Zooming Optimistic Optimization Method to solve the Threshold Estimation Problem

Julien Audiffren
Department of Computer Science
Fribourg University, Switzerland
JULIEN.AUDIFFREN@UNIFR.CH

Abstract

This paper introduces a new global optimization algorithm that solves the threshold estimation problem. In this active learning problem, underlying many empirical neuroscience and psychophysics experiments, the objective is to estimate the input values that would produce the desired output value from an unknown, noisy, non-decreasing response function. Compared to previous approaches, ZOOM (Zooming Optimistic Optimization Method) offers the best of both worlds: ZOOM is model-agnostic, benefits from stronger theoretical guarantees and faster convergence rate, and also quickly jumps between arms, offering strong performance even for small sampling budgets.

1. Introduction

The Threshold Estimation Problem (TEP) is an active learning problem, where an agent is given an interval of possible input values $\mathbb{I}$, a desired output probability $\mu_* \in [0, 1]$, a sampling budget T, and an unknown black-box non-decreasing response function ψ that can only be accessed through noisy observations, where the noise can depend on the input value. The objective is to provide an estimator $\hat{s}$ of an input value that will produce a desired output value, i.e. $\hat{s} \approx s_* \in \psi^{-1}(\mu_*)$.

The TEP is central to Psychophysics[1] and in particular the evaluation of human perception, which is generally assessed by performing psychometric experiments. During these, an experimenter (in our setting, the agent) presents to an observer, a sequence of stimuli of varying intensities s_t, $1 \leq t \leq T$ (for instance, the contrast of a visual stimulus, see e.g. (Audiffren et al., 2022)). After each stimulus, the observer signals to the experimenter whether she was able to see the stimulus – which is modeled as a Bernoulli random variable whose mean (the probability of perception) is $\psi(s_t)$, i.e. the output of a black-box non-decreasing response function called the psychometric function. The objective of many such experiments is to find the *sensitivity threshold*, where the stimulus is just noticeable (Kontsevich and Tyler, 1999), which is often defined as $\mu_* = \frac{1}{2}$, $\frac{2}{3}$ or $\frac{3}{4}$, which correspond respectively to the 50 %, 66.6 % and 75 % detection threshold.

Related Work. The staircase algorithm is arguably the most popular adaptive method to solve the TEP and has been discussed and improved upon significantly in recent years (Wichmann and Jäkel, 2018). However, this method can only be used for a very limited list of possible target probability (such as $\mu_* = 0.5$) (Brown, 1996), and convergence is only guaranteed for specific shapes of the psychometric function (Levitt, 1971). Parametric Bayesian adaptive algorithms (Kontsevich

1. Psychophysics is a research field that studies the relationship between physical stimuli and the perceptions they produce

and Tyler, 1999; Shen and Richards, 2012; Watson, 2017), whose purpose is generally to estimate the entire function ψ, require prior knowledge of the response function parametric model, which limits their applications. More recently, several works have proposed to use Gaussian Processes (GP) to approximate ψ (Gardner et al., 2015a,b; Song et al., 2017). GP addresses some of the shortcomings of the Bayesian methods by being more flexible, but these methods tend to be more costly and require significantly more samples to approximate s_*. Arguably the closest method to ours is DOS (Audiffren, 2021a). However, as the authors acknowledged themselves and observed in their following works (Audiffren and Bresciani, 2022), DOS is very conservative in its estimates, leading to subpar performance when the sampling budget T is small and the response function is strongly smooth. This limits the usability of DOS in practical applications, in particular for some experimental psychophysics experiments where T is very small (Averbeck et al., 2017).

Contributions. In this paper, we introduce and analyze ZOOM (Zooming Optimistic Optimization Method), a new algorithm that solves the threshold estimation problem. ZOOM builds on DOS, by taking inspiration from hierarchical bandits, where the agent uses concentration inequalities to explore a hierarchical partition of the arm space, and global optimization of noisy black-box functions of smoothness (see e.g. (Grill et al., 2015; Shang et al., 2019b)). Compared to previous approaches, ZOOM offers the best of both worlds, as it is model-agnostic, has strong theoretical guarantees, and shows optimal or close to optimal performance in all our experiments.

2. The Threshold Estimation Problem

Notation. We borrow the notation and vocabulary of the multi-armed bandit version of the problem introduced by (Audiffren, 2021a). Let T denote the time horizon (i.e. the sampling budget), $\mathbb{I} = [0, 1]$ the bounded, closed input interval, $\psi : \mathbb{I} \mapsto [0, 1]$ the continuous[2], non-decreasing response function, $\mu_* \in \psi(\mathbb{I})^o$ the target probability (i.e. μ_* is in the interior of the image of ψ) and $s_* \doteq \psi^{-1}(\mu_*)$ the desired threshold[3].

