ISSN 1066-5307, Mathematical Methods of Statistics, 2018, Vol. 27, No. 1, pp. 1–31.

Boundary Crossing Probabilities for General Exponential Families

O.-A. Maillard[1*]

[1] *Inria Lille—Nord Europe, Villeneuve d'Ascq, France*

Received June 12, 2017; in final form, December 22, 2017

Abstract—We consider parametric exponential families of dimension K on the real line. We study a variant of *boundary crossing probabilities* coming from the multi-armed bandit literature, in the case when the real-valued distributions form an exponential family of dimension K. Formally, our result is a concentration inequality that bounds the probability that $\mathcal{B}^\psi(\hat{\theta}_n, \theta^\star) \geq f(t/n)/n$, where $\theta^\star$ is the parameter of an unknown target distribution, $\hat{\theta}_n$ is the empirical parameter estimate built from n observations, ψ is the log-partition function of the exponential family and $\mathcal{B}^\psi$ is the corresponding Bregman divergence. From the perspective of stochastic multi-armed bandits, we pay special attention to the case when the boundary function f is logarithmic, as it is enables to analyze the regret of the state-of-the-art `KL-ucb` and `KL-ucb+` strategies, whose analysis was left open in such generality. Indeed, previous results only hold for the case when $K = 1$, while we provide results for arbitrary finite dimension K, thus considerably extending the existing results. Perhaps surprisingly, we highlight that the proof techniques to achieve these strong results already existed three decades ago in the work of T. L. Lai, and were apparently forgotten in the bandit community. We provide a modern rewriting of these beautiful techniques that we believe are useful beyond the application to stochastic multi-armed bandits.

Keywords: exponential families, Bregman concentration, multi-armed bandits, optimality.

2000 Mathematics Subject Classification: 62L10, 62L12, 62C12, 62F03.

DOI: 10.3103/S1066530718010015

1. MULTI-ARMED BANDIT SETUP AND NOTATION

Let us consider a stochastic multi-armed bandit problem $(\mathcal{A}, \nu)$, where $\mathcal{A}$ is a finite set of cardinality $A \in \mathbb{N}$ and $\nu = (\nu_a)_{a\in\mathcal{A}}$ is a set of probability distributions over $\mathbb{R}$ indexed by $\mathcal{A}$. The game is sequential and goes as follows:

At each round $t \in \mathbb{N}$, the player picks an arm a_t (based on past observations) and receives a stochastic payoff Y_t drawn independently at random according to the distribution ν_{a_t}. The player only observes the payoff Y_t, and the goal is to maximize the expected cumulated payoff, $\sum_{t=1} Y_{a_t}$, over a possibly unknown number of steps.

Although the term multi-armed bandit problem was probably coined during the 60's with reference to the casino slot machines of the 19th century, the formulation of this problem is due to Herbert Robbins — one of the most brilliant minds of his time, see [18], and takes its origin in earlier questions about optimal stopping policies for clinical trials, see [20, 21, 22]. We refer the interested reader to [12] regarding the legacy of the immense work of H. Robbins in mathematical statistics for the sequential design of experiments, compiling his most outstanding research for his 70's birthday. Since then, the field of multi-armed bandits has grown large and bold, and we humbly refer to the introduction of [13] for key historical aspects about the development of the field. Most notably, they include first the introduction of dynamic allocation indices (or Gittins indices, [12]) suggesting that an optimal strategy can be found in the form of an index strategy (that at each round selects an arm with highest "index"); second, the seminal work of Lai and Robbins [14] that showed that indexes can be chosen as "upper confidence bounds" on the mean reward of each arm and provided the first asymptotic lower bound on the achievable performance for specific distributions; third, the generalization of this lower bound in the 90's to generic distributions

*E-mail: odalricambrym.maillard@inria.fr

by Burnetas and Katehakis [7] (see also the recent work [11]) as well as the asymptotic analysis [11] of generic classes of upper-confidence-bound based index policies and finally [4] that popularized a simple sub-optimal index strategy termed `UCB` and most importantly opened the quest for finite-time, as opposed to asymptotic, performance guarantees. For the purpose of this paper, we now remind the formal definitions and notation for the stochastic multi-armed bandit problem following [8].

Quality of a strategy. For each arm $a \in \mathcal{A}$, let μ_a be the expectation of the distribution ν_a, and let $a^\star$ be any optimal arm in the sense that

$$a^\star \in \operatorname*{Argmax}_{a\in\mathcal{A}} \mu_a.$$

We write $\mu^\star$ as a short-hand notation for the largest expectation $\mu_{a^\star}$ and denote the *gap* of the expected payoff μ_a of an arm a to $\mu^\star$ as $\Delta_a = \mu^\star - \mu_a$. In addition, we denote the number of times each arm a is pulled between the rounds 1 and T by $N_a(T)$,

$$N_a(T) \stackrel{\text{def}}{=} \sum_{t=1}^{T} \mathbf{1}_{\{a_t=a\}}.$$

Definition 1 (Expected regret)**.** The quality of a strategy is evaluated using the notion of expected regret (or simply, regret) at round $T \geq 1$, defined as

$$\mathfrak{R}_T \stackrel{\text{def}}{=} \mathbb{E}\Big[T\mu^\star - \sum_{t=1}^{T} Y_t\Big] = \mathbb{E}\Big[T\mu^\star - \sum_{t=1}^{T} \mu_{a_t}\Big] = \sum_{a\in\mathcal{A}} \Delta_a\, \mathbb{E}\big[N_a(T)\big], \tag{1}$$

where we used the tower rule for the first equality. The expectation is with respect to the random draws of the Y_t according to the ν_{a_t} and to the possible auxiliary randomization introduced by the decision-making strategy.

Empirical distributions. We denote empirical distributions in two related ways, depending on whether random averages indexed by the global time t or averages of given numbers t of pulls of a given arm are considered. The first series of averages will be referred to by using a functional notation for the indexation in the global time: $\widehat{\nu}_a(t)$, while the second series will be indexed with the local times t in subscripts: $\widehat{\nu}_{a,t}$. These two related indexations, functional for global times and random averages versus subscript indexes for local times, will be consistent throughout the paper for all quantities at hand, not only empirical averages.

Definition 2 (Empirical distributions)**.** For each $m \geq 1$, we denote by $\tau_{a,m}$ the round at which arm a was pulled for the mth time, that is,

$$\tau_{a,m} = \min\big\{t \in \mathbb{N} \colon N_a(t) = m\big\}.$$

For each round t such that $N_a(t) \geq 1$, we then define the following two empirical distributions

$$\widehat{\nu}_a(t) = \frac{1}{N_a(t)} \sum_{s=1}^{t} \delta_{Y_s} \mathbf{1}_{\{a_s=a\}} \quad \text{and} \quad \widehat{\nu}_{a,n} = \frac{1}{n}\sum_{m=1}^{n} \delta_{X_{a,m}}, \quad \text{where} \quad X_{a,m} \stackrel{\text{def}}{=} Y_{\tau_{a,m}}.$$

where δ_x denotes the Dirac distribution on $x \in \mathbb{R}$.

Lemma 1. *The random variables $X_{a,m} = Y_{\tau_{a,m}}$, $m = 1, 2, \ldots$, are independent and identically distributed according to ν_a. Moreover, we have $\widehat{\nu}_a(t) = \widehat{\nu}_{a,N_a(t)}$.*

Proof. For means based on local times we consider the filtration $(\mathcal{F}_t)$, where for all $t \geq 1$, the σ-algebra $\mathcal{F}_t$ is generated by $a_1, Y_1, \ldots, a_t, Y_t$. In particular, a_{t+1} and all $N_a(t+1)$ are $\mathcal{F}_t$-measurable. Likewise, $\big\{\tau_{a,m} = t\big\}$ is $\mathcal{F}_{t-1}$-measurable. That is, each random variable $\tau_{a,m}$ is a (predictable) stopping time. Hence the result follows by a standard result in probability theory (see, e.g., Section 5.3 in [9]). □

2. BOUNDARY CROSSING PROBABILITIES FOR THE GENERIC KL-ucb STRATEGY

The first appearance of the `KL-ucb` strategy can be traced at least to [15] although it was not given an explicit name at that time. It seems the strategy was forgotten after the work [4] that opened a decade of intensive research on finite-time analysis of bandit strategies and extensions to variants of the problem ([2, 3], see also [6] for a survey of relevant variants of bandit problems), until the work of Honda and Takemura [13] shed a novel light on the asymptotically optimal strategies. Thanks to their illuminating work, the first finite-time regret analysis of `KL-ucb` was obtained in [17] for discrete distributions, soon extended to handle exponential families of dimension 1 as well, in the unifying work [8]. However, as we will see in this paper, we should all be much in debt to the outstanding work of T.L. Lai regarding the analysis of this index strategy, both asymptotically and in finite-time, as a second look at his papers shows how to bypass the limitations of the state-of-the-art regret bounds for the control of *boundary crossing probabilities* in this context (see Theorem 3 below). Actually, the first focus of the present paper is not stochastic bandits but boundary crossing probabilities, and the bandit setting that we provide here should be considered only as giving a solid motivation for the contribution of this paper.

