Dichotomous Optimistic Search to Quantify Human Perception

Julien Audiffren [1]

Abstract

In this paper we address a variant of the continuous multi-armed bandits problem, called the threshold estimation problem, which is at the heart of many psychometric experiments. Here, the objective is to estimate the sensitivity threshold for an unknown psychometric function Ψ, which is assumed to be non decreasing and continuous. Our algorithm, Dichotomous Optimistic Search (DOS), efficiently solves this task by taking inspiration from hierarchical multi-armed bandits and Black-box optimization. Compared to previous approaches, DOS is model free and only makes minimal assumption on Ψ smoothness, while having strong theoretical guarantees that compares favorably to recent methods from both Psychophysics and Global Optimization. We also empirically evaluate DOS and show that it significantly outperforms these methods, both in experiments that mimics the conduct of a psychometric experiment, and in tests with large pulls budgets that illustrate the faster convergence rate.

1. Introduction

Psychophysics investigates the connection between physical stimuli and the subjective responses (such as sensations, or perceptions) they produce. This field of research has widespread applications, including the study of attention (Scheuerman et al., 2017), and the evaluation of treatments for pain relief (Nir et al., 2011). One of the key aspect of Psychophysics is the evaluation of human perception, which is generally assessed by performing psychometric experiments. They unfold as follows : the experimenter presents to an individual, called the observer, a sequence of stimuli of varying intensities (for instance, the volume of a specific sound, see e.g. (Hirahara, 2004)), and try to measure how often the different intensities are perceived by the observer. In particular, many experiments are interested in measuring the *sensitivity threshold*, where the stimulus is just noticeable (Kontsevich & Tyler, 1999).

Human perception is generally modeled as follows. For any given stimulus intensity s, the stimulus is perceived by the observer with (unknown) probability μ_s, and the relation between s and μ_s for a given observer is called the psychometric function, noted Ψ, and is defined such that $\Psi(s) = \mu_s$ for any stimulus intensity s. By definition, Ψ is assumed to be non-decreasing (the stronger the stimulus, the easier it is to perceive it) and continuous (Leek, 2001). Given a target perception probability μ_*, the objective of these psychometric experiments is to estimate the sensitivity threshold s_* such that $\Psi(s_*) \approx \mu_*$. To achieve this, the experimenter must both choose a sequence of stimuli to present to the observer (which must be as short as possible, to limit the fatigue of the observer (Wichmann & Hill, 2001a)) and estimate s_* given the observer responses. In this work, this task will be referred to as the *threshold estimation problem*.

One of the most commonly used technique for solving this problem is the *constant stimuli* method (Wichmann & Hill, 2001a), where the observer is presented with a fixed sequence of stimuli, spanning the range of sensation from imperceptible to consistently perceptible. After collecting the observer responses, the parameters of Ψ are estimated using the maximum likelihood method – where the shape of Ψ is assumed to be known, e.g. a Gaussian cumulative distribution function (c.d.f.). Finally, s_* is estimated as $\Psi^{-1}(\mu_*)$. However, this approach suffers from many limitations. First, the parametric models have been shown to be frequently inconsistent with the observations, or to require many small empirical corrections to fit the data with an acceptable accuracy (Wichmann & Hill, 2001b). Second, the fixed sequence does not account for individual specificity, a key problem as it has been shown that the sensitivity threshold can vary by a factor of ten between individuals (Benson et al., 1989).

Consequently, there has been an increased interest in using adaptive algorithms (Leek, 2001), where an agent A) aims at estimating directly the sensitivity threshold without estimating Ψ, therefore reducing the difficulty of the problem and B) adapts the sequence of stimulus intensity based on the observer responses. One popular such method in

[1]Departement of Neuroscience, University of Fribourg, Fribourg, Switzerland. Correspondence to: Julien Audiffren <julien.audiffren@unifr.ch>.

Proceedings of the 38^{th} International Conference on Machine Learning, PMLR 139, 2021.

Psychophysics is arguably the staircase (and its iterations) (Cornsweet, 1962; Lengyel & Fiser, 2019). However these methods generally rely on strong assumptions about the smoothness and shape of Ψ, and seldomly provide theoretical guarantees on the the estimator they provide. Therefore, proposing a novel, model free and principled algorithm that addresses the threshold estimation problem may bring significant improvements to psychometric experiments.

Fortunately, in recent years multiple techniques have been developed for the stochastic adaptive optimization of black-box functions of unknown smoothness (see e.g. (Grill et al., 2015; Shang et al., 2019)). In particular, in hierarchical bandits methods, the agent uses various concentration inequalities to explore a hierarchical partition of the arm space, progressively narrowing the candidate subspace that may contain the maximum. While the task of locating the extremum of a noisy function is different from threshold estimation, the two problems present similarities with respect to the difficulties they encounter. Therefore, methods and algorithms developed in the former may give valuable hindsights into addressing the latter.

Following this idea, in this paper, we show that the threshold estimation problem can be rewritten as a pure exploration continuous multi-armed bandit problem, with interesting twists (Section 2). Then, we introduce the Dichotomous Optimistic Search algorithm (DOS), that takes inspiration from hierarchical bandits, and black box optimization to solve this problem (Section 4). The idea behind DOS is to perform a stochastic continuous binary search, while achieving the correct trade off between the depth of the binary tree, and the confidence in its noisy comparisons.

This method has multiple advantage over existing algorithms. First, DOS is model free, in the sense that it does not requires the knowledge of the shape of the psychometric function and only assumes that Ψ is mildly smooth around s_* (in fact, we show that local Hölder continuity is a sufficient condition). Moreover, the agent does not require the knowledge of the smoothness of Ψ, or the use of a suited hierarchical partition. Second, DOS has advantageous theoretical guarantees: we prove (see Theorem 1) that the simple regret $\mathcal{R}_T$ of DOS is upper bounded as follows

$$\mathbb{E}(\mathcal{R}_T) \leq \mathcal{O}\left(\sqrt{\frac{(\log T)^2(\log\log T)}{T}}\right).$$

This highlights the advantage of DOS over existing methods used in psychometric experiments, which to the best of our knowledge, do not have similar guarantee except in narrow settings. Moreover, this upper bound compares favorably to state of the art results, such as the simple regret bounds of POO (Grill et al., 2015), as it is independent of the local near-optimality dimension (a measure of the optimization problem difficulty, see e.g. (Shang et al., 2019)).

Third, we also extensively evaluate DOS in a wide range of experiments (Section 5). We show that our agent significantly outperforms adaptive psychometric methods and recent global optimization methods in both experiments that mimic the conduct of a psychometric experiment, and in tests with large pulls budgets that illustrate the faster convergence rate of our agent.

2. The Threshold Estimation Problem

Notation. Let T denote the time horizon (i.e. the maximum number of stimulus presented during the experiment), $\mathbb{I} \subset \mathbb{R}$ the (closed) interval of possible stimuli, $\Psi : \mathbb{I} \mapsto [0,1]$ the psychometric function, $\mu_* \in [0,1]$ the target probability, $s_* \doteq \Psi^{-1}(\mu_*)$ the sensitivity threshold. Finally, let $\mu_{\min} = \inf_{s\in\mathbb{I}} \Psi(s)$ and $\mu_{\max} = \sup_{s\in\mathbb{I}} \Psi(s)$.

In the following we assume without any loss of generality that $\mathbb{I} = [0,1]$. We also assume that the target threshold is strictly reachable, i.e. $\mu_{\min} < \mu_* < \mu_{\max}$. Due to the nature of the task (detecting stimuli of various intensity), the psychometric function is commonly assumed to be continuous and strictly increasing (see e.g. (Leek, 2001)).

The objective of the threshold estimation problem is to find an estimator $\hat{s}$ of the sensitivity threshold s_* with at most T stimuli. $\mathbb{I}$, T and μ_* are known to the agent (here the experimenter), but Ψ and s_* are not. The process unfolds as follows. For each round $t \in [1,\ldots,T]$, first the agent pulls an arm (i.e. chooses an intensity) $s \in \mathbb{I}$ and then the environment (here the observer) draws an independent Bernoulli random variable with mean $\Psi(s)$, and communicates the result to the agent, representing the detection of the stimulus. At time $t = T$, the agent returns the arm $\hat{s}$ that is her best guess for the target stimulus s_*. The performance of the agent is then evaluated using the simple regret $\mathcal{R}$, defined as

$$\mathcal{R}(\hat{s}) = |\mu_* - \Psi(\hat{s})|. \tag{1}$$

Note that (1) is similar to the definition of simple regret in hierarchical bandits (Valko et al., 2013). The relation between the two is discussed later in this section.