The objective of the TEP is to find an estimator $\hat{s}$ of the sensitivity threshold s_* with at most T samples. $\mathbb{I}$, T and μ_* are known to the agent, but ψ is not. The process unfolds as follows. For each round $t \in [1, \ldots, T]$, the agent chooses an arm (i.e. chooses a value to sample) $s \in \mathbb{I}$. Then the environment draws an independent Bernoulli random variable with mean $\psi(s)$, and communicates the result to the agent. At time $t = T$, the agent returns the arm $\hat{s}$ that is her best guess for the threshold s_*. The performance of the agent is then evaluated using the simple regret $\mathcal{R}$, defined as

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

2.1 Smoothness of the response function

Additional smoothness assumptions for ψ are required to provide theoretical guarantees for the convergence of an estimator – this is to avoid the classical “needle in a haystack” problem of global optimization of black box function (see e.g. (Valko et al., 2013)). As presented below, in this work we assume that the Dini derivative of ψ on s_* are finite and strictly positive, which is a mild assumption.

2. Note that to keep the notation simple, we assume here that ψ is continuous, but all of our results can easily be extended to more general case where ψ is only measurable.
3. s_* is unique and well defined when ψ satisfies Assumption 1.

Dini Derivatives. The Dini derivatives are a generalization of the notion of derivative for real valued functions. Let $s \in \mathbb{I}$, the four different Dini derivatives of ψ on s are:

$$D^{-}\psi(s) = \inf_{\delta>0} \sup_{h\in[0,\delta]} \frac{\psi(s-h)-\psi(s)}{-h} \quad D^{+}\psi(s) = \inf_{\delta>0} \sup_{h\in[0,\delta]} \frac{\psi(s+h)-\psi(s)}{h}$$
$$D_{-}\psi(s) = \sup_{\delta>0} \inf_{h\in[0,\delta]} \frac{\psi(s-h)-\psi(s)}{-h} \quad D_{+}\psi(s) = \sup_{\delta>0} \inf_{h\in[0,\delta]} \frac{\psi(s+h)-\psi(s)}{h}$$

$D_{-}\psi(s)$, $D_{+}\psi(s)$, $D^{-}\psi(s)$ and $D^{+}\psi(s)$ take value in $[-\infty, \infty]$ and are always well defined for any point of the function domain. Since in the TEP, ψ is assumed to be increasing, the Denjoy–Young–Saks theorem (Saks, 1937) applies, and the Dini derivates of ψ are always positive and finite almost everywhere. Hence in this work, we make the following mildly stronger assumption:

Assumption 1 (ψ Dini Derivatives on s_*)

$$\infty > D^{-}\psi(s_*) \geq D_{-}\psi(s_*) > 0$$
$$\infty > D^{+}\psi(s_*) \geq D_{+}\psi(s_*) > 0.$$

Interestingly, if ψ is differentiable on s_* and ψ is strictly increasing on s_*, then ψ satisfies Assumption 1, as in this case $D^{-}\psi(s_*) = D^{+}\psi(s_*) = D_{-}\psi(s_*) = D_{+}\psi(s_*) = \psi'(s_*) > 0$.

3. Contributions

3.1 ZOOM strategy.

Let $K \in \mathbb{N}^+$ be a positive integer. ZOOM relies on a partition tree which is made of uniform grids $\mathcal{G}_{d,n}$, $d \geq 1$, $0 \leq n < K^{d-1}$. Each grid contains $K+1$ elements, and is of resolution $1/K^d$:

$$\mathcal{G}_{d,n} = \{s_{d,n,0},\ s_{d,n,1},\ \ldots,\ s_{d,n,K}\} \qquad \text{where } s_{d,n,k} = \frac{nK+k}{K^d} \tag{2}$$

When using ZOOM, the agent only samples arms that are part of a grid, that is to say the $s_{d,n,K}$. We denote by $N_{d,n,k}(t)$ the the number of time the arm $s_{d,n,k}$ has been pulled at time t, and $\hat{\mu}_{d,n,k}(t)$ the corresponding empirical average of its observations. Moreover we use $\mu_{d,n,k} = \psi(s_{d,n,k})$ and $\Delta_{d,n,k} = |\mu_{d,n,k} - \mu_*|$. The general idea of ZOOM can be summarized as follows: the agent starts with the grid $\mathcal{G}_{1,0} = \{0, 1/K, \ldots, 1\}$ that covers the entire input interval $\mathbb{I} = [0,1]$ with resolution $1/K$. Then, at each time t, the agent starts from first grid, i.e. $d = 1$; $n = 0$ and then :

1. Find the most promising interval k (see Section 3.1.2). If the agent is confident enough that the desired input is in this interval, she "zooms" on the sub-grid which spawns this interval ($d \leftarrow d+1$, and $n \leftarrow nK+k$) and repeat this step.

