Second, our results assume that noise is added by a memoryless transmission channel. In many applications, however, the structure of the noise depends upon the questions asked, which calls for generalizing the results herein to this more complex style of noise dependence. We feel that these and other questions will be fruitful areas for further study.



(a) (b)

FIGURE 7: Central Limit Theorem: (a) Distribution of $\frac{H(p_n)-(H(p_0)-nC)}{\sqrt{n}}$, with mean -0.01. The distribution is closely Gaussian as the Q-Q plot (b) shows.

References

[1] BEN-OR, M. AND HASSIDIM, A. (2008). The Bayesian Learner is Optimal for Noisy Binary Search. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*. IEEE. pp. 221–230.

[2] BERRY, D. AND FRISTEDT, B. (1985). *Bandit Problems: Sequential Allocation of Experiments*. Chapman & Hall, London.

[3] BLUM, J. (1954). Multidimensional stochastic approximation methods. *The Annals of Mathematical Statistics* **25,** 737–744.

[4] BURNASHEV, M. V. AND ZIGANGIROV, K. S. (1974). An interval estimation problem for controlled observations. *Problemy Peredachi Informatsii* **10,** 51–61. (originally in Russian).

[5] CASTRO, R. AND NOWAK, R. (2008). Active learning and sampling. *Foundations and Applications of Sensor Management* 177–200.

[6] COVER, T. M. AND THOMAS, J. A. (1991). *Elements of information theory.* Wiley-Interscience, New York, NY, USA.

[7] DEGROOT, M. H. (1970). *Optimal Statistical Decisions.* McGraw Hill, New York.

[8] Dynkin, E. and Yushkevich, A. (1979). *Controlled Markov Processes.* Springer, New York.

[9] Frazier, P., Powell, W. and Dayanik, S. (2008). A knowledge gradient policy for sequential information collection. *SIAM Journal on Control and Optimization* **47,**.

[10] Frazier, P. I. and Henderson, S. (2010). Optimal entropic bisection search in noisy environment. *Technical report.* Cornell University.

[11] Geman, D. and Jedynak, B. (1996). An active testing model for tracking roads in satellite images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **18,** 1–14.

[12] Gittins, J. (1989). *Multi-Armed Bandit Allocation Indices.* John Wiley and Sons, New York.

[13] Horstein, M. (1963). Sequential decoding using noiseless feedback. *IEEE Transactions on Information Theory* **9,** 136–143.

[14] Horstein, M. (2002). Sequential transmission using noiseless feedback. *IEEE Transactions on Information Theory* **9,** 136–143.

[15] Karp, R. and Kleinberg, R. (2007). Noisy binary search and its applications. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms.* SIAM, Philadelphia. pp. 881–890.

[16] Kushner, H. J. and Yin, G. G. (1997). *Stochastic Approximation Algorithms and Applications.* Springer-Verlag, New York.

[17] Lai, T. and Robbins, H. (1985). Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics* **6,** 4–22.

[18] Lampert, C. H., Blaschko, M. B. and Hofmann, T. (2009). Efficient subwindow search: A branch and bound framework for object localization. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31,** 2129–2142.

[19] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* **60,** 91–110.

[20] Pelc, A. (2002). Searching games with errors: fifty years of coping with liars. *Theoretical Computer Science* **270,** 71–109.

[21] Polyak, B. (1990). New method of stochastic approximation type. *Automation and Remote Control* **51,** 937–946.

[22] Robbins, H. (1952). Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society* **58,** 527–535.

[23] Robbins, H. and Monro, S. (1951). A stochastic approximation method. *Annals of Mathematical Statistics* **22,** 400–407.

[24] Ruppert, D. (1988). Efficient estimators from a slowly convergent robbins-monro procedure. Technical Report 781. School of Operations Research and Industrial Engineering, Cornell University.

[25] Schapire, R. E. (1990). The strength of weak learnability. *Machine Learning* **5,** 197–227–227.

[26] Sznitman, R. and Jedynak, B. (2010). Active testing for face detection and localization. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32,** 1914–1920.

[27] Vapnik, V. N. (1995). *The nature of statistical learning theory.* Springer-Verlag New York, Inc., New York, NY, USA.

[28] Vedaldi, A., Gulshan, V., Varma, M. and Zisserman, A. (2009). Multiple kernels for object detection. In *Proceedings of the International Conference on Computer Vision (ICCV).*

[29] Viola, P. and Jones, M. J. (2004). Robust real-time face detection. *International Journal of Computer Vision* **57,** 137–154.

[30] Whittle, P. (1981). Arm-acquiring bandits. *The Annals of Probability* **9,** 284–292.

[31] Whittle, P. (1988). Restless bandits: Activity allocation in a changing world. *Journal of Applied Probability* 287–298.