Let us now introduce formally the `KL-ucb` strategy. We assume that the learner is given a family $\mathcal{D} \subset \mathfrak{M}_1(\mathbb{R})$ of probability distributions that satisfies $\nu_a \in \mathcal{D}$ for each arm $a \in \mathcal{A}$, where $\mathfrak{M}_1(\mathcal{X})$ denotes the set of all probability distributions over the set $\mathcal{X}$. For two distributions $\nu, \nu' \in \mathfrak{M}_1(\mathbb{R})$, we denote by $\mathtt{KL}(\nu, \nu')$ their Kullback–Leibler divergence and by $E(\nu)$ and $E(\nu')$ their expectations. (This operator is denoted by E while expectations of a function f with respect to underlying randomizations are referred to as $\mathbb{E}[f]$ or $\mathbb{E}_{X\sim\nu}[f(X)]$ to make explicit the law of the random variable X).

The generic form of the algorithm of interest in this paper is described as Algorithm 1. It relies on two parameters: an operator $\Pi_{\mathcal{D}}$ (in spirit, a projection operator) that associates with each empirical distribution $\widehat{\nu}_a(t)$ an element of the model $\mathcal{D}$; and a nondecreasing function f, which is typically such that $f(t) \approx \log(t)$.

At each round $t \geq A+1$ (recall that $A = |\mathcal{A}|$), an upper confidence bound $U_a(t)$ is associated with the expectation μ_a of the distribution ν_a of each arm; an arm a_{t+1} with highest upper confidence bound is then played.

Algorithm 1. The `KL-ucb` algorithm (generic form).

Parameters: An operator $\Pi_{\mathcal{D}} \colon \mathfrak{M}_1(\mathbb{R}) \to \mathcal{D}$; a nondecreasing function $f \colon \mathbb{N} \to \mathbb{R}$

Initialization: Pull each arm of $\{1, \dots, A\}$ once

for each round $t+1$, where $t \geq A$, **do**

for each arm a compute the quantity

$$U_a(t) = \sup\left\{E(\nu)\colon \nu \in \mathcal{D} \quad \text{and} \quad \mathtt{KL}\Big(\Pi_{\mathcal{D}}\big(\widehat{\nu}_a(t)\big),\, \nu\Big) \leq \frac{f(t)}{N_a(t)}\right\};$$

pick an arm $\quad a_{t+1} \in \operatorname*{argmax}_{a\in\mathcal{A}} U_a(t).$

In the literature, another variant of `KL-ucb` is introduced where the term $f(t)$ is replaced with $f(t/N_a(t))$. We refer to this algorithm as `KL-ucb+`. While `KL-ucb` has been analyzed and shown to be provably near-optimal, the variant `KL-ucb+` has not been analyzed yet.

Alternative formulation of `KL-ucb`. We wrote the `KL-ucb` algorithm so that the optimization problem resulting from the computation of $U_a(t)$ is easy to handle. Now, under some assumption, one can rewrite this term in an equivalent form more suited for the analysis. We refer to [8]:

Assumption 1. *There is a known interval $\mathcal{I} \subset \mathbb{R}$ with boundary $\mu^- \leq \mu^+$, for which each model $\mathcal{D} = \mathcal{D}_a$ of probability measures is contained in $\mathfrak{M}_1(\mathcal{I})$ and such that $\forall \nu \in \mathcal{D}_a\ \forall \mu \in \mathcal{I} \setminus \{\mu^+\}$,*

$$\inf\big\{KL(\nu, \nu')\colon \nu' \in \mathcal{D}_a \ \text{ s.t. } \ E(\nu') > \mu\big\} = \min\big\{KL(\nu, \nu')\colon \nu' \in \mathcal{D}_a \ \text{ s.t. } \ E(\nu') \geq \mu\big\}.$$

Lemma 2 (Rewriting). *Under Assumption 1, the upper bound used by the* KL-ucb *algorithm satisfies the following equality*

$$U_a(t) \;=\; \max\left\{\mu\in\mathcal{I}\setminus\{\mu^+\}\colon \mathcal{K}_a\big(\Pi_a(\hat\nu_a(t)),\mu\big)\le \frac{f(t)}{N_a(t)}\right\}$$

$$\textit{where}\quad \mathcal{K}_a(\nu_a,\mu^\star)\stackrel{\text{def}}{=}\inf_{\nu\in\mathcal{D}_a\colon E(\nu)>\mu^\star} \mathtt{KL}(\nu_a,\nu)\quad\textit{and}\quad \Pi_a\stackrel{\text{def}}{=}\Pi_{\mathcal{D}_a}.$$

Likewise, a similar result holds for KL-ucb+ *whith $f(t)$ replaced by $f(t/N_a(t))$.*

Remark 1. Assumption 1 is valid, for instance, when $\mathcal{D}_a=\mathfrak{M}_1([0,1])$ and $\mathcal{I}=[0,1]$. Indeed, we can replace the strict inequality with an inequality provided that $\mu<1$ by [13], and the infimum is attained by lower semi-continuity of the KL divergence and convexity and closure of the set $\{\nu'\in\mathfrak{M}_1([0,1])\;\text{s.t.}\;E(\nu')\ge\mu\}$.

Using boundary-crossing probabilities for regret analysis. We continue by restating a convenient way to decompose the regret and make appear the *boundary crossing probabilities* that are at the heart of this paper. The following lemma is a direct adaptation from [8].

Lemma 3 (From Regret to Boundary Crossing Probabilities). *Let $\epsilon\in\mathbb{R}^+$ be a small constant such that $\epsilon\in(0,\min\{\,\mu^\star-\mu_a,\,a\in\mathcal{A}\,\})$. For $\mu,\gamma\in\mathbb{R}$, introduce the set*

$$\mathcal{C}_{\mu,\gamma}=\big\{\nu'\in\mathfrak{M}_1(\mathbb{R})\colon \mathcal{K}_a(\Pi_a(\nu'),\mu)<\gamma\big\}.$$

Then the number of pulls of a sub-optimal arm $a\in\mathcal{A}$ by Algorithm KL-ucb *satisfies*

$$\mathbb{E}[N_T(a)]\le 2+\inf_{n_0\le T}\left\{n_0+\sum_{n\ge n_0+1}^{T}\mathbb{P}\big\{\hat\nu_{a,n}\in\mathcal{C}_{\mu^\star-\epsilon,f(T)/n}\big\}\right\}$$
$$+\sum_{t=|\mathcal{A}|}^{T-1}\underbrace{\mathbb{P}\big\{N_{a^\star}(t)\;\mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\hat\nu_{a^\star,N_{a^\star}(t)}),\,\mu^\star-\epsilon\big)>f(t)\big\}}_{\textit{Boundary Crossing Probability}}.$$

Likewise, the number of pulls of a sub-optimal arm $a\in\mathcal{A}$ by Algorithm KL-ucb+ *satisfies*

$$\mathbb{E}[N_T(a)]\le 2+\inf_{n_0\le T}\left\{n_0+\sum_{n\ge n_0+1}^{T}\mathbb{P}\big\{\hat\nu_{a,n}\in\mathcal{C}_{\mu^\star-\epsilon,f(T/n)/n}\big\}\right\}$$
$$+\sum_{t=|\mathcal{A}|}^{T-1}\underbrace{\mathbb{P}\big\{N_{a^\star}(t)\;\mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\hat\nu_{a^\star,N_{a^\star}(t)}),\,\mu^\star-\epsilon\big)>f(t/N_{a^\star}(t))\big\}}_{\textit{Boundary Crossing Probability}}.$$

Proof. The first part of this lemma for KL-ucb is proved in [8]. The second part that is about KL-ucb+ can be proved straightforwardly following the very same lines. We thus only provide the main steps here for clarity. For $\epsilon>0$ satisfying $\epsilon<\min\{\,\mu^\star-\mu_a,\,a\in\mathcal{A}\,\}$, consider the following inclusion of events:

$$\big\{a_{t+1}=a\big\}\subseteq\big\{\mu^\star-\epsilon<U_a(t)\text{ and }a_{t+1}=a\big\}\;\cup\;\big\{\mu^\star-\epsilon\ge U_{a^\star}(t)\big\}.$$