2. Once she has found a promising interval but where she can't "zoom", she samples the extremities of the interval for which she has the least amount of information (see Section 3.1.3).

The process is repeated until $t = T$ and the sampling budget is elapsed. Then, the arm that was pulled the most in the deepest reached subgrid is returned (see Section 3.1.4).

Choice of the parameter K. K directly quantify the coarseness of the grid, and the larger the K, the closer s_* will be to one of the grid point. As stated by Theorem 1 (Section 3.2), the following value of K, which is adaptive in T, leads to a fast convergence rate of the estimator:

$$K = \left\lfloor \sqrt{\frac{T}{\log T \log\log T}} \right\rfloor \tag{3}$$

Importantly ZOOM is *model-agnostic*, in the sense that it does not use prior knowledge regarding ψ, such as the values of the Dini derivatives of ψ. Moreover, the choice of K defined by (3) lead to favorable regret guarantees for all ψ satisfying Assumption 1 – a phenomenon that was confirmed in our experiments (see Section 4).

3.1.1 Confidence Intervals

Let $d \geq 1$ and $n, k \geq 0$. Since ψ is assumed to be non-decreasing, we have $s_* \in [s_{d,n,k}, s_{d,n,k+1}]$ i.f.f. $\psi(s_{d,n,k}) \leq \mu_* \leq \psi(s_{d,n,k+1})$. However, while the agent can observe if $\hat{\mu}_{d,n,k} \leq \mu_*$, she can never be sure if $\mu_{d,n,k} \leq \mu_*$. To address this problem, our method uses a concentration inequality derived from the Azuma-Hoeffding inequality, and is a confidence interval commonly used for the optimism against uncertainty principle, see e.g. (Auer et al., 2007). In this case, the agent decides that that $\mu_{d,n,k} \leq \mu_*$ if $\hat{\mu}_{d,n,k} \leq \mu_*$ and

$$|\mu_* - \hat{\mu}_i(t)| > \mathcal{B}_T(N_i(t)) \doteq \sqrt{\frac{3\log T}{2N_i(t)}}, \tag{4}$$

The second strategy is derived from the work of (Audibert et al., 2009), which was further extended in (Garivier et al., 2022). Here, the agent decide that that $\mu_{d,n,k} \leq \mu_*$ if $\hat{\mu}_{d,n,k} \leq \mu_*$ and

$$\mathsf{kl}(\hat{\mu}_i(t), \mu_*) > \mathcal{K}(N_i(t)) \doteq 2\frac{\log(T/N_i(t))}{N_i(t)}. \tag{5}$$

where kl denotes the Kullback-Leibler divergence. Contrarily to (4), this inequality is not related to a concentration inequality that is true with a probability close to one. However, (Garivier et al., 2022) have shown that by using this inequality, arms that are significantly worse than the optimal arm are only pulled a small number of times, resulting in a lower cumulative regret than UCB (Auer, 2002). ZOOM uses the two inequalities (4) and (5) alternatively, to leverage the advantages of both conditions: a strong theoretical guarantees (4) (Section 3.2), and better experimental results (5) (Section 4).

3.1.2 Finding the most promising interval

Let $\mathcal{C}$ be the condition used by ZOOM to reach a decision when comparing arms to the target threshold s_* (alternately (4) and (5)). To simplify the notation, in the following we use: $s_{d,n,k} \ll s_*$ if $\hat{\mu}_{d,n,k} < \mu_*$ and $\mathcal{C}$ is satisfied. By construction, $s_{d,n,0} \ll s_* \ll s_{d,n,K}$. Indeed, the agent zooms on an interval only if she has deduced that s_* is in the interior of the interval. For other values of $0 < k < K$, if the arm has never been pulled, i.e. $N_{d,n,k} = 0$, then $\mathcal{C}$ is not satisfied and none of the above notations apply.

Let $d \geq 1$, $n \geq 0$. To find the most promising interval of $\mathcal{G}_{d,n}$, the agent starts with its middle element $s_{d,n,K/2}$. If $N_{d,n,k} = 0$, then this interval is chosen. Otherwise, if $\hat{\mu}_{d,n,k} \geq \mu_*$, then the agent chooses the interval d, n, k' such that

$$k' = \max \left\{ \ell \geq 0, \text{ s.t. } N_{d,n,\ell} = 0 \text{ or } \hat{\mu}_{d,n,\ell} < \mu_* \right\}.$$

In the other case where $\hat{\mu}_{d,n,k} < \mu_*$, the agent chooses the interval d, n, k' where

$$k' = \min \left\{ \ell \geq 0, \text{ s.t. } N_{d,n,\ell} = 0 \text{ or } \hat{\mu}_{d,n,\ell} > \mu_* \right\}.$$

k' is well defined in both cases as $\ell = 0$ (resp. $\ell = K$) always satisfies the aforementioned condition.