Remark 1 (Lapses and Guessing). Due to the subjective nature of perception, in general $\mu_{\min} > 0$ and $\mu_{\max} < 1$, even for completely undetectable (resp. unmissable) stimuli (Wichmann & Hill, 2001a). Indeed, $\mu_{\min}$ is identified as the guess rate, i.e. the chance that the observer correctly guesses the answer independently of the stimulus. Similarly $1 - \mu_{\max}$ is called the lapse rate and represents the probability of the observer missing the stimulus due to factors external to the experiment (such as blinking for a visual stimulus). These values are generally unknown at the beginning of the experiment, and most methods in psychophysics require the use of heuristics to estimate $\mu_{\min}$ and $\mu_{\max}$ (Wichmann & Hill, 2001a), before using these estimate to *normalize* the data. Importantly, this is not the case for DOS, which does not require any prior information on $\mu_{\min}$ and $\mu_{\max}$.

In the rest of this paper, we make the following mild assumption on the smoothness of Ψ, which is required to prove DOS theoretical guarantees.

Assumption 1 (Ψ is smooth around s_*)**.** There exists $\nu > 0$, and $0 < \rho < 1$ such that $\forall h > 0$, $\forall s \in \mathbb{I}$,

$$|s - s_*| \leq 2^{-h} \implies |\Psi(s) - \Psi(s_*)| \leq \nu\rho^h$$

This hypothesis implies that Ψ is smooth enough around s_*, and prevents the well known "find the needle in a haystack" problem of global optimization (Valko et al., 2013). This assumption is similar to the ones used in recent black-box optimization methods of function of unknown smoothness such as (Grill et al., 2015). Notably, Assumption 1 does not restrict the choice of possible Ψ, as shown by Lemma 1, whose proof can be found in the supplementary materials.

Lemma 1. *Let $\Psi : \mathbb{I} \mapsto [0, 1]$, and $\mathcal{V}$ be a neighborhood of s_* Then, $\forall \alpha > 0$*

$$\begin{array}{l}\Psi \textit{ is locally } \alpha\textit{−Hölder} \\ \textit{continuous on } \mathcal{V}\end{array} \implies \Psi \textit{ satisfies Assumption 1.}$$

Proof. By definition, $\exists r > 0$ such that $\mathcal{V}$ contains a ball $\mathbb{B}$ of radius $r > 0$ centered on s_*. Ψ is $\alpha-$Hölder on $\mathbb{B}$, thus $\exists C > 0$ such that $\forall s' \in \mathbb{B}$, $|\Psi(s) - \Psi(s')| \leq C|s - s'|^\alpha$, hence it is easy to see that Assumption 1 is satisfied on $\mathbb{B}$ for $\rho = 2^{-\alpha}$ and $\nu = C$. For $\mathbb{I} \setminus \mathbb{B}$, we just observe that

$$\forall s' \in \mathbb{I} \setminus \mathbb{B}, \quad |s - s'| > r = 2^{\frac{\log r}{\log 2}}.$$

and thus that Assumption 1 is satisfied on $\mathbb{I} \setminus \mathbb{B}$ for $\rho = 2^{-\alpha}$ and $\nu = 2\rho^{\frac{\log r}{\log 2}}$. Hence the conclusion. □

In particular, all continuously differentiable Ψ (and consequently all the commonly used psychometric functions) satisfy Assumption 1. Importantly, DOS does not require the knowledge of the smoothness parameters (ν, ρ).

Relation with Global Optimization Let Ψ be a psychometric function and μ_* a target probability. Define f as:

$$\begin{aligned} f: \quad & \mathbb{I} \mapsto [-1, 0] \\ & s \rightarrow -|\mu_* - \Psi(s)|. \end{aligned} \tag{2}$$

It is easy to see that f admits s_* as its unique maximum, and that $f(s_*) = 0$. Moreover, the regret defined by (1) is equivalent to the usual definition of simple regret for f (see e.g. (Bubeck et al., 2011)). Similarly, Assumption 1 implies a similar smoothness condition for f around its maximum. Therefore, (2) draws a link between the black box optimization of f and the threshold estimation of Ψ. However, since Ψ is unknown and can only be observed through the realizations of Bernoulli random variables, it is impossible to directly use global optimization strategies to solve the threshold optimization problem. Nevertheless, this transformation is useful to draw parallels between the two problems and their solutions.

3. Related Works

Threshold Estimation in Psychophysics. Several adaptive algorithms have proposed to solve the threshold estimation problem in psychometric experiments. On the one hand, the staircase algorithm, arguably the most popular adaptive method, has been discussed and improved upon significantly in recent years (Wichmann & Jäkel, 2018). However, this method can only be used for a very limited list of target probability (such as $\mu_* = 0.5$) (Brown, 1996), and convergence is only guaranteed for specific shape of the psychometric function (such as Gaussian c.d.f.) (Levitt, 1971). On the other hand, there has been increasing interest in parametric Bayesian adaptive algorithms. The purpose of these methods is generally to estimate the entire function Ψ, using a model based approach (e.g. Ψ is assumed to be a Gaussian c.d.f.). Notably, in (Kontsevich & Tyler, 1999) the authors proposed a method which aims at each step to minimize the entropy of the distribution of possible parameters for Ψ, while in (Shen & Richards, 2012), the authors introduced a sampling method that aims at minimizing the variance of each parameter. However, these methods also require the prior knowledge on the psychometric function shape, which significantly limits their applications, and (García-Pérez & Alcalá-Quintana, 2007; Hatzfeld et al., 2016) have empirically shown that all the aforementioned methods produce significantly worse estimations and might even diverge when this assumption is false. More recently, several works have proposed to use Gaussian Processes (GP) to approximate Ψ (Gardner et al., 2015a;b; Song et al., 2017). These algorithms are compatible with a much wider range of possible Ψ and can outperform other methods for some applications (see e.g. (Gardner et al., 2015a)). However these methods have a different, more general objective as they generally aim to estimate Ψ, and thus are more costly than methods that only approximate s_*. Moreover, they require the choice of a proper kernel and a mean function, as well as a grid of hyperparameters; and they perform best when these elements are hand crafted for the problem (Gardner et al., 2015a), using prior information regarding Ψ – and may perform poorly when these assumptions are false and when the kernel is ill-suited, see e.g. Section 5. Conversely, our method, DOS, is completely model free, a significant advantage when nothing is known about the psychometric function, a situation that occurs frequently in psychophysics research (see e.g. (Schütz et al., 2008)). Finally, and contrarily to the aforementioned methods, DOS has strong theoretical guarantees regarding its estimation of

s_* (see Theorem 1).

Global Optimization. To address the problem of black box optimization of an unknown function f in presence of noise, two families of solutions have been proposed. In the first, f is assumed to have some strong global smoothness, such as a Lipschitz condition (see e.g. the Lipschitz multi armed bandit problem (Kleinberg et al., 2008; 2019)). In the second category, f is only assumed to have some local smoothness around its maximum (see e.g. (Valko et al., 2013)). This framework leads to arguably more difficult problems, in particular when the smoothness is unknown. However it has been shown that even in this setting it is possible to achieve near optimal regret bounds, for instance by using a hierarchical bandit approach such as POO – Parallel Optimistic Optimization, see e.g. (Grill et al., 2015; Shang et al., 2019). This latest setting is the closest to our problem. Indeed, Assumption 1 can be seen as similar to their minimal assumption (see e.g. (Grill et al., 2015)) – but Assumption 1 is slightly weaker in the sense that in our case the agent does not have access to a hierarchical partition that is well suited for the function f. Importantly, in both cases the smoothness parameters (ν, ρ) are unknown. One other significant difference between the two settings is the non decreasing property of Ψ. There is no equivalent hypothesis in the global optimization setting; and this property of Ψ is key to DOS, our solution to the threshold optimization problem, and its significantly better regret bound (see Theorem 1).

Other related works. The threshold estimation problem shares some similarities with the noisy bissection problem (Chakraborty et al., 2011) and the learning the demand curve problem (Chhabra & Das, 2011). However, they are multiple crucial differences between these topics. For instance, in the later, the objective can be reformulated as minimizing the cumulative regret, instead of the simple regret (1) – leading to very different, non equivalent solutions (Bubeck et al., 2011). Additionally, in both cases strong assumptions are made on the properties of the noise (e.g. Gaussian, (Jedynak et al., 2012)), the shape of the function or its smoothness (Chakraborty et al., 2011) – which is very different from our model free approach, and cannot be easily adapted. Finally, the closest works to ours are arguably (Fontaine et al., 2020; Audiffren, 2021). The first has been developed simultaneously and independently, and while their algorithm also uses repeated pull of a candidate arm to obtain high probability comparison, it significantly differs from ours as : A) their algorithm was designed to optimize the cumulative regret, and the simple regret bound that can be derived from their work is significantly worse than ours, and B) their method cannot be applied to our problem (and conversely), as they rely on different, non equivalent hypotheses as well as different types of feedback (noisy gradient in their case). The second is an extended abstract that discusses general ideas related to DOS, but does not provide any theoretical analysis or regret bound and only limited empirical evaluation, which are main contributions of this work.