Indeed, on the event $\big\{a_{t+1}=a\big\}\cap\big\{\mu^\star-\epsilon<U_{a^\star}(t)\big\}$, we have $\mu^\star-\epsilon<U_{a^\star}(t)\le U_a(t)$ (where the last inequality is by definition of the strategy). Moreover, note that

$$\big\{\mu^\star-\epsilon<U_a(t)\big\}\subseteq\Big\{\exists\nu'\in\mathcal{D}\colon E(\nu')>\mu^\star-\epsilon\text{ and }N_a(t)\;\mathcal{K}_a\big(\Pi_a(\hat\nu_{a,N_a(t)}),\,\mu^\star-\epsilon\big)\le f(t/N_a(t))\Big\},$$

$$\text{and}\quad\big\{\mu^\star-\epsilon\ge U_{a^\star}(t)\big\}\subseteq\Big\{\exists\nu'\in\mathcal{D}\colon N_{a^\star}(t)\;\mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\hat\nu_{a^\star,N_{a^\star}(t)}),\,\mu^\star-\epsilon\big)>f(t/N_{a^\star}(t))\Big\},$$

since $\mathcal{K}_a$ is a nondecreasing function in its second argument and $\mathcal{K}_a\big(\nu, E(\nu)\big) = 0$ for all distributions ν. Therefore we have the following decomposition:

$$\mathbb{E}[N_T(a)] \le 1 + \sum_{t=|\mathcal{A}|}^{T-1} \mathbb{P}\big\{N_{a^\star}(t)\ \mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\hat\nu_{a^\star,N_{a^\star}(t)}),\ \mu^\star - \epsilon\big) > f(t/N_{a^\star}(t))\big\}$$

$$+ \sum_{t=|\mathcal{A}|}^{T-1} \mathbb{P}\big\{N_a(t)\ \mathcal{K}_a\big(\Pi_a(\hat\nu_{a,N_a(t)}),\ \mu^\star - \epsilon\big) \le f(t/N_a(t)) \text{ and } a_{t+1} = a\big\}.$$

Using the remaining steps of the proof of the result from [8], equation (10) can now be straightforwardly modified to work with $f(t/N_a(t))$ instead of $f(t)$, thus concluding this proof. □

Lemma 3 shows that two terms need to be controlled in order to derive regret bounds for the considered strategy. The *boundary crossing probability* term is arguably the most difficult to handle and is the focus of the next sections. The other term involves the probability that an empirical distribution belongs to a convex set, which can be handled either directly as in [8] or by resorting to finite-time Sanov-type results such as Theorem 2.1 and comments on p. 372 in [10] or its version in Lemma 1 in [17]. For completeness, we state the exact result from [10].

Lemma 4 (Non-asymptotic Sanov's lemma). *Let $\mathcal{C}$ be an open convex subset of $\mathfrak{M}_1(\mathcal{X})$ such that $\Lambda_\nu(\mathcal{C}) = \inf_{\kappa\in\mathcal{C}} \mathtt{KL}(\kappa, \nu)$ is finite. Then, for all $t \ge 1$, $\mathbb{P}_\nu\{\hat\nu_t \in \mathcal{C}\} \le \exp\big(-t\Lambda_\nu(\overline{\mathcal{C}})\big)$, where $\overline{\mathcal{C}}$ is the closure of $\mathcal{C}$.*

Scope and focus of this work. We focus on the setting of stochastic multi-armed bandits because this gives a strong and natural motivation for studying boundary crossing probabilities. However, one should understand that the primary goal of this paper is to give credit to the work of T.L. Lai regarding the neat understanding of boundary crossing probabilities and not necessarily to provide a regret bound for such bandit algorithms as `KL-ucb` or `KL-ucb+`. Also, we believe that results on boundary crossing probabilities are useful beyond the bandit problem in hypothesis testing. Thus, and in order to avoid obscuring the main result regarding boundary crossing probabilities, we choose not to provide regret bounds here and to leave them has an exercise for the interested reader; controlling the remaining term appearing in the decomposition of Lemma 3 is indeed mostly technical and does not seem to require especially illuminating or fancy idea. We refer to [8] for an example of bound in the case of exponential families of dimension 1.

High-level overview of the contribution. We are now ready to explain the main results of this paper. For the purpose of clarity, we provide them as an informal statement before proceeding with the technical material.

Our contribution is about the behavior of the *boundary crossing probability* term for exponential families of dimension K when choosing the threshold function $f(x) = \log(x) + \xi\log\log(x)$. Our result reads as follows.

Theorem (Informal statement). *Assuming that the observations are generated from a distribution that belongs to an exponential family of dimension K that satisfies some mild conditions, for any nonnegative ϵ and some class-dependent but fully explicit constants c, C (also depending on ϵ) it holds*

$$\mathbb{P}\big\{N_{a^\star}(t)\ \mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\hat\nu_{a^\star,N_{a^\star}(t)}),\ \mu^\star - \epsilon\big) > f(t)\big\} \le \frac{C}{t}\log(t)^{K/2-\xi} e^{-c\sqrt{f(t)}},$$

$$\mathbb{P}\big\{N_{a^\star}(t)\ \mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\hat\nu_{a^\star,N_{a^\star}(t)}),\ \mu^\star - \epsilon\big) > f(t/N_{a^\star}(t))\big\} \le \frac{C}{t}\log(tc)^{K/2-\xi-1},$$

where the first inequality holds for all t and the second one for large enough $t \ge t_c$, where t_c is class dependent but explicit and "reasonably" small.

We provide the rigorous statement in Theorem 3 and Corollaries 1 and 2 below. The main interest of this result is that it shows how to tune ξ with respect to the dimension K of the family. Indeed, in order to ensure that the probability term is summable in t, the bound suggests that ξ should be at least larger than $K/2-1$. The case of exponential families of dimension 1 ($K=1$) is especially interesting, as it supports the fact that both `KL-ucb` and `KL-ucb+` can be tuned using $\xi=0$ (and even negative ξ for `KL-ucb`). This was observed in numerical experiments in [8] although not theoretically supported until now.

The rest of the paper is organized as follows. Section 3 provides the required background and notation about exponential families, Section 4 provides the precise statements as well as previous results, Section 5 details the proof of Theorem 3, and Section 6 details the proofs of Corollaries 1 and 2.

3. GENERAL EXPONENTIAL FAMILIES, PROPERTIES AND EXAMPLES

Before focusing on the boundary crossing probabilities, we require a few tools and definitions related to exponential families. The purpose of this section is thus to present them and prepare for the main result of this paper. In this section, for a set $\mathcal{X}\subset\mathbb{R}$, we consider a multivariate function $F\colon\mathcal{X}\to\mathbb{R}^K$ and denote $\mathcal{Y}=F(\mathcal{X})\subset\mathbb{R}^K$.

Definition 3 (Exponential families)**.** The exponential family generated by the function F and the reference measure ν_0 on the set $\mathcal{X}$ is

$$\mathcal{E}(F;\nu_0)=\big\{\nu_\theta\in\mathfrak{M}_1(\mathcal{X});\ \forall x\in\mathcal{X}\ \nu_\theta(dx)=\exp\big(\langle\theta,F(x)\rangle-\psi(\theta)\big)\nu_0(dx),\ \theta\in\mathbb{R}^K\big\},$$

where $\psi(\theta)\stackrel{\text{def}}{=}\log\int_{\mathcal{X}}\exp\big(\langle\theta,F(x)\rangle\big)\,\nu_0(dx)$ is the normalization function (or log-partition function) of the exponential family. The vector θ is called the vector of canonical parameters. The parameter set of the family is the domain $\Theta_{\mathcal{D}}\stackrel{\text{def}}{=}\big\{\theta\in\mathbb{R}^K;\ \psi(\theta)<\infty\big\}$, and the invertible parameter set of the family is $\Theta_I\stackrel{\text{def}}{=}\big\{\theta\in\mathbb{R}^K;\ 0<\lambda_{\mathtt{MIN}}(\nabla^2\psi(\theta))\le\lambda_{\mathtt{MAX}}(\nabla^2\psi(\theta))<\infty\big\}\subset\Theta_{\mathcal{D}}$, where $\lambda_{\mathtt{MIN}}(M)$ and $\lambda_{\mathtt{MAX}}(M)$ denote the minimum and maximum eigenvalues of a positive semi-definite matrix M.

Remark 2. When $\mathcal{X}$ is compact, which is the usual assumption in multi-armed bandits ($\mathcal{X}=[0,1]$) and F is continuous, then we automatically get $\Theta_{\mathcal{D}}=\mathbb{R}^K$.

In the sequel, we always assume that the family is regular, that is $\Theta_{\mathcal{D}}$ has a nonempty interior. Another key assumption is that the parameter $\theta^\star$ of the optimal arm belongs to the interior of Θ_I and is away from its boundary, which essentially avoids degenerate distributions, as we illustrate below.