After the interval d, n, k' is chosen, the agent decides to zoom on the interval, if and only if $s_{d,n,k} \ll s_* \ll s_{d,n,k+1}$, in which case the agent repeats the process using $d \leftarrow d+1$, $n \leftarrow nK + k$ and $k = \lfloor K/2 \rfloor$. Otherwise the agent moves to the next step of the algorithm (Section 3.1.3).

3.1.3 Choosing the interval extremity

After the interval of interest d, n, k has been selected, the agent decides which endpoint – either $s_{d,n,k}$ or $s_{d,n,k+1}$ – to explore as follows. First, if $s_{d,n,k} \ll s_*$ (resp. $s_* \ll s_{d,n,k+1}$), then the other endpoint $s_{d,n,k+1}$ (resp. $s_{d,n,k}$) is sampled. This represents the case where the agent has already enough information regarding one endpoint of the interval. Else, if $N_{d,n,k} = 0$ (resp. $N_{d,n,k+1} = 0$), then this endpoint $s_{d,n,k}$ (resp. $s_{d,n,k+1}$) is sampled, as there is no information regarding this arm. Finally, if none of the conditions above apply, the endpoint to sample is selected uniformly at random.

3.1.4 ZOOM Output estimator

When the time horizon is reached ($t = T$), then ZOOM outputs $\hat{s} = s_{d_*,n_*,k_*}$ as its estimator, where

$$\begin{aligned} d_* &= \max\{d \geq 0, \exists n, k \text{ s.t. } N_{d,n,k} > 0\}, \\ n_*, k_* &= \arg\max_{n,k} \left\{ N_{d_*,n,k} \right\}, \end{aligned}$$

with in case of equality, the agent chooses at random. In other words, the agent returns the arm that was sampled the most among the deepest subgrid that can be reached using (4).

3.2 Regret Upper Bound

The following upper bound holds on the regret incurred by ZOOM (see the appendix for the proof).

Theorem 1 (Regret Upper Bound) *Let $T > 0$, and ψ a response function that satisfies Assumption 1.Then, if ZOOM is run with parameter K as in* (3)*, there exists a constant $C_\psi > 0$, that only depends on ψ, such that its regret $\mathcal{R}_T$ is upper bounded with probability at least $1 - 2/T$ by*

$$\mathcal{R}_T \leq C_\psi \sqrt{\frac{\log T \log \log T}{T}} \tag{6}$$

4. Experiments

In this section we perform an empirical evaluation of ZOOM. We compare ZOOM to Staircase (Wichmann and Jäkel, 2018), QuestPlus (Watson, 2017), GP (Gardner et al., 2015a), DOS (Audiffren,

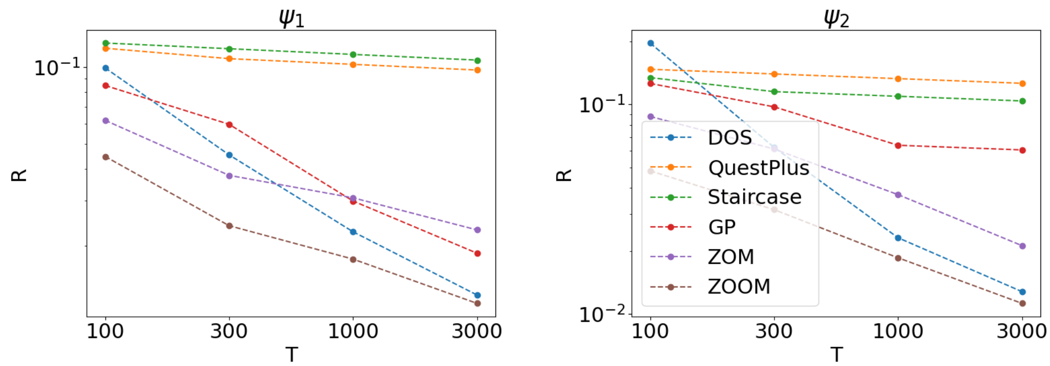

Figure 1: Average of the regret incurred by ZOOM compared to state-of-the-art methods (DOS, Staircase, POO, QuestPlus, GP) for multiple time horizons ($T = 100$, 300, 1000 and 3000) and response functions $(\psi_i)_{1 \le i \le 2}$. The regret is displayed on a log scale