4. Contributions

In this section, we introduce DOS and discuss its key ideas before proving an upper bound for its simple regret which compares favorably to the existing regret bounds in global optimization. We provide sketches of proof for the different results – the detailed proofs can be found in the supplementary materials.

4.1. DOS

In the following, we use $\log^2(T) \doteq \log(\log(T))$. Let κ denote the number of different arms that are pulled by DOS during the attributed time budget T. Since there is a continuous set of possible arms, and $\kappa \leq T$, most arms will never get pulled; and in the following we say that the agent activates an arm when she pulls it for the first time. For any $1 \leq i \leq \kappa$, we use s_i (resp. $N_i(t)$, $\hat{\mu}_i(t)$ and μ_i) to denote the stimulus value (resp. the number of pulls, the empirical average and the true probability value) associated to the i-th activated arm at time t. Finally, let $\Delta_i = |\mu_i - \mu_*|$ – i.e. the regret obtained by the agent if she chooses to return the arm i when $t = T$.

DOS strategy. The general idea of DOS is inspired by the deterministic dichotomous search algorithm : the agent aims to produce a sequence of intervals $\mathbb{I}_1, \ldots, \mathbb{I}_\kappa \subset \mathbb{I}$ such that

$$\forall 1 \leq i \leq \kappa, \quad |\mathbb{I}_i| \leq 2^{-i} \text{ and } s_* \in \mathbb{I}_i \tag{3}$$

Note that if (3) is true, then $|s_\kappa - s_*| \leq 2^{-\kappa}$ and Assumption 1 implies that

$$\Delta_* \leq \nu\rho^\kappa \tag{4}$$

in other words, the sequence s_κ (resp μ_κ) converges exponentially fast toward s_* (resp μ_*). To produce this sequence, DOS proceeds as follows. To obtain $\mathbb{I}_{i+1}$ given the interval $\mathbb{I}_i$, the agent activates $s_{i+1} = \frac{2k_i+1}{2^{i+1}}$ – the arm located at the center of $\mathbb{I}_i$ – and repeatedly pulls this new arm, until the time budget is elapsed ($t = T$) or one of the two possible new arm activation criteria is satisfied. Then she compares μ_*, the target probability, and $\hat{\mu}_{i+1}(N_{i+1})$, i.e. the empirical proportion of stimuli of intensity s_{i+1} that were detected. Depending on the result, the agent defines the next interval $\mathbb{I}_{i+1}$ in the sequence as :

$$\mathbb{I}_{i+1} = \begin{cases} \left[\dfrac{2k_i}{2^{i+1}}, \dfrac{2k_i+1}{2^{i+1}}\right] & \text{if } \mu_* < \hat{\mu}_{i+1}, \\ \left[\dfrac{2k_i+1}{2^{i+1}}, \dfrac{2k_i+2}{2^{i+1}}\right] & \text{otherwise.} \end{cases}$$

Algorithm 1 DOS

Parameters μ_* (objective), T (time horizon)
Initialization $i \leftarrow 1$ (current arm), $s_1 \leftarrow 1/2$ (current stimulus), $N_1 \leftarrow 0$ (number of pulls of s_1), $\hat{\mu}_1 \leftarrow 0$ (empirical average of s_1), $t \leftarrow 0$ (total pulls), $\mathcal{S} = \text{NULL}$ the latest promising arm, N_* as in (8) and $\mathcal{B}_T(\cdot)$ as in (7).
Main Loop
While $t \leq T$:
 If $|\mu_* - \hat{\mu}_i(t)| > 2\mathcal{B}_T(N_i(t))$ **or** $N_i(t) > N_*$:
  If $N_i(t) > N_*$ **Then** $\mathcal{S} \leftarrow i$ **EndIf**
  Activate new arm: $i \leftarrow i + 1$ and

$$s_i \leftarrow \begin{cases} s_{i-1} + (1/2^i) & \text{if } \mu_* > \hat{\mu}_{i-1} \\ s_{i-1} - (1/2^i) & \text{if } \mu_* \leq \hat{\mu}_{i-1} \end{cases}$$

 EndIf
 Sample arm s_i, update $t, N_i, \hat{\mu}_i$
EndWhile
Output: s_{i_*}, where $i_* = \begin{cases} \mathcal{S} & \text{if } \mathcal{S} \neq \text{NULL}, \\ \kappa & \text{otherwise.} \end{cases}$

Here the agent leverages the fact that Ψ is monotonically increasing. Importantly, this strategy presents key differences with the deterministic dichotomous search, which are discussed below. The pseudocode for DOS can be found in Algorithm 1. Note that DOS *is completely model-free*, in the sense that it does not uses any parameter or prior knowledge regarding Ψ, including (ν, ρ), the parameters of Ψ local smoothness.

DOS noisy comparisons. Contrarily to the deterministic setting, here the agent has only access to noisy observations of $\Psi(s_i)$. Therefore, for any arm s_i the agent can only compare $\hat{\mu}_i$ and μ_*, and can never be sure if $\Psi(s_i) \geq \mu_*$. To quantify this uncertainty, DOS uses a Hoeffding-Chernoff concentration bound (see e.g. (Auer et al., 2007)):

$$\mathbb{P}\Big(|\mu_i - \hat{\mu}_i(t)| > \varepsilon\Big) \leq 2\exp\Big(-2N_i(t)\varepsilon^2\Big). \tag{5}$$

Let $\mathcal{Q}_i$ be the event where DOS reaches the wrong conclusion about the position of the arm s_i with respect to s_* :

$$\mathcal{Q}_i = \begin{cases} \big\{\hat{\mu}_i(T) < \mu_*\big\} & \text{if } \mu_i \geq \mu_*, \\ \big\{\hat{\mu}_i(T) \geq \mu_*\big\} & \text{if } \mu_i < \mu_*, \end{cases} \tag{6}$$

and let $q_i \doteq \mathbb{P}(\mathcal{Q}_i)$. While q_i can be reduced by pulling the arm s_i multiple times – and thus reducing the confidence interval – the number of pulls required increases drastically as the distance Δ_i decreases. This is compounded by (4), as Δ_i can be expected to decrease exponentially as i increases.

Meanwhile, decreasing the uncertainty of the comparisons comes at a cost on the number of activated arms, as more time is spent on each arm – and (4) gives insight into the link between the number of activated arms κ and the quality of the arm s_κ. Hence, the agent must achieve a proper trade-off between two opposite objective:

- **Confidence**: Pull each activated arm more to increase confidence in the comparison between $\hat{\mu}$ and μ_*,
- **Depth**: Increase the number of activated arm to improve the bound for Δ_κ provided by (4).

Moreover, identifying the correct trade-off between the two objectives is complicated by the fact that while Ψ is assumed to satisfy Assumption 1, the parameters ν and ρ are unknown to the agent. This raise additional difficulties as these parameters are crucial to the quality of the upper regret bound (4) and to assess the behavior of Ψ around s_*. This conundrum is discussed below.

Activation criteria. DOS uses two different activation rules to achieve the proper trade-off between Confidence and Depth. These two rules both rely on (5), but with different perspectives. The first rule forces the activation of a new arm if

$$|\mu_* - \hat{\mu}_i(t)| > \mathcal{B}_T(N_i(t)) \doteq \frac{3}{2}\sqrt{\frac{\log(T)}{N_i(t)}}. \tag{7}$$

Note that (7) is a confidence interval commonly used for the optimism against uncertainty principle, see e.g. (Auer et al., 2007). If (7) is achieved, then the agent is considered confident enough to activate the next arm, regardless of the number of pulls N_i, as stated by the following Lemma:

Lemma 2. *If* (7) *is satisfied for the arm* s_i*, then* $q_i < \dfrac{2}{T^3}$.

Proof. Without any loss of generality, assume that $\hat{\mu}_i > \mu_*$. Using the second triangular inequality $\mu_i - \mu_* > \hat{\mu}_i - \mu_* - |\hat{\mu}_i - \mu_i|$, hence the conclusion by using (5) and (7). □

However, the number of pull required to achieve (7) can be too large, in particular if Δ_i is small. Hence the second rule plays an important role in achieving the aforementioned exploration trade-off, by setting a maximum number of pulls for the arm i before the activation of the next arm. Indeed, this rules forces the activation of a new arm if

$$N_i(t) > N_* \doteq \left\lfloor \frac{T}{(\log T)(\log^2 T)} \right\rfloor. \tag{8}$$

Therefore, (8) provides a lower bound on the depth of the search: independently of observed results, DOS activates at least $\kappa = \lfloor(\log T)(\log^2 T)\rfloor$ arms. Moreover, it can be shown (see Lemma 3) that if (8) occurs, then s_i is a *promising arm*, i.e. Δ_i is small enough for s_i to be a good estimator of s_*. These arms are key to DOS estimation of the sensitivity threshold, as discussed below.