Examples. Bernoulli distributions form an exponential family with $K=1$, $\mathcal{X}=\{0,1\}$, $F(x)=x$, $\psi(\theta)=\log(1+e^\theta)$. The Bernoulli distribution with mean μ has parameter $\theta=\log(\mu/(1-\mu))$. Note that $\Theta_{\mathcal{D}}=\mathbb{R}$ and that degenerate distributions with mean 0 or 1 correspond to parameters $\pm\infty$.

Gaussian distributions on $\mathcal{X}=\mathbb{R}$ form an exponential family with $K=2$, $F(x)=(x,x^2)$, and for each $\theta=(\theta_1,\theta_2)$, $\psi(\theta)=-\frac{\theta_1^2}{4\theta_2}+\frac{1}{2}\log\big(-\frac{\pi}{\theta_2}\big)$. The Gaussian distribution $\mathcal{N}(\mu,\sigma^2)$ has parameter $\theta=\big(\frac{\mu}{\sigma^2},-\frac{1}{2\sigma^2}\big)$. It is immediate to check that $\Theta_{\mathcal{D}}=\mathbb{R}\times\mathbb{R}_\star^-$. Degenerate distributions with variance 0 correspond to a parameter θ with both infinite components, while as θ approaches the boundary $\mathbb{R}\times\{0\}$, the variance tends to infinity. It is natural to consider only parameters that correspond to a not too large variance.

3.1. Bregman Divergence Induced by the Exponential Family

An interesting property of exponential families is the following straightforward rewriting of the Kullback–Leibler divergence:

$$\forall\theta,\theta'\in\Theta_{\mathcal{D}},\quad \mathtt{KL}(\nu_\theta,\nu_{\theta'})=\langle\theta-\theta',\mathbb{E}_{X\sim\nu_\theta}(F(X))\rangle-\psi(\theta)+\psi(\theta').$$

In particular, the vector $\mathbb{E}_{X\sim\nu_\theta}(F(X))$ is called the vector of *dual* (*or expectation*) *parameters*. It is equal to the vector $\nabla\psi(\theta)$. Note that $\mathtt{KL}(\nu_\theta,\nu_{\theta'})=\mathcal{B}^\psi(\theta,\theta')$, where $\mathcal{B}^\psi$ is known as the Bregman divergence with potential function ψ and is defined (see [5] for further details) by

$$\mathcal{B}^\psi(\theta,\theta')\stackrel{\text{def}}{=}\psi(\theta')-\psi(\theta)-\langle\theta'-\theta,\nabla\psi(\theta)\rangle.$$

Thus, if Π_a is chosen to be the projection on the exponential family $\mathcal{E}(F;\nu_0)$ and ν is a distribution with projection given by $\nu_\theta = \Pi_a(\nu)$, then we can rewrite the definition of $\mathcal{K}_a$ in the simpler form

$$\mathcal{K}_a\big(\Pi_a(\nu),\, \mu\big) = \inf\big\{\, \mathcal{B}^\psi(\theta,\theta');\, \mathbb{E}_{\nu_{\theta'}}(X) > \mu \,\big\}. \tag{2}$$

We continue by providing a powerful rewriting of the Bregman divergence.

Lemma 5 (Bregman duality). *For all $\theta^\star \in \Theta_\mathcal{D}$ and $\eta \in \mathbb{R}^K$ such that $\theta^\star + \eta \in \Theta_\mathcal{D}$, let $\Phi(\eta) = \psi(\theta^\star + \eta) - \psi(\theta^\star)$. Further, let us introduce the Fenchel–Legendre dual of Φ defined by*

$$\Phi^\star(y) = \sup_{\eta\in\mathbb{R}^K} \langle \eta, y\rangle - \Phi(\eta).$$

Then $\log \mathbb{E}_{X\sim\nu_{\theta^\star}} \exp\big(\langle \eta, F(X)\rangle\big) = \Phi(\eta)$. *Further, for all F such that $F = \nabla\psi(\theta)$ for some $\theta \in \Theta_\mathcal{D}$, one has $\Phi^\star(F) = \mathcal{B}^\psi(\theta,\theta^\star)$.*

Lemma 6 (Bregman and Smoothness). *The following inequalities hold true*

$$\mathcal{B}^\psi(\theta,\theta') \le \frac{\|\theta-\theta'\|^2}{2}\sup\{\lambda_{MAX}(\nabla^2\psi(\tilde\theta));\, \tilde\theta \in [\theta,\theta']\},$$
$$\|\nabla\psi(\theta) - \nabla\psi(\theta')\| \le \sup\{\lambda_{MAX}(\nabla^2\psi(\tilde\theta));\, \tilde\theta \in [\theta,\theta']\}\, \|\theta-\theta'\|,$$
$$\mathcal{B}^\psi(\theta,\theta') \ge \frac{\|\theta-\theta'\|^2}{2}\inf\{\lambda_{MIN}(\nabla^2\psi(\tilde\theta));\, \tilde\theta \in [\theta,\theta']\},$$
$$\|\nabla\psi(\theta) - \nabla\psi(\theta')\| \ge \inf\{\lambda_{MIN}(\nabla^2\psi(\tilde\theta));\, \tilde\theta \in [\theta,\theta']\}\, \|\theta-\theta'\|,$$

where $\lambda_{MAX}(\nabla^2\psi(\tilde\theta))$ and $\lambda_{MIN}(\nabla^2\psi(\tilde\theta))$ are the largest and smallest eigenvalues of $\nabla^2\psi(\tilde\theta)$.

Proof of Lemma 5. The second equality holds by simple algebra. Now the first equality is immediate, since

$$\log \mathbb{E}_{\theta^\star} \exp(\langle \eta, F(X)\rangle) = \log \int \exp(\langle \eta, F(x)\rangle + \langle \theta^\star, F(x)\rangle - \psi(\theta^\star))\nu_0(dy) = \psi(\eta+\theta^\star) - \psi(\theta^\star).$$

□

Proof of Lemma 6. We have by definition that $\mathcal{B}^\psi(\theta,\theta') = \psi(\theta) - \psi(\theta') - \langle \theta-\theta', \nabla\psi(\theta')\rangle$. Then, by a Taylor expansion, there exists $\tilde\theta' \in [\theta,\theta']$ such that

$$\psi(\theta) = \psi(\theta') + \langle \theta-\theta', \nabla\psi(\theta')\rangle + \frac{1}{2}(\theta-\theta')^T\nabla^2\psi(\tilde\theta)(\theta-\theta').$$

Likewise, there exists $\tilde\theta \in [\theta,\theta']$ such that $\nabla\psi(\theta) = \nabla\psi(\theta') + \nabla^2\psi(\tilde\theta)(\theta-\theta')$. □

3.2. Dual Formulation of the Optimization Problem

Using Bregman divergence enables us to rewrite the K-dimensional optimization problem (2) in a slightly more convenient form thanks to a dual formulation. Indeed, introducing a Lagrangian parameter $\lambda \in \mathbb{R}^+$ and using Karush–Kuhn–Tucker conditions, one gets the following necessary optimality conditions

$$\nabla\psi(\theta') - \nabla\psi(\theta) - \lambda\partial_{\theta'}\mathbb{E}_{\nu_{\theta'}}(X) = 0 \quad \text{with} \quad \lambda(\mu - \mathbb{E}_{\nu_{\theta'}}(X)) = 0, \quad \lambda \ge 0, \quad \mathbb{E}_{\nu_{\theta'}}(X) \ge \mu,$$

and by definition of the exponential family, we can make use of the fact that

$$\partial_{\theta'}\mathbb{E}_{\nu_{\theta'}}(X) = \mathbb{E}_{\nu_{\theta'}}(XF(X)) - \mathbb{E}_{\nu_{\theta'}}(X)\nabla\psi(\theta') \in \mathbb{R}^K,$$

where $X \in \mathbb{R}$ and $F(X) \in \mathbb{R}^K$. Combining these two equations, we obtain the system

$$\begin{cases} \nabla\psi(\theta')(1+\lambda\mathbb{E}_{\nu_{\theta'}}(X)) - \nabla\psi(\theta) - \lambda\mathbb{E}_{\nu_{\theta'}}(XF(X)) = 0 \in \mathbb{R}^K \\ \text{with } \lambda(\mu - \mathbb{E}_{\nu_{\theta'}}(X)) = 0, \quad \lambda \ge 0, \quad \mathbb{E}_{\nu_{\theta'}}(X) \ge \mu. \end{cases} \tag{3}$$

For minimal exponential family, this system admits for each fixed θ, μ a unique solution in θ', which we write for clarity $\theta(\lambda^\star;\theta,\mu)$ to indicate its dependence on the optimal value $\lambda^\star$ of the dual parameter as well as the constraints.