2021a), using the implentation provided by (Peirce et al., 2019), and the parameters recommended by their respective papers. We also study ZOM, a version of ZOOM that only relies of the concentration inequality (4). For each experiment, we tested four different time horizons $T = 100$, 300, 1000 and 3000, three possible target values $\mu_* = 0.5$, $2/3$ and $3/4$, and two different functions ψ_0 –the cdf of $\mathcal{N}(\mu, \sigma^2 = 0.25)$– and $\psi_1(s) = \mu_* - 5(s_* - s)\mathbf{1}_{s<s_*} + 20(s - s_*)\mathbf{1}_{s>s_*}$.[4]

Results. Figure 1 shows the results of our experiment. Interestingly the version of ZOOM that uses only the Azuma- Hoeffding concentration inequality, denoted ZOM, performed substantially worse than its counterpart for most time horizons and response functions. Staircase performed poorly in our experiments, as $mu_* \neq 0.5$. It is a known limitation for the method, as its performance tend to worsen in this case (Wichmann and Hill, 2001b). Questplus, a Bayesian method, converged significantly slower than non-bayesian methods such as DOS or ZOOM. Moreover, Questplus is unable to estimate the response function ψ_2, as it significantly differs from its parametric model. Conversely, GP and DOS were able to estimate all the tested response functions, but still suffers from sub-optimal performance for small T – highlighting the slow rate of convergence of its more general model. Overall this experiment highlights the advantage of ZOOM, which has the best – or close to the best – regret in all settings for all considered time budgets.

5. Conclusion

In this work, we introduced a new method for solving the threshold estimation problem, ZOOM. Compared to previous methods, we showed that ZOOM has the best of both worlds: it has a strong regret upper bounds, is model-agnostic and performed the best or close to the best in all our experiments, even for small sampling budget. As a result, ZOOM has the advantages of other approaches such as DOS without their drawbacks. Consequently, we argue that it is currently the best off-the-shelf method to solve the threshold estimation problem.

4. Note that $\psi_1(s_*) = \mu_*$, that ψ_1 is not differentiable on s_* but satisfies Assumption 1. s_* was chosen randomly for each set of experiment.

References

Jean-Yves Audibert, Sébastien Bubeck, et al. Minimax policies for adversarial and stochastic bandits. In *COLT*, volume 7, pages 1–122, 2009.

Julien Audiffren. Dichotomous optimistic search to quantify human perception. In *International Conference on Machine Learning*, pages 414–424. PMLR, 2021a.

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

Julien Audiffren and Jean-Pierre Bresciani. Model based or model free? comparing adaptive methods for estimating thresholds in neuroscience. *Neural Computation*, 34(2):338–359, 2022.

Julien Audiffren, Jean-Luc Bloechle, and Jean-Pierre Bresciani. Influence of mental workload on motion perception: A direct comparison of luminance-based and contrast-based stimuli. *Vision Research*, 193:107977, 2022.

Peter Auer. Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine Learning Research*, 3(Nov):397–422, 2002.

Peter Auer, Ronald Ortner, and Csaba Szepesvári. Improved Rates for the Stochastic Continuum-Armed Bandit Problem. In Nader H. Bshouty and Claudio Gentile, editors, *Learning Theory*, Lecture Notes in Computer Science, pages 454–468, Berlin, Heidelberg, 2007. Springer. ISBN 978-3-540-72927-3.

Beate Averbeck, Lena Seitz, Florian P Kolb, and Dieter F Kutz. Sex differences in thermal detection and thermal pain threshold and the thermal grill illusion: a psychophysical study in young volunteers. *Biology of sex differences*, 8(1):1–13, 2017.

Peter L Bartlett, Victor Gabillon, and Michal Valko. A simple parameter-free and adaptive approach to optimization under a minimal local smoothness assumption. In *Algorithmic Learning Theory*, pages 184–206. PMLR, 2019.

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

Sébastien Bubeck, Rémi Munos, and Gilles Stoltz. Pure exploration in finitely-armed and continuous-armed bandits. *Theoretical Computer Science*, 412(19):1832–1852, April 2011. ISSN 0304-3975. URL `http://www.sciencedirect.com/science/article/pii/S030439751000767X`.

Rui M Castro and Robert D Nowak. Upper and lower error bounds for active learning. In *The 44th Annual Allerton Conference on Communication, Control and Computing*, volume 2, page 1, 2006.

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

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

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

WOUTER A Dreschler and J Verschuure. *Psychophysical evaluation of fast compression systems*. World Scientific Singapore, 1996.

Miguel A. García-Pérez and Rocío Alcalá-Quintana. 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. URL https://onlinelibrary.wiley.com/doi/abs/10.1348/000711006X104596.