Lemma 3. *If* (7) *is false, but* (8) *is satisfied, then with probability at least* $1 - \frac{2}{T^3}$*,*

$$\Delta_i < 3\sqrt{\frac{(\log T)^2(\log^2 T)}{T}}$$

Proof. Note that since (7) is not satisfied for $N_i = N_*$, we have $|\mu_* - \hat{\mu}_i(t)| < \mathcal{B}_T(N_*)$. Hence, using the triangular inequality, $\Delta_i = |\mu_* - \mu_i| \leq |\mu_* - \hat{\mu}_i| + |\mu_i - \hat{\mu}_i|$. The result is then obtained by using the concentration inequality (5) and the definition of N_* (8). □

DOS final output. When the time horizon is reached ($t = T$), there are two possible scenarios. In the first case, the activation rule (8) was used at least once, and the agent found at least one promising arm. Then the agent returns the last encountered promising arm, and the regret incurred is controlled by Lemma 3. In the second case, no promising arm was found during the exploration process and all arms were activated using (7). In this scenario, the agent returns the last activated arm s_κ. To upper bound Δ_κ, let

$$\mathcal{A}^* \doteq \{\forall t \leq T, \forall i \leq \kappa,\, |\mu_i - \hat{\mu}_i(t)| \leq \mathcal{B}_T(N_i)\}\,.$$

In other words, $\mathcal{A}^*$ is the event where the empirical average of all activated arms are concentrated around their true mean. First note that on $\mathcal{A}^*$, if an arm is activated using (7), then the algorithm necessarily the right conclusion when comparing $\hat{\mu}_i$ and μ_* – and thus (3) is true. Thus, in this scenario, the event $\mathcal{A}^*$ has a probability close to one, as stated by the following Lemma.

Lemma 4. $\mathbb{P}(\mathcal{A}^*) \geq 1 - \frac{2}{T}$.

Proof. First note that on $\mathcal{A}^*$, DOS always reach the right conclusion when comparing $\hat{\mu}_i$ and μ_* after the activation rule (7). Thus in this scenario the sequence of arms s_i is fixed (but not their respective number of pulls) and the result is then obtained by taking the union bound on all arms and on all times $t < T$. □

Since (3) is true, then (4) is also true and the regret of s_κ is directly related to the number of activated arms κ. The following proposition proves a lower bound for κ.

Proposition 1. *Let* i_* *defined as*

$$i_* \doteq \left\lfloor \log\left(\frac{T}{(\log^2 T)(\log T)^2}\right)\frac{1}{2\log(1/\rho)}\right\rfloor. \quad (9)$$

Then all the following properties are true

(A) if $\log^2 T > -1/\log\rho$ *and* $T > 16$*, then* $i_* \leq \kappa$

(B) $\forall i > i_*$, $\mathcal{A}^*$ *a.s.* $\Delta_i \leq \nu\sqrt{\frac{(\log^2 T)(\log T)^2}{T}}$,

Proof. (A) is proved by using the fact that the definition of N_* (8) implies that $\kappa \geq \log(T)\log^2(T)$. (B) is derived from (4) by using the definition of i_* (9). □

In other words, (B) upper bounds DOS regret provided that the agent has activated at least i_* arms, and (A) states that for T large enough, more than i_* arms are activated. By combining all the previous results, it is possible to prove the following upper bound on the regret incurred by DOS:

Theorem 1 (Upper Bound on simple regret)**.** *Assume that Hypothesis 1 is true. Then,* $\forall T > 0$*, the simple regret of DOS* $\mathcal{R}_T$ *is upper bounded by*

$$\mathbb{E}(\mathcal{R}_T) \leq (3+\nu)\sqrt{\frac{(\log T)^2\log^2(T)}{T}}. \quad (10)$$

Proof. This results from noting that by definition $\mathcal{R}_T \leq 1$, and then using Lemma 3, Lemma 4 and Proposition 1. □

Note that (10) does not depend on ρ, i.e. the bound is uniform for any value of $\rho < 1$. This is a very important property, as ρ is directly linked with the difficulty of the problem. Indeed, for a value of ρ close to one, the set of arms that are both A) close enough to require very large number of comparisons to be eliminated and B) not close enough to be a sufficiently good estimator of the sensitivity threshold may be very large. This problem is related to the notion of near optimality dimension, discussed below.

Comparison with POO regret bound. Using the transformation described in (2), it is interesting to compare (10) to the upper regret bound for POO, which achieves state of the art performance in black box optimization problems (Shang et al., 2019). POO regret bound relies on the notion of near optimality dimension (Grill et al., 2015), for which we provide below an equivalent definition in the threshold estimation setting.

Definition 1 (Near optimality dimension)**.** Let $\nu > 0$ and $0 < \rho < 1$. The near optimality dimension of Ψ, noted $d(\nu, \rho)$, is defined as

$$d(\nu,\rho) \doteq \inf\Big\{d' \in \mathbb{R}^+ : \exists C, h > 0, \\ \Psi^{-1}(\mu_* + 2\nu\rho^h) - \Psi^{-1}(\mu_* - 2\nu\rho^h) \leq C(2\rho^{d'})^{-h}\Big\}$$

Intuitively, the $d(\nu, \rho)$ represents the measure of the size of the near optimal set; the larger, the more candidates for the optimal arm. The following regret bound has been shown for POO (Grill et al., 2015; Shang et al., 2019):

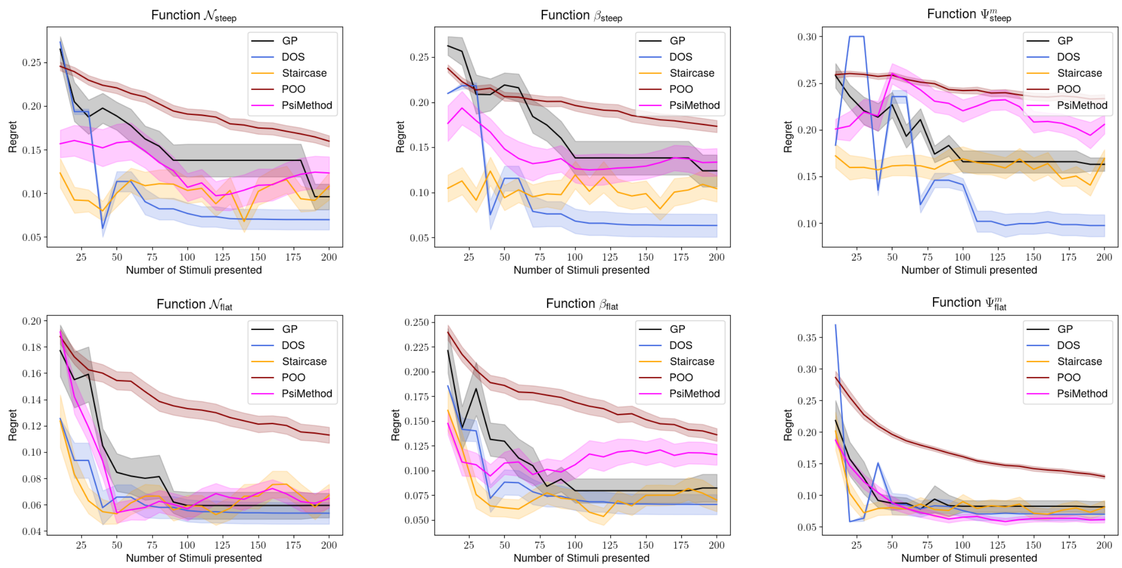

Figure 1. Comparison of the evolution of the average regret over 100 runs as a function of the number of stimuli presented to the observer, for a time horizon of $T = 200$, for each psychometric function. The standard deviation is reported using the shaded area.

Theorem 2 (Theorem 1 from (Grill et al., 2015))**.** *Let* $\rho_{\max} < 1$, $\nu_{\max} \in \mathbb{R}$. *Then* $\exists \mathcal{K} > 0 \in \mathbb{R}$, *such that* $\forall \rho \leq \rho_{\max}, \forall \nu \leq \nu_{\max}$,

$$\mathbb{E}(\mathcal{R}_T) \leq \mathcal{K}\left[\left(\frac{\log^2(T)}{T}\right)^{\frac{1}{2+d(\nu,\rho)}}\right] \tag{11}$$

Note that while (10) has an additional $\sqrt{\log^2 T}$ term, (11) is worse than (10) in the threshold estimation problem for two reasons. First, the constant $\mathcal{K}$ in (11) depends on $\rho_{\max}$, and $\mathcal{K} \to +\infty$ when $\rho_{\max} \to 1$. Therefore, (11) is not uniform on ρ, and requires additional information on Ψ smoothness to be used. Second, the rate of convergence of (11) is strictly worse than (10) for any Ψ that have a strictly positive near optimality dimension. The following proposition shows the existence of such function in the threshold estimation problem, highlighting the advantage of DOS.

Proposition 2 (Non zero optimality dimension)**.** *There exist psychometric functions* Ψ *satisfying Assumption 1 such that* $\min_{\nu,\rho} d_\Psi(\nu, \rho) > \log 2 > 0$.