Remark 3. For $\theta \in \Theta_I$, when the optimal value of λ is $\lambda^\star = 0$, then $\nabla\psi(\theta') = \nabla\psi(\theta)$ and thus $\theta' = \theta$, which is only possible if $\mathbb{E}_{\nu_\theta}(X) \geq \mu$. Thus whenever $\mu > \mathbb{E}_{\nu_\theta}(X)$, the dual constraint is active, i.e., $\lambda > 0$, and we get the vector equation

$$\nabla\psi(\theta')(1+\lambda\mu) - \nabla\psi(\theta) - \lambda\mathbb{E}_{\nu_{\theta'}}(XF(X)) = 0 \quad \text{and} \quad \mathbb{E}_{\nu_{\theta'}}(X) = \mu.$$

The example of discrete distributions. In many cases, the above optimization problem reduces to a simpler one-dimensional optimization problem, where we optimize over the dual parameter λ. We illustrate this phenomenon by a family of discrete distributions. Let $\mathbb{X} = \{x_1, \ldots, x_K, x_\star\}$ be a set of distinct real numbers. Without loss of generality, assume that $x_\star > \max_{k\leq K} x_k$. The family of distributions p with support in $\mathbb{X}$ is a specific K-dimensional family. Indeed, let F be the feature function with kth component $F_k(x) = \mathbf{1}\{x = x_k\}$, for all $k \in \{1,\ldots,K\}$. Then the parameter $\theta = (\theta_k)_{1\leq k\leq K}$ of the distribution $p = p_\theta$ has components $\theta_k = \log(\frac{p(x_k)}{p(x_\star)})$ for all k. Note that $p(x_k) = \exp(\theta_k - \psi(\theta))$ for all k, and $p(x_0) = \exp(-\psi(\theta))$. Then it follows that $\psi(\theta) = \log(\sum_{k=1}^K e^{\theta_k} + 1)$, $\nabla\psi(\theta) = (p(x_1), \ldots, p(x_K))^\top$ and $\mathbb{E}(XF_k(X)) = x_k p_\theta(x_k)$. Further, $\Theta_{\mathcal{D}} = (\mathbb{R}\cup\{-\infty\})^K$ and $\theta \in \Theta_{\mathcal{D}}$ corresponds to the condition $p_\theta(x_\star) > 0$. Now, for a nontrivial value μ such that $\mathbb{E}_{p_\theta}(X) < \mu < x_\star$, it can be readily checked that the system (3) specialized to this family is equivalent (with no surprise) to the one considered for instance in [13] for discrete distributions. After some tedious but simple steps detailed in [13], one obtains the following easy-to-solve *one-dimensional* optimization problem (see also [8]), although the family is of dimension K:

$$\mathcal{K}_a\big(\Pi_a(\nu),\, \mu\big) = \mathcal{K}_a\big(\nu_\theta,\, \mu\big) = \sup\Big\{\sum_{x\in\mathbb{X}} p_\theta(x)\log\Big(1 - \lambda\frac{x-\mu}{x_\star - \mu}\Big);\ \lambda\in[0,1]\Big\}.$$

3.3. Empirical Parameter and Definition

In this section we discuss the definition of the empirical parameter corresponding to the projection of the empirical distribution on the exponential family. While this is innocuous for most settings, in full generality, one needs to take some specific care to ensure that all the objects we deal with are well defined and that all parameters θ we talk about indeed belong to the set Θ_D (or better Θ_I).

An important property is that if the family is regular, then $\nabla\psi(\Theta_{\mathcal{D}})$ is an open set that coincides with the interior of realizable values of $F(x)$ for $x \sim \nu$ for any ν absolutely continuous with respect to ν_0. In particular, by convexity of the set $\nabla\psi(\Theta_{\mathcal{D}})$ this means that the empirical average $\frac{1}{n}\sum_{i=1}^n F(X_i) \in \mathbb{R}^K$ belongs to $\overline{\nabla\psi(\Theta_{\mathcal{D}})}$ for all $\{X_i\}_{i\leq n} \sim \nu_\theta$ with $\theta \in \Theta_{\mathcal{D}}$. Thus, for the observed samples $X_1, \ldots, X_n \in \mathcal{X}$ coming from $\nu_{a^\star}$, the projection $\Pi_{a^\star}(\hat\nu_{a^\star,n})$ on the family can be represented by a sequence $\{\hat\theta_{n,m}\}_{m\in\mathbb{N}} \in \Theta_{\mathcal{D}}$ such that

$$\nabla\psi(\hat\theta_{n,m}) \stackrel{m}{\to} \hat F_n \ \text{ where } \ \hat F_n \stackrel{\text{def}}{=} \frac{1}{n}\sum_{i=1}^n F(X_i) \in \mathbb{R}^K.$$

In the sequel, we want to ensure that whenever $\nu_{a^\star} = \nu_{\theta^\star}$ with $\theta^\star \in \mathring\Theta_I$, we have $\hat F_n \in \nabla\psi(\mathring\Theta_I)$, which means that there is a unique $\hat\theta_n \in \mathring\Theta_I$ such that $\nabla\psi(\hat\theta_n) = \hat F_n$, or equivalently $\hat\theta_n = \nabla\psi^{-1}(\hat F_n)$. To this end, we assume that $\theta^\star$ is away from the boundary of Θ_I. In many cases, it is then sufficient to assume that n is larger than a small constant (roughly K) to ensure that we can find a unique $\hat\theta_n \in \mathring\Theta_I$ such that $\nabla\psi(\hat\theta_n) = \hat F_n$.

Example. Let us consider Gaussian distributions on $\mathcal{X} = \mathbb{R}$ with $K = 2$. We consider a parameter $\theta^\star = (\frac{\mu}{\sigma^2}, -\frac{1}{2\sigma^2})$ corresponding to a Gaussian finite mean μ and positive variance σ^2. Now, for any $n \geq 2$, the empirical mean $\hat\mu_n$ is finite and the empirical variance $\hat\sigma_n^2$ is positive, and thus $\theta_n = \nabla\psi^{-1}(\hat F_n)$ is well defined.

The case of Bernoulli distributions is interesting as it shows a slightly different situation. Let us consider a parameter $\theta^\star = \log(\mu/(1-\mu))$ corresponding to the Bernoulli distribution with mean μ. Before $\hat F_n$ can be mapped to a point in $\mathring\Theta_I = \mathbb{R}$, one needs to wait that the number of observations for both 0 and 1 is positive. Whenever $\mu \in (0,1)$, the probability that this does not happen is controlled by $\mathbb{P}(n_0(n)=0 \text{ or } n_1(n)=0) = \mu^n + (1-\mu)^n \le 2\max(\mu, 1-\mu)^n$, where $n_x(n)$ denotes the number of observations of symbol $x \in \{0,1\}$ after n samples. For $\mu \ge 1/2$, the latter quantity is less than $\delta_0 \in (0,1)$ for $n \ge \frac{\log(2/\delta_0)}{\log(1/\mu)}$, which depends on the probability level δ_0 and cannot be considered to be especially small when μ is close[1] to 1. That is, even when the parameter $\hat{\theta}_n$ does not belong to $\mathbb{R}$, the event $n_0(n)=0$ corresponds to having empirical mean equal to 1. This is a favorable situation since any optimistic algorithm should pull the corresponding arm. Thus we only need to control $\mathbb{P}(n_1(n)=0) = (1-\mu)^n$, which is less than $\delta_0 \in (0,1)$ for $n \ge \frac{\log(1/\delta_0)}{\log(1/(1-\mu))}$, which is essentially a constant. As a matter of illustration, when $\delta = 10^{-3}$ and $\mu = 0.9$, this condition is met for $n \ge 3$.

Following this discussion, we assume in the sequel that n is always large enough so that $\hat{\theta}_n = \nabla\psi^{-1}(\hat{F}_n) \in \mathring{\Theta}_I$ can be uniquely defined. Now, to discuss the separation between the parameter and the boundary more formally, we introduce the following definition.

Definition 4 (Enlarged parameter set). Let $\Theta \subset \Theta_{\mathcal{D}}$ and $\rho > 0$ be a constant. The enlargement of size ρ of Θ in Euclidean norm (i.e. the ρ-neighborhood) is defined by

$$\Theta_\rho \stackrel{\text{def}}{=} \Big\{\theta \in \mathbb{R}^K; \inf_{\theta' \in \Theta_{\mathcal{D}}} \|\theta - \theta'\| < \rho\Big\}.$$

Further, for each ρ such that $\Theta_\rho \subset \Theta_I$ we introduce the quantities

$$v_\rho = v_{\Theta_\rho} \stackrel{\text{def}}{=} \inf_{\theta \in \Theta_\rho} \lambda_{\texttt{MIN}}(\nabla^2\psi(\theta)), \qquad V_\rho = V_{\Theta_\rho} \stackrel{\text{def}}{=} \sup_{\theta \in \Theta_\rho} \lambda_{\texttt{MAX}}(\nabla^2\psi(\theta)).$$

Using the notion of enlarged parameter set, we highlight an especially useful property to prove concentration inequalities, summarized in the following result.