Jacob Gardner, Gustavo Malkomes, Roman Garnett, Kilian Q Weinberger, Dennis Barbour, and John P Cunningham. Bayesian active model selection with an application to automated audiometry. *Advances in neural information processing systems*, 28:2386–2394, 2015a.

Jacob R Gardner, Xinyu Song, Kilian Q Weinberger, Dennis L Barbour, and John P Cunningham. Psychophysical detection testing with bayesian active learning. In *UAI*, pages 286–295, 2015b.

Aurélien Garivier, Hédi Hadiji, Pierre Ménard, and Gilles Stoltz. Kl-ucb-switch: optimal regret bounds for stochastic bandits from both a distribution-dependent and a distribution-free viewpoints. *Journal of Machine Learning Research*, 23(179):1–66, 2022.

Jean-Bastien Grill, Michal Valko, and Rémi Munos. Black-box optimization of noisy functions with unknown smoothness. December 2015. URL https://hal.inria.fr/hal-01222915.

Christian Hatzfeld, Viet Quoc Hoang, and Mario Kupnik. It's All About the Subject - Options to Improve Psychometric Procedure Performance. In Fernando Bello, Hiroyuki Kajimoto, and Yon Visell, editors, *Haptics: Perception, Devices, Control, and Applications*, volume 9774, pages 394–403. Springer International Publishing, Cham, 2016. ISBN 978-3-319-42320-3 978-3-319-42321-0. URL http://link.springer.com/10.1007/978-3-319-42321-0-36.

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

Robert Kleinberg, Aleksandrs Slivkins, and Eli Upfal. Multi-armed bandits in metric spaces. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 681–690, 2008.

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

T Langheinrich, L Tebartz Van Elst, WA Lagreze, M Bach, CH Lücking, and Mark W Greenlee. Visual contrast response functions in parkinson's disease: evidence from electroretinograms, visually evoked potentials and psychophysics. *Clinical neurophysiology*, 111(1):66–74, 2000.

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

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

Andrea Locatelli and Alexandra Carpentier. Adaptivity to smoothness in x-armed bandits. In *Conference on Learning Theory*, pages 1463–1492. PMLR, 2018.

Rémi Munos. Optimistic optimization of a deterministic function without the knowledge of its smoothness. *Advances in neural information processing systems*, 24, 2011.

Jonathan Peirce, Jeremy R. Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindeløv. PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1):195–203, February 2019. ISSN 1554-3528. URL https://doi.org/10.3758/s13428-018-01193-y.

Stanisław Saks. Theory of the integral. 1937.

Jaelle Scheuerman, Kristen Brent Venable, Maxwell T Anderson, and Edward J Golob. Modeling spatial auditory attention: Handling equiprobable attended locations. In *CAID@ IJCAI*, pages 1–7, 2017.

Xuedong Shang, Emilie Kaufmann, and Michal Valko. General parallel optimization a without metric. In *Algorithmic Learning Theory*, pages 762–788. PMLR, 2019a.

Xuedong Shang, Emilie Kaufmann, and Michal Valko. General parallel optimization a without metric. In *Algorithmic Learning Theory*, pages 762–788, March 2019b. URL http://proceedings.mlr.press/v98/xuedong19a.html.

Yi Shen and Virginia M. Richards. 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. URL https://asa.scitation.org/doi/full/10.1121/1.4733540.

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

Léonard Torossian, Aurélien Garivier, and Victor Picheny. X-armed bandits: Optimizing quantiles, cvar and other risks. In *Asian Conference on Machine Learning*, pages 252–267. PMLR, 2019.

Alexander B Tsybakov. Optimal aggregation of classifiers in statistical learning. *The Annals of Statistics*, 32(1):135–166, 2004.

Alexander B Tsybakov. *Introduction to Nonparametric Estimation*. Springer New York, NY, 2009.

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

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

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

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

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

Di Yuan, Tiesong Zhao, Yiwen Xu, Hong Xue, and Liqun Lin. Visual jnd: a perceptual measurement in video coding. *IEEE Access*, 7:29014–29022, 2019.

Tilmann Zwicker. Psychoacoustics as the basis for modern audio signal data compression. *The Journal of the Acoustical Society of America*, 107(5):2875–2875, April 2000. ISSN 0001-4966. URL https://asa.scitation.org/doi/abs/10.1121/1.428677.

Appendix A. Regret Upper Bound: Proof of Theorem 1

We begin by proving some useful Lemmas.