5. Experiments

We now evaluate the performance of DOS in three different settings, with multiple psychometric functions. First, we use a small time budget ($T = 200$) – this aims at reproducing the constraints of real psychometric experiments, where only a few hundred stimuli can be presented to the observer before the fatigue and learning effects significantly interfere with the experiment (Wichmann & Hill, 2001a). Second, we take interest in DOS performance for large values of T, to illustrate its faster convergence rate. The final set of experiments, described below, aim at illustrating the advantages and limitations of the different procedure with very small time budget ($T = 50$ and 100). All experiments were performed with custom script using Python 3.7, unless mentioned otherwise.

Baselines. We compare DOS to three methods used in Psychophysics : *Staircase*, arguably the most commonly used methods in Psychophysics (see e.g. (Lengyel & Fiser, 2019)), *Psimethod* (Kontsevich & Tyler, 1999) is a popular Bayesian algorithm that assumes that Ψ is e.g. a Gaussian c.d.f., and *GP*, a recent algorithm for psychometrics experiments based on Gaussian Processes – see e.g.(Gardner et al., 2015a; Song et al., 2017). We also compared DOS to the black box optimization algorithm POO (Parallel Optimistic Optimization)(Grill et al., 2015), using (2) to transform the threshold estimation problem. Finally, in the last round of experiment, we compare also DOS to Quest+ (Watson, 2017), a recent algorithm that improves over Psimethod and is particularly efficient for small time budgets. For Staircase, PsiMethod and Quest+, we used the implementation provided by (Peirce et al., 2019), and the parameters recommanded by their respective papers. We used the dichotomous partition of [0,1] for POO and DOS, and ran GP with the kernel and hyperparameters recommended by (Song et al., 2017).

Psychometric Functions We used six different psychometric functions to assess the behavior of DOS. They can be roughly divided into two families, “flat” and “steep”,

Table 1. Average ($\pm$ standard deviation) over 100 runs of the simple regret, for each method and Psychometric functions with time horizon $T = 500$, 2000 or 5000. The reported regret has been multiplied by 10 for readability purpose. The best results are written in bold.

T	Ψ	GP	DOS	Staircase	POO	PsiMethod
500	steep $\mathcal{N}$	1.01 ($\pm$0.68)	**0.48** ($\pm$0.33)	0.81 ($\pm$0.70)	1.30 ($\pm$0.23)	1.52 ($\pm$1.20)
	steep β	0.85 ($\pm$0.56)	**0.43** ($\pm$0.36)	0.98 ($\pm$0.76)	1.24 ($\pm$0.25)	1.29 ($\pm$0.85)
	steep Ψ_m	0.80 ($\pm$0.69)	**0.25** ($\pm$0.26)	0.86 ($\pm$0.90)	1.17 ($\pm$0.27)	0.62 ($\pm$0.45)
	flat $\mathcal{N}$	0.54 ($\pm$0.47)	**0.46** ($\pm$0.36)	0.52 ($\pm$0.44)	0.86 ($\pm$0.23)	0.61 ($\pm$0.29)
	flat β	1.02 ($\pm$0.84)	**0.49** ($\pm$0.38)	0.67 ($\pm$0.54)	1.06 ($\pm$0.23)	1.12 ($\pm$0.35)
	flat Ψ_m	0.84 ($\pm$1.30)	**0.35** ($\pm$0.35)	0.65 ($\pm$0.64)	0.88 ($\pm$0.18)	0.58 ($\pm$0.13)
2000	steep $\mathcal{N}$	0.85 ($\pm$0.95)	**0.31** ($\pm$0.26)	1.05 ($\pm$0.77)	1.04 ($\pm$0.30)	1.09 ($\pm$0.87)
	steep β	0.78 ($\pm$0.94)	**0.40** ($\pm$0.31)	1.09 ($\pm$0.71)	0.98 ($\pm$0.21)	1.22 ($\pm$0.97)
	steep Ψ_m	0.89 ($\pm$0.70)	**0.21** ($\pm$0.24)	0.73 ($\pm$0.78)	0.73 ($\pm$0.15)	0.72 ($\pm$0.36)
	flat $\mathcal{N}$	0.48 ($\pm$0.36)	**0.44** ($\pm$0.34)	0.65 ($\pm$0.42)	0.73 ($\pm$0.15)	0.59 ($\pm$0.25)
	flat β	0.87 ($\pm$0.74)	**0.43** ($\pm$0.30)	0.71 ($\pm$0.52)	0.80 ($\pm$0.16)	1.03 ($\pm$0.17)
	flat Ψ_m	0.58 ($\pm$0.46)	**0.28** ($\pm$0.29)	0.48 ($\pm$0.50)	0.64 ($\pm$0.14)	0.58 ($\pm$0.13)
5000	steep $\mathcal{N}$	0.76 ($\pm$0.90)	**0.27** ($\pm$0.21)	1.10 ($\pm$0.75)	0.84 ($\pm$0.18)	0.96 ($\pm$1.02)
	steep β	0.7 ($\pm$0.79)	**0.28** ($\pm$0.26)	0.93 ($\pm$0.73)	0.75 ($\pm$0.13)	1.19 ($\pm$0.91)
	steep Ψ_m	0.92 ($\pm$0.61)	**0.12** ($\pm$0.16)	0.79 ($\pm$0.85)	0.54 ($\pm$0.13)	0.70 ($\pm$0.28)
	flat $\mathcal{N}$	0.44 ($\pm$0.65)	**0.31** ($\pm$0.27)	0.54 ($\pm$0.54)	0.53 ($\pm$0.11)	0.50 ($\pm$0.16)
	flat β	0.64 ($\pm$0.64)	**0.34** ($\pm$0.26)	0.63 ($\pm$0.57)	0.67 ($\pm$0.10)	1.02 ($\pm$0.17)
	flat Ψ_m	0.55 ($\pm$0.27)	**0.17** ($\pm$0.20)	0.58 ($\pm$0.56)	0.47 ($\pm$0.12)	0.68 ($\pm$0.13)

depending on their behavior around the objective s_* – flatter functions varying less around s_*. They include two functions, $\mathcal{N}_{\text{flat}}$ and $\mathcal{N}_{\text{steep}}$, based on a Gaussian c.d.f., with mean and standard deviation of respectively $m = 0.35$ and $\sigma = 0.5$ for $\mathcal{N}_{\text{flat}}$ (resp. $m = 0.66$ and $\sigma = 0.2$ for $\mathcal{N}_{\text{steep}}$). This is the most advantageous setting for Staircase and PsiMethod, as this setting satisfies the Gaussian hypothesis. Two other functions, β_{flat} and β_{steep}, are based on the c.d.f. of Beta variables, with respectively $\alpha = 2$ and $\beta = 1$ for β_{flat} (resp. $\alpha = 2$ and $\beta = 5$ for $\mathcal{N}_{\text{steep}}$). While not Gaussian, these functions are usual c.d.f. and are strongly smooth (continuously differentiable). The last two functions, Ψ^m_{flat} and Ψ^m_{steep}, are non decreasing Hölder continuous functions (the hardest setting) defined as :

$$\Psi^m(s^* + x) = \mu_* + 1_{x>0}|x|^{k_+} - 1_{x<0}|x|^{k_-}$$

where $k_+ = 1.5$ and $k_- = 0.5$ for Ψ^m_{flat} (resp. $k_+ = 1$ and $k_- = 0.3$ for Ψ^m_{steep}). For each function, the objective is to identify the stimulus s_* such that $\mu_* = 0.707$ (for flat functions) or $\mu_* = 0.5$ (for steep functions) – values that are reachable by Staircase. Importantly, the psychometric functions were clipped to the probability interval $[0.2, 0.8]$ for steep functions, $[0.1, 0.9]$ for flat functions, to represent arbitrary guess and lapse rates. During the experiments, only the value of μ_* was provided to the different algorithms.

Results

Small Time Horizon. Figure 1 reports the evolution average simple regret over 100 runs for $T = 200$. First, note that all methods tend to perform better on "flat" functions than on their "steep" counterpart – since they vary less around s_*, it is easier to find a reasonably good estimator. In particular, all methods achieve good results for Ψ^m_{flat}, which is significantly flat to the right of s_*. Second, PsiMethod performs poorly for non Gaussian Ψ, as they violate the assumption on its shape. Moreover, $\mu_{\min}$ and $\mu_{\max}$ were considered unknown – which is generally the case in practice– and this lack of prior information is known to impact the performance of Bayesian methods (García-Pérez & Alcalá-Quintana, 2007). Third, while POO seems to converge toward the solution for every function, it achieves the worst regret in all the studied settings, as the rate of convergence is slow (POO cannot take advantage of the monotonic property of Ψ). Finally, while GP and Staircase achieve reasonable performance, DOS provides one of the best estimation – if not the best – in all these settings, particularly for "difficult" functions such as Ψ^m_{steep}.Interestingly, the regret trajectory of DOS sometimes increases for a short time. This is due to the fact that when the agent activates a new arm, she might moves from a $s_i > s_*$ to a $s_{i+1} < s_*$, (or vice versa), and the regret may increase temporarily. However, as additional arms are pulled, the sequence s_i converges toward s_*, resulting in a small $\mathcal{R}_T$.