Lemma 7 (Log-Laplace control). *Let $\Theta \subset \Theta_{\mathcal{D}}$ be a convex set and $\rho > 0$ be such that $\theta^\star \in \Theta_\rho \subset \Theta_I$. Then, for all $\eta \in \mathbb{R}^K$ such that $\theta^\star + \eta \in \Theta_\rho$, it holds*

$$\log \mathbb{E}_{\theta^\star} \exp(\eta^\top F(X)) \le \eta^\top \nabla\psi(\theta^\star) + \frac{V_\rho}{2}\|\eta\|^2.$$

Proof. It holds by simple algebra

$$\begin{aligned}\log \mathbb{E}_{\theta^\star} \exp(\eta^\top F(X)) &= \psi(\theta^\star + \eta) - \psi(\theta^\star)\\ &\le \eta^\top \nabla\psi(\theta^\star) + \max_{\theta \in H(\theta^\star+\eta, \theta^\star)} \frac{1}{2}\eta^\top \nabla^2\psi(\theta)\eta \le \eta^\top \nabla\psi(\theta^\star) + \frac{V_\rho}{2}\|\eta\|^2,\end{aligned}$$

where $H(\theta, \theta') = \{\alpha\theta + (1-\alpha)\theta', \alpha \in [0,1]\}$. The equality holds by definition and basic rewriting. In the inequalities, we used that Θ_ρ is convex as an enlargement of a convex set, and thus that $H(\eta + \theta^\star, \theta^\star) \subset \Theta_\rho$. □

In the sequel, we are interested in sets Θ such that $\Theta_\rho \subset \mathring{\Theta}_I$ for some specific ρ. This comes essentially from the fact that we require some room around Θ and Θ_I to ensure all quantities remain finite and well defined. Before proceeding, it is convenient to introduce the notation $d(\Theta', \Theta) = \inf_{\theta \in \Theta, \theta' \in \Theta'} \|\theta - \theta'\|$, as well as the Euclidean ball $B(y, \delta) = \{y' \in \mathbb{R}^K : \|y' - y\| \le \delta\}$. Using this notation, the following lemma whose proof is immediate provides conditions which are satisfied in all future technical considerations.

[1] This also suggests replacing $\hat{F}_n$ with a Laplace or Krichevsky–Trofimov estimates that provide initial bonus to each symbol and, as a result, map any $\hat{F}_n$ for $n \ge 0$ to a parameter in $\hat{\theta}_n \in \mathbb{R}$.

Lemma 8 (Well-defined parameters). *Let $\theta^\star \in \mathring{\Theta}_I$ and $\rho^\star = d(\{\theta^\star\}, \mathbb{R}^K \setminus \Theta_I) > 0$. Now for any convex set $\Theta \subset \Theta_I$ such that $\theta^\star \in \Theta$ and $d(\Theta, \mathbb{R}^K \setminus \Theta_I) = \rho^\star$ and any $\rho < \rho^\star/2$, it holds $\Theta_{2\rho} \subset \mathring{\Theta}_I$. Further, for any δ such that $\hat{F}_n \in B(\nabla\psi(\theta^\star), \delta) \subset \nabla\psi(\Theta_\rho)$ there exists $\hat{\theta}_n \in \Theta_\rho \subset \mathring{\Theta}_I$ such that $\nabla\psi(\hat{\theta}_n) = \hat{F}_n$.*

In the sequel, we will restrict our analysis to the slightly more restrictive case when $\hat{\theta}_n \in \Theta_\rho$ with $\Theta_{2\rho} \subset \mathring{\Theta}_I$. This is mostly for convenience and to avoid dealing with rather specific situations.

Remark 4. Again let us remind that when $\mathcal{X}$ is compact and F is continuous, then $\Theta_I = \Theta_{\mathcal{D}} = \mathbb{R}^K$.

Illustration. We now illustrate the definition of v_ρ and V_ρ. For Bernoulli distributions with parameter $\mu \in [0,1]$, $\nabla\psi(\theta) = 1/(1+e^{-\theta})$ and $\nabla^2\psi(\theta) = e^{-\theta}/(1+e^{-\theta})^2 = \mu(1-\mu)$. Thus, v_ρ is away from 0 whenever Θ_ρ excludes the means μ close to 0 or 1 and $V_\rho \leq 1/4$.

Now $\psi(\theta) = -\frac{\theta_1^2}{4\theta_2} + \frac{1}{2}\log\left(\frac{-\pi}{\theta_2}\right)$, where $\theta = (\frac{\mu}{\sigma^2}, -\frac{1}{2\sigma^2})$, for a family of Gaussian distributions with unknown mean and variance. Thus, $\nabla\psi(\theta) = (-\frac{\theta_1}{2\theta_2}, \frac{\theta_1^2}{4\theta_2^2} - \frac{1}{2\theta_2})$, $\nabla^2\psi(\theta) = (-\frac{1}{2\theta_2}, \frac{\theta_1}{2\theta_2^2}; \frac{\theta_1}{2\theta_2^2}, -\frac{\theta_1^2}{2\theta_2^3} + \frac{1}{2\theta_2^2}) = 2\mu\sigma^2(\frac{1}{2\mu}, 1; 1, 2\mu + \frac{\sigma^2}{\mu})$. The smallest eigenvalue is larger than $\sigma^4/(1/2 + \sigma^2 + 2\mu^2)$ and the largest is upper bounded by $\sigma^2(1 + 2\sigma^2 + 4\mu^2)$, which enables us to control V_ρ and v_ρ.

4. BOUNDARY CROSSING FOR K-DIMENSIONAL EXPONENTIAL FAMILIES

In this section, we study the boundary crossing probability term appearing in Lemma 3 for a K-dimensional exponential family $\mathcal{E}(F;\nu_0)$. We overview the existing results before detailing our main contribution. As explained in the introduction, the key technical tools that enable us to obtain the novel results were already known three decades ago, and thus even though the novel result is impressive due to its generality and tightness, it should be regarded as a modernized version of an existing but almost forgotten result that enables us to solve a few long-lasting open questions as a by-product.

4.1. Previous Work on Boundary-Crossing Probabilities

The existing results used in the bandit literature about boundary-crossing probabilities are restricted to a few specific cases. For instance in [8], the authors provide the following control.

Theorem 1 (`KL-ucb`). *In the case of canonical (that is $F(x) = x$) exponential families of dimension $K = 1$, it holds for $f(x) = \log(x) + \xi\log\log(x)$ that for all $t > A$*

$$\mathbb{P}_{\theta^\star}\Big\{\bigcup_{n=1}^{t-A+1} n\,\mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\hat{\nu}_{a^\star,n}),\,\mu^\star\big) > f\big(t\big) \cap \mu_{a^\star} > \hat{\mu}_{a^\star,n}\Big\} \leq e\lceil f(t)\log(t)\rceil e^{-f(t)}.$$

Further, in the special case of distributions with finitely many K atoms, it holds for all $t > A$, $\epsilon > 0$

$$\mathbb{P}_{\theta^\star}\Big\{\bigcup_{n=1}^{t-A+1} n\mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\hat{\nu}_{a^\star,n}),\,\mu^\star - \epsilon\big) > f\big(t\big)\Big\} \leq e^{-f(t)}\big(3e + 2 + 4\epsilon^{-2} + 8e\epsilon^{-4}\big).$$

In contrast to [16], the authors of [8] provide an asymptotic control in the more general case of exponential families of dimension K with some basic regularity condition, as we explained earlier. We now restate this beautiful result from [16] in a way that is suitable for a more direct comparison with other results.