Lemma 2 *Let* $d, n, k \geq 0$. *Then if* $\mu_{d,n,k} \leq \mu_*$,

$$\mathbb{P}\Big(\hat{\mu}_{d,n,k} \geq \mu_* + \mathcal{B}_T(N_{d,n,k}(t))\Big) < \frac{2}{T^3}$$

Proof

Note that

$$\{\hat{\mu}_{d,n,k}(T) > \mu_* + \mathcal{B}_T(N_{d,n,k}(t))\} \subset \{\hat{\mu}_{d,n,k}(T) > \mu_{d,n,k} + \mathcal{B}_T(N_{d,n,k}(t))\}$$

Hence the resulting using the Chernoff Hoeffding concentration inequality .

■

In other words, Lemma 2 states that the probability of ZOOM to reach the wrong decision about an arm is controlled. In the following, we examine the behavior of ZOOM on $\mathcal{A}^*$, where by definition ZOOM never reaches the wrong conclusion regarding arms comparisons when using (4). Formally, let

$$\mathcal{A}^* \doteq \{\forall t \leq T, \forall i \leq \kappa, |\mu_{d,n,k} - \hat{\mu}_{d,n,k}(t)| \leq \mathcal{B}_T(N_{d,n,k}(t))\} ,$$

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 $\mathbb{P}_{\mathcal{A}^*}(\mathcal{E}) = 1$. Under $\mathcal{A}^*$, ZOOM never reach a wrong subgrid when using (4). Therefore ZOOM only explores one subgrid per depth, and we can prove that the event $\mathcal{A}^*$ has high probability.

Lemma 3 $\mathbb{P}(\mathcal{A}^*) \geq 1 - \frac{2}{T}$.

Proof This directly results from Lemma 2 by taking the union bound. ■

In the following, we only refer to subgrids that are reached using (4), and we call d_* the deepest grid reached. As a consequence of the previous lemma, ZOOM only reaches one subgrid per level, and it is the one that contains s_*. Since only one grid per level is reached, in the following the drop the n from the notation to ease the reading : w denote by $\mathcal{G}_d$, the subgrid of depth d reached, and $s_{d,k}$ its arms. The convergence speed depends on the smoothness of ψ. The following Lemma details the smoothness property of ψ.

Lemma 4 *Let* ψ *be a response function that satisfies Assumption 1. Then* $\exists C_1 > 0$*, such that:*

$$\forall s \in \mathbb{I}, \quad |\psi(s_*) - \psi(s)| \leq C_1 |s_* - s| \tag{7}$$

Moreover $\exists C_2 > 0$*, such that:*

$$\forall s \in \mathbb{I}, \quad |\psi(s_*) - \psi(s)| \geq C_2 |s_* - s| \tag{8}$$

Proof We start by proving (7). By definition of the Dini derivatives, we have :

$$\lim_{s \to s_*} \frac{\psi(s) - \psi(s_*)}{s - s_*} \leq \underbrace{\max\left(D^- \psi(s_*), D^+ \psi(s_*)\right)}_{\overline{D}}$$

$$\lim_{s \to s_*} \frac{\psi(s) - \psi(s_*)}{s - s_*} \geq \underbrace{\min\left(D_-\psi(s_*), D_+\psi(s_*)\right)}_{\overline{D}}$$

Now by assumption $\overline{D}, \underline{D} < \infty$. By definition of $\overline{D}$, there exists $\delta > 0$, s.t. $\forall s \in [s_* - \delta, s_* + \delta]$

$$\left| \frac{\psi(s) - \psi(s_*)}{s - s_*} \right| \leq 2\overline{D}.$$

Moreover, $\forall s \notin [s_* - \delta, s_* + \delta]$, we trivially have

$$\left| \frac{\psi(s) - \psi(s_*)}{s - s_*} \right| \leq \frac{1}{\delta}.$$

Hence (7) with $C_1 = \max(2\overline{D}, 1/\delta)$. We now prove (8). By Assumption 1, we have $\underline{D}, \overline{D} > 0$. Thus there exists $\delta > 0$, s.t. $\forall s \in [s_* - \delta, s_* + \delta]$

$$\left| \frac{\psi(s) - \psi(s_*)}{s - s_*} \right| \geq \frac{1}{2}\underline{D}.$$

Moreover, $\forall s \notin [s_* - \delta, s_* + \delta]$, we trivially have

$$\left| \frac{\psi(s) - \psi(s_*)}{s - s_*} \right| \geq \frac{\delta \underline{D}}{2}.$$

Hence (8) with $C_2 = \min(\underline{D}/2, \delta\underline{D}/2)$.