Large Time Horizon. Table 1 reports the average simple regret over 100 runs for different algorithm and psychometric function for three different time horizons : $T = 500$, 2000 and 5000. It can be seen that DOS outperforms its competitor in every case, and its advantage increases with T. Interestingly, Staircase and PsiMethod do not appear to converge, as they display little improvement between $T = 500$ and $T = 5000$. This is due to the fact that these

Table 2. Average (± standard deviation) over 100 runs of the simple regret in the psychometric setting, for each method and Psychometric functions with time horizon $T = 50$ and 100 . The reported regret has been multiplied by 10 for readability purpose.

Setting	ψ	T	DOS	Stair.	PsiM.	Quest+	GP	POO
Yes/No	Gaussian	50	1.01(0.50)	0.98(0.72)	0.94(0.58)	**0.90(0.41)**	1.52(0.46)	1.72(0.66)
		100	**0.70(0.47)**	0.90(0.64)	0.72(0.32)	0.70(0.50)	1.22(0.34)	1.51(0.60)
	Hölder	50	**1.08(0.56)**	1.17(0.55)	1.22(0.91)	1.31(1.11)	1.73(1.66)	1.68(0.70)
		100	**0.83(0.58)**	1.03(0.62)	1.12(0.72)	1.15(0.87)	1.25(0.75)	1.38 (0.62)
2-AFC	Gaussian	50	**1.14(0.60)**	1.28(0.73)	1.53(0.90)	1.33(0.62)	1.67(0.64)	1.70(0.67)
		100	**0.90(0.56)**	1.15(0.69)	1.33(0.69)	1.20(0.55)	1.27(0.44)	1.52(0.37)
	Hölder	50	**1.05(0.58)**	1.41(0.71)	1.74(1.10)	1.70(1.23)	1.65(0.63)	1.69(0.81)
		100	**0.85(0.47)**	1.25(0.72)	1.68(1.08)	1.71(1.25)	1.45(0.54)	1.40(0.79)

methods rely on manually constructed grids of hyperparameter (discretization of loc and scale for PsiMethod; scale of step size for Staircase) – which need to be tuned to Ψ to ensure convergence for large values of T, and thus requires prior knowledge of the function shape. Conversely, POO and GP tend to converge toward s_*, but DOS appears to consistently outperforms them. For POO, this can be seen as an illustration of DOS better regret bounds, while GP has no such theoretical guarantees. Finally, GP appears to be performing poorly for Ψ^m_{steep}; this can be explained by the fact that its kernel was suboptimal to estimate non standard functions such as Ψ^m_{steep}.

Mimicking Psychometric Experiments. In Table 2, we compared all the procedures in a setting which mimics the behavior of a small budget psychometric experiment ($T = 50$ and 100). We used two frameworks, which illustrates two common type of experimental settings : YES/NO ($\mu_{\min} = 0$, $\mu_{\max} = 1$, $\mu_* = 0.5$) and 2-AFC ($\mu_{\min} = 0.5$, $\mu_{\max} = 0.96$, $\mu_* = 0.707$), and two psychometric functions : the steep Gaussian c.d.f. and Hölder function previously described. The Yes/No Gaussian (resp. 2-AFC Hölder) setting was expected to be the most (resp. least) favorable for Bayesian methods (i.e. PsiMethod and Quest+).

Table 2 shows that indeed, for the first setting, Bayesian methods are slightly better than DOS, and that all methods are very close for $T = 100$ Yes/No Gaussian c.d.f.. Conversely, DOS is slightly better than its counterparts for $T = 50$ Yes/No Hölder and 2-AFC Gaussian, and significantly better in the other cases. In summary, these experiments show that DOS performs almost as good as its Bayesian counterparts in their ideal setting, and quickly outperforms them in non ideal settings or when T increases.

6. Discussion

In this works, we discussed a new method for solving the threshold estimation problem, DOS. Compared to previous works in global optimization such as POO (Grill et al., 2015), we showed that DOS has better regret bounds, and consistently performs better in our experiments. This is due to the fact that POO was developed for a different, more general setting and thus cannot take advantage of the properties of psychometric functions. Compared to other methods used in Psychophysics (Staircase, PsiMethod, GP), DOS is completely *model free*, does not require any assumption on the shape of Ψ, is parameter free, has strong theoretical guarantees and performs better empirically. Importantly, DOS is not a replacement for Bayesian psychometric methods such as PsiMethod and GP. Indeed, these model-based methods are designed to estimate the entire Ψ function, while DOS only estimates s_* (and thus is able to do it more efficiently). Consequently, they remain the tool of choice to use when the shape of the Ψ is known (e.g. Gaussian c.d.f.), and when the general behavior of the function is of interest to the experimenter – for instance when establishing a person's audiogram (Gardner et al., 2015b). However, when little is known about Ψ, or when the objective is to estimate s_* – the most frequent case in Psychophysics research, see e.g. (Schütz et al., 2008; García-Pérez & Alcalá-Quintana, 2007) – then DOS is a significantly better solution.

Future works might include the application of DOS to other problems outside the domain of Psychophysics, such as the handover of signal on a cellular network, see e.g. (Sun et al., 2019).

References

Audiffren, J. Quantifying human perception with multi-armed banditss. In *The 20th International Conference on Autonomous Agents and Multiagent Systems*, 2021.

Auer, P., Ortner, R., and Szepesvári, C. Improved Rates for the Stochastic Continuum-Armed Bandit Problem. In Bshouty, N. H. and Gentile, C. (eds.), *Learning Theory*, Lecture Notes in Computer Science, pp. 454–468, Berlin, Heidelberg, 2007. Springer. ISBN 978-3-540-72927-3. doi: 10.1007/978-3-540-72927-3_33.

Benson, A. J., Hutt, E. C., and Brown, S. F. Thresholds for the perception of whole body angular movement about a vertical axis. *Aviation, Space, and Environmental Medicine*, 60(3):205–213, 1989. ISSN 0095-6562(Print).

Brown, L. G. Additional rules for the transformed up-down method in psychophysics. *Perception & Psychophysics*, 58(6):959–962, 1996.

Bubeck, S., Munos, R., and Stoltz, G. Pure exploration in finitely-armed and continuous-armed bandits. *Theoretical Computer Science*, 412(19):1832–1852, April 2011. ISSN 0304-3975. doi: 10.1016/j.tcs.2010.12.059. URL `http://www.sciencedirect.com/science/article/pii/S030439751000767X`.

Chakraborty, M., Das, S., and Magdon-Ismail, M. Near-optimal target learning with stochastic binary signals. In *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence*, pp. 69–76, 2011.

Chhabra, M. and Das, S. Learning the demand curve in posted-price digital goods auctions. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pp. 63–70, 2011.

Cornsweet, T. N. The Staircase-Method in Psychophysics. *The American Journal of Psychology*, 75 (3):485, September 1962. ISSN 00029556. doi: 10.2307/1419876. URL `https://www.jstor.org/stable/1419876?origin=crossref`.

Fontaine, X., Mannor, S., and Perchet, V. An adaptive stochastic optimization algorithm for resource allocation. In *Algorithmic Learning Theory*, pp. 319–363, 2020.

García-Pérez, M. A. and Alcalá-Quintana, R. Bayesian adaptive estimation of arbitrary points on a psychometric function. *British Journal of Mathematical and Statistical Psychology*, 60(1):147–174, 2007. ISSN 2044-8317. doi: 10.1348/000711006X104596. URL `https://onlinelibrary.wiley.com/doi/abs/10.1348/000711006X104596`.

Gardner, J., Malkomes, G., Garnett, R., Weinberger, K. Q., Barbour, D., and Cunningham, J. P. Bayesian active model selection with an application to automated audiometry. *Advances in neural information processing systems*, 28:2386–2394, 2015a.

Gardner, J. R., Song, X., Weinberger, K. Q., Barbour, D. L., and Cunningham, J. P. Psychophysical detection testing with bayesian active learning. In *UAI*, pp. 286–295, 2015b.

Grill, J.-B., Valko, M., and Munos, R. Black-box optimization of noisy functions with unknown smoothness. December 2015. URL `https://hal.inria.fr/hal-01222915`.

Hatzfeld, C., Hoang, V. Q., and Kupnik, M. It's All About the Subject - Options to Improve Psychometric Procedure Performance. In Bello, F., Kajimoto, H., and Visell, Y. (eds.), *Haptics: Perception, Devices, Control, and Applications*, volume 9774, pp. 394–403. Springer International Publishing, Cham, 2016. ISBN 978-3-319-42320-3 978-3-319-42321-0. doi: 10.1007/978-3-319-42321-0-36. URL `http://link.springer.com/10.1007/978-3-319-42321-0-36`.