Theorem 2 (Lai [16]). *Consider an exponential family of dimension K. Define for $\gamma > 0$ the cone $\mathcal{C}_\gamma(\theta) = \{\theta' \in \mathbb{R}^K : \langle \theta', \theta\rangle \geq \gamma\|\theta\| \cdot \|\theta'\|\}$. Then, for $f(x) = \alpha\log(x) + \xi \log\log(x)$ it holds for all $\theta^\dagger \in \Theta$ such that $\|\theta^\dagger - \theta^\star\|^2 \geq \delta_t$, where $\delta_t \to 0$, $t\delta_t \to \infty$ as $t \to \infty$,*

$$\mathbb{P}_{\theta^\star}\Big\{\bigcup_{n=1}^{t} \hat\theta_n \in \Theta_\rho \cap n\mathcal{B}^\psi(\hat\theta_n, \theta^\dagger) \geq f\Big(\frac{t}{n}\Big) \cap \nabla\psi(\hat\theta_n) - \nabla\psi(\theta^\dagger) \in \mathcal{C}_\gamma(\theta^\dagger - \theta^\star)\Big\}$$

$$\overset{t\to\infty}{=} O\Big(t^{-\alpha}\|\theta^\dagger - \theta^\star\|^{-2\alpha}\log^{-\xi-\alpha+K/2}(t\|\theta^\dagger - \theta^\star\|^2)\Big)$$

$$= O\Big(e^{-f(t\|\theta^\dagger - \theta^\star\|^2)}\log^{-\alpha+K/2}(t|\theta^\dagger - \theta^\star|^2)\Big).$$

Discussion. The quantity $\mathcal{B}^\psi(\hat\theta_n, \theta^\dagger)$ is the direct analog of $\mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\hat\nu_{a^\star,n}), \mu^\star - \epsilon\big)$ in Theorem 1. Note however that $f(t/n)$ replaces the larger quantity $f(t)$, which means that Theorem 2 controls a larger quantity than Theorem 1, and is thus in this sense stronger. It also holds for general exponential families of dimension K. Another important difference is the order of magnitude of the right-hand side terms of both theorems. Indeed, since $e\lceil f(t)\log(t)\rceil e^{-f(t)} = O(\frac{\log^{2-\xi}(t)+\xi\log(t)^{1-\xi}\log\log(t)}{t})$, Theorem 1requires that $\xi > 2$ in order that this term is $o(1/t)$, and $\xi > 0$ for the second term of Theorem 1. In contrast, Theorem 2 shows that it is enough to consider $f(x) = \log(x) + \xi\log\log(x)$ with $\xi > K/2 - 1$ to ensure a $o(1/t)$ bound. For $K = 1$, this means we can even use $\xi > -1/2$ and in particular $\xi = 0$, which corresponds to the value they recommend in the experiments.

Thus Theorem 2 improves over Theorem 1 in three ways: it is an extension to dimension K, it provides a bound for $f(t/n)$ (and thus for `KL-ucb+`) and not only $f(t)$, and finally allows for smaller values of ξ. These improvements are partly due to the fact that Theorem 1 controls a concentration with respect to $\theta^\dagger$, not $\theta^\star$, which takes advantage of the fact that there is some gap when going from $\mu^\star$ to distributions with mean $\mu^\star - \epsilon$. The proof of Theorem 2 directly takes advantage of this, contrary to that of the first part of Theorem 1.

On the other hand, Theorem 2 is only asymptotic whereas Theorem 1 holds for finite t. Furthermore, we notice two restrictions on the control event. First, it requires $\hat\theta_n \in \Theta_\rho$, but we showed in the previous section that this is a minor restriction. Second, there is the restriction to a cone $\mathcal{C}_\gamma(\theta^\dagger - \theta^\star)$ which simplifies the analysis, but is a more dramatic restriction. This restriction cannot be removed trivially since, as can be seen from the complete statement of Theorem 2 in [16], the right-hand side blows up to ∞ when $\gamma \to 0$. As we will see, it is possible to overcome this restriction by resorting to a smart covering of the space with cones, and sum the resulting terms via a union bound over the covering. We explain the precise way of proceeding in the proof of Theorem 3 in Section 5.

Hint about proving the first part of Theorem 1. It may be interesting to give a hint about the proof of the first part of Theorem 1, as it involves an elegant step, despite relying quite heavily on two specific properties of the canonical exponential family of dimension 1. Indeed in the special case of the canonical one-dimensional family (that is $K = 1$ and $F_1(x) = x \in \mathbb{R}$), $\hat F_n = \frac{1}{n}\sum_{i=1}^n X_i$ coincides with the empirical mean and it can be shown that $\Phi^\star(F)$ is strictly decreasing on $(-\infty, \mu^\star]$. Thus for any $F \leq \mu^\star$, it holds

$$\big\{\hat F_n \leq \mu^\star \cap \Phi^\star(\hat F_n) \geq \Phi^\star(F)\big\} \subset \big\{\hat F_n \leq F\big\}. \tag{4}$$

Further, using the notation of Section 3.1, it also holds in that case $\mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\hat\nu_{a^\star,n}),\ \mu^\star\big) = \mathcal{B}^\psi(\hat\theta_n, \theta^\star) = \Phi^\star(\hat F_n)$, where $\hat\theta_n = \dot\psi^{-1}(\hat F_n)$ is uniquely defined. A second nontrivial property that is shown in [8] is that for all $F \leq \mu^\star$, we can localize the supremum as

$$\Phi^\star(F) = \sup\big\{xF - \Phi(x) \colon x < 0 \text{ and } xF - \Phi(x) > 0\big\}. \tag{5}$$

Armed with these two properties, the proof reduces almost trivially to the following elegant lemma.

Lemma 9 (Dimension 1)**.** *Consider a canonical one-dimensional family (that is $K = 1$ and $F_1(x) = x \in \mathbb{R}$). Then, for all f such that $f(t/n)/n$ is nonincreasing in n,*

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{m\le n<M} \mathcal{B}^{\psi}(\hat{\theta}_n, \theta^\star) \ge f(t/n)/n \Big\} \le \exp\Big(-\frac{m}{M} f(t/M)\Big).$$

The proof of this lemma is provided in the Appendix and is directly adapted from the proof of Theorem 1. The first statement of Theorem 1 is obtained by a peeling argument, using $m/M = (f(t)-1)/f(t)$. However this argument does not seem to extend nicely to using $f(t/n)$, which explains why there is no statement regarding this threshold.

4.2. Main Results and Contributions

In this section, we provide several results on boundary crossing probabilities that we prove in detail in the next section. We first provide a nonasymptotic bound with explicit terms for the control of the boundary crossing probability term. We then provide two corollaries that can be used directly for the analysis of `KL-ucb` and `KL-ucb+` and that better highlight the asymptotic scaling of the bound with t, which helps seeing the effect of the parameter ξ on the bound.

Theorem 3 (Boundary crossing for exponential families)**.** *Let $\epsilon > 0$ and define $\rho_\epsilon = \inf\{\|\theta' - \theta\| \colon \mu_{\theta'} = \mu^\star - \epsilon, \mu_\theta = \mu^\star\}$. Let $\rho^\star = d(\{\theta^\star\}, \mathbb{R}^K \setminus \Theta_I)$ and $\Theta \subset \Theta_{\mathcal{D}}$ be a set such that $\theta^\star \in \Theta$ and $d(\Theta, \mathbb{R}^K \setminus \Theta_I) = \rho^\star$. Thus $\theta^\star \in \Theta \subset \Theta_\rho \subset \mathring{\Theta}_I$ for each $\rho < \rho^\star$. Assume that $n \to f(t/n)/n$ is nonincreasing and $n \to nf(t/n)$ is nondecreasing. Then, for every $b > 1, p, q, \eta \in [0,1]$, and $n_i = b^i$ if $i < I_t = \lceil \log_b(qt)\rceil$, $n_{I_t} = t+1$, it holds*

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{1\le n\le t} \hat{\theta}_n \in \Theta_\rho \cap \mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\hat{\nu}_{a^\star,n}), \mu^\star - \epsilon\big) \ge f(t/n)/n \Big\}$$
$$\le C(K,b,\rho,p,\eta) \sum_{i=0}^{I_t-1} \exp\bigg(-n_i\rho_\epsilon^2\alpha^2 - \rho_\epsilon\chi\sqrt{n_i f(t/n_i)} - f\Big(\frac{t}{n_{i+1}-1}\Big)\bigg) f\Big(\frac{t}{n_{i+1}-1}\Big)^{K/2},$$

where we introduced the constants $\alpha = \eta\sqrt{v_\rho/2}$, $\chi = p\eta\sqrt{2v_\rho^2/V_\rho}$ and

$$C(K,b,\rho,p,\eta) = C_{p,\eta,K}\bigg(2\frac{\omega_{p,K-2}}{\omega_{\max\{p,\frac{2}{\sqrt{5}}\},K-2}} \max\Big\{\frac{2bV_\rho^4}{p\rho^2 v_\rho^6}, \frac{V_\rho^3}{v_\rho^4}, \frac{b^2V_\rho^5}{pv_\rho^6(\frac{1}{2}+\frac{1}{K})}\Big\}^{K/2} + 1\bigg).$$

Here $C_{p,\eta,K}$ is the cone-covering number of $\nabla\psi\big(\Theta_\rho \setminus \mathcal{B}_2(\theta^\star, \rho_\epsilon)\big)$ with minimal angular separation p and not intersecting the set $\nabla\psi\big(\Theta_\rho \setminus \mathcal{B}_2(\theta^\star, \eta\rho_\epsilon)\big)$; and $\omega_{p,K} = \int_p^1 \sqrt{1-z^2}^K dz$ if $K \ge 0$, 1 else.