■

Corollary 5 *If Assumption 1 is true, then* $\forall d \leq d_*$*,*

$$\mathbb{P}_{\mathcal{A}^*}\left(\Delta_{d,k} \leq C_1 K^{-d+1}\right) = 1 \tag{9}$$

Proof This corollary is an immediate consequence of Lemma 4. ■

Finally, the following lemma provides bounds to the number of times that an arm will be pulled under $\mathcal{A}^*$

Lemma 6 $\forall d, k \geq 0$, $\mathcal{A}^*$ *almost surely,*

$$N_{d,k} \leq \frac{3 \log T}{2\Delta_{d,k}^2}.$$

Proof This directly result from the inequality (4) and the definition of $\mathcal{B}_T$. ■

PROOF OF THEOREM 1

We are now ready to prove the regret upper bounds of Theorem 1. Note that since the regret is always upper bounded by 1, Lemma 3 implies that it is only necessary to prove the upper bounds on $\mathcal{A}^*$.

Proposition 7 *Assume that ψ satisfies Assumption 1. Let C_1, C_2 as in Lemma 4. Let $K = \lfloor\sqrt{\frac{T}{(\log T)(\log\log T)}}\rfloor$. Then, there exists a constant $C_\psi > 0$, that only depends on ψ, such that $\mathcal{A}^*$-almost surely,*

$$\mathcal{R}_T \leq C_\psi \sqrt{\frac{\log T \log\log T}{T}}$$

Proof Under $\mathcal{A}^*$, only valid grids are reached, and thus ZOOM only explore one grid per depth. For any $d \geq 1$, let $\mathcal{G}_d$ be the valid subgrid of depth d, and $s_{d,0}, \dots, s_{d,K}$ its arms. let d_* be the largest depth reached by ZOOM. First, if $d_* > 1$, i.e. if the algorithm has zoomed, then all arms on the deepest layers satisfy

$$|s_{d_*,k} - s_*| \leq \frac{1}{K},$$

and thus Corollary 5 implies that

$$\Delta_{d_*,k} = |\psi(s_{d_*,k}) - \psi(s_*)| \leq C_1|s_{d_*,k} - s_*| \leq \frac{C_1}{K} = C_1 \left\lfloor\sqrt{\frac{T}{\log T \log\log T}}\right\rfloor^{-1}. \tag{10}$$

Now if $d_* = 1$, the algorithm has never zoomed, and we will show that the arm that was pulled the most (which is the arm returned by the algorithm when the time budget elapsed) is a good estimator of s_*. To ease the notation, in the rest of the proof we will drop the d index (since $d = 1$). For any $0 \leq i \leq K$, let

$$\mathcal{K}_i = \left\{0 \leq k \leq K, \text{ s.t. } \quad \frac{i}{K} \leq |s_k - s_*| < \frac{i+1}{K}\right\}$$

Since the s_k are an uniform grid over $\mathbb{I}$ of step $1/K$, we have

$$\forall 0 \leq i \leq K, \quad |\mathcal{K}_i| \leq 2 \tag{11}$$

$$|\mathcal{K}_0| \geq 1 \tag{12}$$

Note that any arms in $\mathcal{K}_0$ satisfy (10) and there is at least one arm of $\mathcal{K}_0$. Now let $K \geq i > 0$, and $k \in \mathcal{K}_i$. By definition of $\mathcal{K}_i$, we have $0 < \frac{i}{K} \leq |s_k - s_*|$, and thus using (8),

$$\Delta_k \geq C_2 \frac{i}{K}$$

Now, by definition of $\mathcal{A}^*$, $\mathcal{B}_T(N_{d,k} - 1) > \Delta_k$ (otherwise ZOOM would have stopped to sample $s_{d,k}$ before) and thus

$$N_k \leq \underbrace{\frac{3\log T}{2i^2}\left(C_2^{-2}K^2\right)}_{\doteq n_i}.$$

So the total number of pulls that ZOOM uses on the $\mathcal{K}_i$, $i > 1$ is upper bounded by

$$\begin{aligned}\sum_{i=1}^{K}\sum_{k\in\mathcal{K}_i} N_k &\le \sum_{i=1}^{K}\sum_{k\in\mathcal{K}_i} n_i \le 2\sum_{i=1}^{K} n_i \\ &\le 2\sum_{i=1}^{K}\frac{3\log T}{2C_2^2 i^2}K^2 \le \frac{\pi^2\log T}{2C_2^2}K^2 \le \frac{\pi^2 T}{2C_2^2\log\log T} = o(T)\end{aligned}$$

Now since the algorithm has not zoomed, we have

$$T = \sum_{i=0}^{K}\sum_{k\in\mathcal{K}_i} N_k \le \sum_{k\in\mathcal{K}_0} N_k + \underbrace{\frac{\pi^2\log T}{2C_2^2}K^2}_{=o(T)}$$

and thus for T large enough,

$$\sum_{k\in\mathcal{K}_0} N_k \ge 2/3$$

and thus the conclusion using (11) and (10). ■