Hirahara, T. Physical characteristics of headphones used in psychophysical experiments. *Acoustical science and technology*, 25(4):276–285, 2004.

Jedynak, B., Frazier, P. I., and Sznitman, R. Twenty questions with noise: Bayes optimal policies for entropy loss. *Journal of Applied Probability*, 49(1):114–136, 2012.

Kleinberg, R., Slivkins, A., and Upfal, E. Multi-armed bandits in metric spaces. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pp. 681–690, 2008.

Kleinberg, R., Slivkins, A., and Upfal, E. Bandits and experts in metric spaces. *Journal of the ACM (JACM)*, 66 (4):1–77, 2019.

Kontsevich, L. L. and Tyler, C. W. Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, 39(16):2729–2737, August 1999. ISSN 0042-6989. doi: 10.1016/S0042-6989(98)00285-5. URL `http://www.sciencedirect.com/science/article/pii/S0042698998002855`.

Leek, M. R. Adaptive procedures in psychophysical research. *Perception & Psychophysics*, 63(8):1279–1292, November 2001. ISSN 1532-5962. doi: 10.3758/BF03194543. URL `https://doi.org/10.3758/BF03194543`.

Lengyel, G. and Fiser, J. The relationship between initial threshold, learning, and generalization in perceptual learning. *Journal of Vision*, 19(4):28–28, April 2019. ISSN 1534-7362. doi: 10.1167/19.4.28. URL https://jov.arvojournals.org/article.aspx?articleid=2732298.

Levitt, H. Transformed Up-Down Methods in Psychoacoustics. *The Journal of the Acoustical Society of America*, 49 (2B):467–477, February 1971. ISSN 0001-4966. doi: 10.1121/1.1912375. URL https://asa.scitation.org/doi/abs/10.1121/1.1912375.

Nir, R.-R., Granovsky, Y., Yarnitsky, D., Sprecher, E., and Granot, M. A psychophysical study of endogenous analgesia: the role of the conditioning pain in the induction and magnitude of conditioned pain modulation. *European journal of pain*, 15(5):491–497, 2011.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., and Lindeløv, J. K. PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1):195–203, February 2019. ISSN 1554-3528. doi: 10.3758/s13428-018-01193-y. URL https://doi.org/10.3758/s13428-018-01193-y.

Scheuerman, J., Venable, K. B., Anderson, M. T., and Golob, E. J. Modeling spatial auditory attention: Handling equiprobable attended locations. In *CAID@ IJCAI*, pp. 1–7, 2017.

Schütz, A. C., Braun, D. I., Kerzel, D., and Gegenfurtner, K. R. Improved visual sensitivity during smooth pursuit eye movements. *Nature neuroscience*, 11(10):1211–1216, 2008.

Shang, X., Kaufmann, E., and Valko, M. General parallel optimization a without metric. In *Algorithmic Learning Theory*, pp. 762–788, March 2019. URL http://proceedings.mlr.press/v98/xuedong19a.html.

Shen, Y. and Richards, V. M. A maximum-likelihood procedure for estimating psychometric functions: Thresholds, slopes, and lapses of attention. *The Journal of the Acoustical Society of America*, 132(2):957–967, August 2012. ISSN 0001-4966. doi: 10.1121/1.4733540. URL https://asa.scitation.org/doi/full/10.1121/1.4733540.

Song, X. D., Garnett, R., and Barbour, D. L. Psychometric function estimation by probabilistic classification. *The Journal of the Acoustical Society of America*, 141(4): 2513–2525, 2017.

Sun, L., Hou, J., and Shu, T. Optimal handover policy for mmwave cellular networks: A multi-armed bandit approach. In *2019 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6. IEEE, 2019.

Valko, M., Carpentier, A., and Munos, R. Stochastic Simultaneous Optimistic Optimization. In *International Conference on Machine Learning*, pp. 19–27, February 2013. URL http://proceedings.mlr.press/v28/valko13.html.

Watson, A. B. Quest+: A general multidimensional bayesian adaptive psychometric method. *Journal of Vision*, 17(3): 10–10, 2017.

Wichmann, F. A. and Hill, N. J. The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, 63(8):1293–1313, November 2001a. ISSN 1532-5962. doi: 10.3758/BF03194544. URL https://doi.org/10.3758/BF03194544.

Wichmann, F. A. and Hill, N. J. The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics*, 63(8):1314–1329, November 2001b. ISSN 1532-5962. doi: 10.3758/BF03194545. URL https://doi.org/10.3758/BF03194545.

Wichmann, F. A. and Jäkel, F. Methods in psychophysics. *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*, 5:1–42, 2018.

A. Technical Proofs

Proof of Lemma 2

Lemma. *If* (7) *is satisfied for the arm s_i, then*

$$q_i < \frac{2}{T^3}$$

Proof. Without any loss of generality suppose that $\mu_i > \mu_*$. Note that

$$\{\hat{\mu}_i(T) + \mathcal{B}_T(N_i(t)) < \mu_*\} \subset \{\mu_i - \hat{\mu}_i(T) > \mathcal{B}_T(N_i(t))\}$$

Hence, using the Chernoff Hoeffding concentration inequality (5) and (7),

$$\begin{aligned} q_i &< 2\exp\left(-2N_i(t)\mathcal{B}_T(N_i(t))^2\right) \\ &\le 2\exp\left(-3N_i(t)\frac{\log T}{N_i(t)}\right) \le \frac{2}{T^3} \end{aligned}$$

□

Proof of Lemma 3

Lemma. *If* (7) *is false, but* (8) *is satisfied, then with probability at least* $1 - \frac{2}{T^3}$,

$$\Delta_i < 3\sqrt{\frac{(\log T)^2(\log^2 T)}{T}}$$

Proof. We have

$$\begin{aligned} \Delta_i &= |\mu_i - \mu_*| \\ &\le \underbrace{|\hat{\mu}_i - \mu_i|}_{A} + \underbrace{|\hat{\mu}_i - \mu_*|}_{B} \end{aligned}$$

Since the first rule was not activated, this implies that (7) is not true, hence

$$B \le \mathcal{B}_T(N_*).$$

Moreover, using the Chernoff Hoeffding concentration inequality (5), we have that with probability at least $1 - \frac{2}{T^3}$,

$$A \le \mathcal{B}_T(N_*).$$

Hence,

$$\begin{aligned} \Delta_i &\le 2\mathcal{B}_T(N_*) \\ &\le 3\sqrt{\frac{\log T}{N*}} \\ &\le 3\sqrt{\frac{(\log T)^2(\log^2 T)}{T}} \end{aligned}$$

□

Proof of Theorem 1

Let $\mathcal{A} = \{\forall 1 \le i \le \kappa, N_i(T) < N_*\}$. In other words, $\mathcal{A}$ is the event where (10) is never satisfied, thus all arms are activated using rule (8).

On $\Omega \setminus \mathcal{A}$, at least one arm satisfy (10), hence a promising arm is returned. Hence by using Lemma 3 we have

$$\mathbb{P}\left(\mathcal{R}_T \le 2\sqrt{\frac{(\log T)^2(\log^2 T)}{T}}\;\middle|\;\Omega \setminus \mathcal{A}\right) \ge 1 - \frac{2}{T^2} \tag{12}$$

In the following, we examine the behavior of the regret on $\mathcal{A}$. Let

$$\mathcal{A}^* \doteq \mathcal{A} \cap \{\forall t \le T, \forall i \le \kappa, \, |\mu_i - \hat{\mu}_i(t)| \le \mathcal{B}_T(N_i)\},$$

Let $\mathbb{P}_{\mathcal{A}^*}(\cdot) \doteq \mathbb{P}(\cdot|\mathcal{A}^*)$. We say that an event $\mathcal{E}$ is $\mathcal{A}^*$ almost sure ($\mathcal{A}^*$ a.s.) if $\mathbb{P}_{\mathcal{A}^*}(\mathcal{E}) = 1$. One difficulty of our setting is that the sequence of activated arms $(s_i)_{i\ge1}$ is *a priori* random, as DOS has two choice for each new arm. However this problem is easily solved on $\mathcal{A}^*$, as stated by Lemma **??**.

Lemma 5. $\mathbb{P}_{\mathcal{A}^*}\left(\forall \kappa > i \ge 1, s_{i+1} = s_i + \frac{1}{2^{i+1}}\textbf{\textit{sign}}(\mu_* - \mu_i)\right) = 1$

Proof. Let $\kappa > i > 0$. Suppose without any loss of generality that $\mu_* > \mu_i$ (the other case is proved similarly). $\mathcal{A}^*$ a.s., we have

$$\begin{aligned} \mu_* - \hat{\mu}_i(T) &= \mu_* - \mu_i + \mu_i - \hat{\mu}_i(t_i) \\ &\ge \mu_i - \hat{\mu}_i(t_{i+1}) \\ &\ge -\mathcal{B}(\delta, T, N_i(t_{i+1})) \end{aligned} \tag{13}$$

But arm $i+1$ was activated using (8) hence

$$|\mu_i - \hat{\mu}_i(T)| > \mathcal{B}_T(N_i(T)).$$

Thus,

$$\mu_* - \hat{\mu}_i(T) > 0.$$

Hence the conclusion. □

An consequence of Lemma **??** is that on $\mathcal{A}^*$ the sequence of arms s_i is *fixed*. Note that the number of pull per arm is still random. Now we can prove that the event $\mathcal{A}^*$ has high probability on $\mathcal{A}$.

Lemma. $\mathbb{P}(\mathcal{A}^*) = \mathbb{P}(\mathcal{A}^* \cap \mathcal{A}) \ge \mathbb{P}(\mathcal{A}) - \frac{2}{T}$.

Proof. This directly results from Lemma 2 and the previous lemma by taking the union bound on all arms and all times.

□

The following Lemma shows that the sequence of activated arms s_i converge exponentially fast toward the threshold s_*, independently of the smoothness of Ψ.

Lemma 6. $\forall i, \mathbb{P}_{\mathcal{A}^*}\left(|s_i - s_*| \leq 2^{-i}\right) = 1$

Proof. We prove Lemma **??** by iteration. For $i = 1$, we have $s_1 = 1/2$. Since $s_* \in [0,1]$, $|s_* - 1/2| \leq 1/2$. Hence the result is true for rank $i = 1$. Now suppose that the result holds for rank i. Then, $\mathcal{A}^*$ a.s.,

$$
\begin{aligned}
|s_{i+1} - s_*| &= |s_i + 2^{-i+1}\mathbf{sign}(\mu_* - \mu_i) - s_*| \\
&= |s_i - s_* + 2^{-i+1}\mathbf{sign}(s_* - s_i)| \\
&\leq 2^{i+1}
\end{aligned}
$$

where we used Lemma **??** in the first line, Ψ non decreasing for the second line, and for the last line the fact that the result holds at rank i and $\mathbf{sign}(s_* - s_i) = -\mathbf{sign}(s_i - s_*)$. Hence the result is true at rank $i + 1$. □

Corollary 1. *If Assumption 1 is true, then*

$$
\mathbb{P}_{\mathcal{A}^*}\left(\forall i > 0, \Delta_i \leq \nu\rho^i\right) = 1 \tag{14}
$$

Proof. This corollary is an immediate consequence of Lemma **??** and Assumption 1. □

Proposition. *Let i_* defined as*

$$
i_* \doteq \left\lfloor \log\left(\frac{T}{(\log^2 T)(\log T)^2}\right)\frac{1}{2\log(1/\rho)}\right\rfloor.
$$

Then all the following properties are true

(A) if $\log^2 T > -1/\log\rho$ and $T > 16$, then $i_* \leq \kappa$

(B) $\forall i > i_*$, $\mathcal{A}^*$ a.s.

$$
\Delta_i \leq \nu\sqrt{\frac{(\log^2 T)(\log T)^2}{T}},
$$

Proof. **Proof of (A).**

$$
\begin{aligned}
i_* &\leq \log\left(\frac{T}{(\log^2 T)(\log T)^2}\right)\frac{1}{2\log(1/\rho)} \\
&\leq \frac{\log(T)}{2\log(1/\rho)} \\
&\leq \frac{(\log T)(\log^2 T)}{2(\log^2 T)\log(1/\rho)} \\
&\leq \frac{(\log T)(\log^2 T)}{2} < \frac{\kappa}{2}.
\end{aligned}
$$

where in the second line we used $\log\left((\log^2 T)(\log T)^2\right) > 0$ and in third line we used the fact that $\log^2 T > -1/\log\rho$.

Proof of (B).

First note that

$$
\begin{aligned}
\rho^{i_*} &= \exp(i_*\log\rho) \\
&\leq \exp\left(-\log\left(\frac{T}{(\log^2 T)(\log T)^2}\right)\frac{\log\rho}{2\log\rho}\right) \\
&\leq \exp\left(\frac{1}{2}\log\left(\frac{(\log^2 T)(\log T)^2}{T}\right)\right) \\
&\leq \sqrt{\frac{(\log^2 T)(\log T)^2}{T}}.
\end{aligned}
$$

Hence, using Corollary **??**,

$$
\Delta_i \leq \nu\rho^{i_*} \quad \leq \nu\sqrt{\frac{(\log^2 T)(\log T)^2}{T}}
$$

□

To conclude proof of Theorem 1, let

$$
\mathcal{E} = \left\{\mathcal{R}_T \leq (3+\nu)\sqrt{\frac{(\log T)^2(\log^2 T)}{T}}\right\}.
$$

Note that

$$
\begin{aligned}
\mathbb{E}\left(\mathcal{R}_T\right) &\leq (3+\nu)\sqrt{\frac{(\log T)^2(\log^2 T)}{T}}\mathbb{P}(\mathcal{E}) + (1 - \mathbb{P}(\mathcal{E})) \\
&\leq (3+\nu)\sqrt{\frac{(\log T)^2(\log^2 T)}{T}} + (1 - \mathbb{P}(\mathcal{E}))
\end{aligned}
$$

where in the first line we used the fact that $\mathcal{R}_T \leq 1$. Moreover,

$$
\begin{aligned}
\mathbb{P}(\mathcal{E}) &= \mathbb{P}(\mathcal{E}\cap\mathcal{A}^*) + \mathbb{P}(\mathcal{E}\cap(\mathcal{A}\setminus\mathcal{A}^*)) + \mathbb{P}(\mathcal{E}\cap(\Omega\setminus\mathcal{A})) \\
&\geq \underbrace{\mathbb{P}(\mathcal{E}\cap\mathcal{A}^*)}_{A_1} + \underbrace{\mathbb{P}(\mathcal{E}\cap(\Omega\setminus\mathcal{A}))}_{A_2}.
\end{aligned}
$$

Using Proposition 1,

$$
\mathbb{P}(\mathcal{E}\cap\mathcal{A}^*) = \mathbb{P}(\mathcal{E}|\mathcal{A}^*)\mathbb{P}(\mathcal{A}^*) = \mathbb{P}(\mathcal{A}^*)
$$

and using (**??**),

$$
\begin{aligned}
\mathbb{P}(\mathcal{E}\cap(\Omega\setminus\mathcal{A})) &= \mathbb{P}(\mathcal{E}|\Omega\setminus\mathcal{A})\mathbb{P}(\Omega\setminus\mathcal{A}) \\
&\geq \left(1 - \frac{2}{T^2}\right)\mathbb{P}(\Omega\setminus\mathcal{A}) \\
&\geq \mathbb{P}(\Omega\setminus\mathcal{A}) - \frac{2}{T^2}
\end{aligned}
$$

thus by using Lemma 4,

$$\begin{aligned}\mathbb{P}(\mathcal{E}) &\geq \mathbb{P}(\mathcal{A}^*) + \mathbb{P}(\Omega \setminus \mathcal{A}) - \frac{2}{T^2}\\ &\geq 1 - \mathbb{P}(\mathcal{A} \setminus \mathcal{A}^*) - \frac{2}{T^2}\\ &\geq 1 - \frac{4}{T}\end{aligned}$$

hence the conclusion.

Proof of Proposition 2

Proposition (Non zero optimality dimension)**.** *There exist psychometric functions* Ψ *satisfying Assumption 1 such that* $\min_{\nu,\rho} d_\Psi(\nu, \rho) > 0$. *In particular, for*

$$\Psi(x) = \begin{cases} \min\left(1, \mu_* + \exp(-1/|x - s_*|)\right) & \text{if } x > s_* \\ \max\left(0, \mu_* + |x - s_*|^{2/5}\right) & \text{if } x \leq s_* \end{cases} \tag{15}$$

we have $d_\Psi \geq \frac{\log 2}{\log 5 - \log 4}$

Proof. Let Ψ as defined in (**??**). It is easy to see that Ψ is strictly increasing. Moreover Ψ α Hölder continuous for $\alpha = 2/5$, and therefore it satisfies Assumption 1. Additionally,

$$\Psi^{-1}(\mu_* + 2\nu\rho^h) - \Psi^{-1}(\mu_* - 2\nu\rho^h) \geq -\frac{1}{\log(\nu\rho^h)} \tag{16}$$

Now let C, d' as in (1). We have, using (**??**)

$$-\frac{2}{\log(\nu\rho^h)} \leq C(2\rho^{d'})^{-h}$$

Note that when $h \to \infty$ the left part is $\mathcal{O}(1/h)$ while the right part is $\mathcal{O}(\left(\frac{1}{2\rho^{d'}}\right)^h)$. Consequently $2\rho^{d'} \leq 1$ is a necessary condition for the inequality to be true when $h \to \infty$, which is turn implies the conclusion. $\square$