Remark 5. The same result holds when replacing all occurrences of $f(\cdot)$ by the constant $f(t)$.

Remark 6. In dimension 1, the theorem takes a simpler form. Indeed $C_{p,\eta,1} = 2$ for all $p, \eta \in (0,1)$ and thus, choosing $b = 2$ for instance, $C(1,2,\rho,p,\eta)$ reduces to $2\Big(2\max\Big\{\frac{2V_\rho^2}{\rho v_\rho^3}, \frac{V_\rho^{3/2}}{v_\rho^2}, \frac{2V_\rho^{5/2}}{v_\rho^3}\Big\} + 1\Big)$. In the case of Bernoulli distributions, if $\Theta_\rho = \{\log(\mu/(1-\mu)), \mu \in [\mu_\rho, 1-\mu_\rho]\}$, then $v_\rho = \mu_\rho(1-\mu_\rho)$, $V_\rho = 1/4$ and $C(1,2,\rho,p,\eta) = 2(\frac{1}{8\mu_\rho^3(1-\mu_\rho)^3} + 1)$.

Remark 7. We believe it is possible to reduce the max term by a factor V_ρ^3/v_ρ^4 in the definition of $C(K,b,\rho,p,\eta)$.

Let $f(x) = \log(x) + \xi\log\log(x)$. We now state two corollaries of Theorem 3. The first one is stated for the case when the boundary is $f(t)/n$ and is thus directly relevant to the analysis of `KL-ucb`. The second corollary is about the more challenging boundary $f(t/n)/n$ that corresponds to the `KL-ucb+` strategy. We note that f is nondecreasing only for $x \ge e^{-\xi}$. When $x = t$, this requires that $t \ge e^{-\xi}$.

Now, when $x = t/N_{a^\star}(t)$, where $N_{a^\star}(t) = t - O(\log(t))$, assuming that f is nondecreasing requires that $\xi \geq \log(1 - O(\log(t)/t))$ for large t, that is $\xi \geq 0$. In the sequel we thus restrict to $t \geq e^{-\xi}$ when using the boundary $f(t)$ and to $\xi \geq 0$ when using the boundary $f(t/n)$. Finally, we remind that the quantity $\chi = p\eta\sqrt{2v_\rho^2/V_\rho}$ is a function of p, η and ρ, and introduce the notation $\chi_\epsilon = \rho_\epsilon \chi$ for convenience.

Corollary 1 (Boundary crossing for $f(t)$). *Let $f(x) = \log(x) + \xi \log\log(x)$. Using the same notation as in Theorem 3, for all $p, \eta \in [0,1], \rho < \rho^\star$ and all $t \geq e^{-\xi}$ such that $f(t) \geq 1$ it holds*

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{1\leq n<t} \hat{\theta}_n \in \Theta_\rho \cap \mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\hat{\nu}_{a^\star,n}), \mu^\star - \epsilon\big) \geq f(t)/n \Big\}$$
$$\leq \frac{C(K,4,\rho,p,\eta)(1+\chi_\epsilon)}{\chi_\epsilon t}\left(1 + \xi\frac{\log\log(t)}{\log(t)}\right)^{K/2} \log(t)^{-\xi+K/2} e^{-\chi_\epsilon\sqrt{\log(t)+\xi\log\log(t)}}.$$

Corollary 2 (Boundary crossing for $f(t/n)$). *Let $f(x) = \log(x) + \xi \log\log(x)$. For all $p, \eta \in [0,1]$, $\rho < \rho^\star$ and $\xi \geq \max(K/2 - 1, 0)$, and for $t \in [85\chi^{-2}, t_\chi]$, where $t_\chi = \chi_\epsilon^{-2}\frac{\exp(\log(4.5)^2/\chi_\epsilon^2)}{4\log(4.5)^2}$, it holds*

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{1\leq n<t} \hat{\theta}_n \in \Theta_\rho \cap \mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\hat{\nu}_{a^\star,n}), \mu^\star - \epsilon\big) \geq f(t/n)/n \Big\} \leq C(K,4,\rho,p,\eta)\Bigg[e^{-\chi_\epsilon\sqrt{tc'}}$$
$$+ \frac{(1+\xi)^{K/2}}{ct\log(tc)}\begin{cases} \frac{16}{3}\log(tc\log(tc)/4)^{K/2-\xi} + 80\log(1.25)^{K/2-\xi} & \text{if } \xi \geq K/2, \\ \frac{16}{3}\log(t/3)^{K/2-\xi} + 80\log(t\frac{c\log(tc)}{4-c\log(tc)})^{K/2-\xi} & \text{if } \xi \in [K/2-1, K/2], \end{cases}\Bigg]$$

where $c = \chi_\epsilon^2/(2\log(5))^2$, and $c' = \sqrt{f(5)/5}$ if $\xi \geq K/2$ and $\sqrt{f(4)/4}$ else. Further, for larger values of t, $t \geq t_\chi$, the second term in the brackets becomes

$$\frac{(1+\xi)^{K/2}}{ct\log(tc)}\begin{cases} 144\log(1.25)^{K/2-\xi} & \text{if } \xi \geq K/2, \\ 144\log(t/3)^{K/2-\xi} & \text{if } \xi \in [K/2-1, K/2] \quad (\text{and } \xi \geq 0). \end{cases}$$

Remark 8. In Corollary 1, since the asymptotic regime of $\chi_\epsilon\sqrt{\log(t)} - (K/2 - \xi)\log\log(t)$ may take a massive amount of time to kick-in when $\xi < K/2 - 2\chi_\epsilon$, we recommend to take $\xi > K/2 - 2\chi_\epsilon$. Note also that the value $\xi = K/2 - 1/2$ is interesting in practice, since then $\log(t)^{K/2-\xi} = \sqrt{\log(t)} < 5$ for all $t \leq 10^9$.

Remark 9. The restriction to $t \geq 85\chi_\epsilon^{-2}$ is merely for $\xi \simeq K/2 - 1$. For instance for $\xi \geq K/2$, the restriction becomes $t \geq 76\chi_\epsilon^{-2}$, and it becomes less restrictive for larger ξ. The term t_χ is virtually infinite: for instance, when $\chi_\epsilon = 0.3$, this is already larger than 10^{12}, while $85\chi_\epsilon^{-2} < 945$.

Remark 10. According to this result, the value $K/2 - 1$ (when it is nonnegative) appears to be a critical value for ξ, since the boundary crossing probabilities are not summable in t for $\xi \leq K/2 - 1$, but are summable for $\xi > K/2 - 1$. Indeed, the terms behind the curved brackets are conveniently $o(\log(t))$ with respect to t, except when $\xi = K/2 - 1$. In practice however, since this asymptotic behavior may take a large time to kick-in, we recommend ξ to be away from $K/2 - 1$.

Remark 11. Achieving a bound for the threshold $f(t/N_a(t))$ is more challenging than for $f(t)$. Only the latter case was analyzed in [8] as the former was out of reach of their analysis. Also, the result is valid with exponential families of dimension K and not only dimension 1, which is a major improvement. It is interesting to note that when $K = 1$, $\max(K/2 - 1, 0) = 0$, and to observe experimentally that a sharp phase transition indeed appears for `KL-ucb+` precisely at the value $\xi = 0$: the algorithm suffers a linear regret when $\xi < 0$ and a logarithmic regret when $\xi = 0$. For `KL-ucb`, no sharp phase transition appears at point $\xi = 0$. Instead, a relatively smooth phase transition appears for a negative ξ dependent on the problem. Both observations are coherent with the statements of the corollaries.

Discussion regarding the proof technique. The proof technique that we consider below significantly differs from the proof from [8] and [13], and combines key ideas disseminated in two works from Tze Leung Lai [16] and [15] with some nontrivial extension that we describe below. Also, we simplify some of the original arguments and improve the readability of the initial proof technique, in order to shed more light on these neat ideas.

Change of measure. At a high level, the first big idea of this proof is to resort to a *change of measure* argument, which is classically used only to prove the *lower bound* on the regret. The work [16] should be given full credit for this idea. This is in stark contrast with the proof techniques developed later for the finite-time analysis of stochastic bandits. The change of measure is actually not used once, but twice. First, to go from $\theta^\star$, the parameter of the optimal arm to some perturbation of it $\theta^\star_c$. Then, which is perhaps more surprising, to go from this perturbed point to a mixture over a well-chosen ball centered in it. Although we have reasons to believe that this second change of measure may not be required (at least choosing a ball in dimension K seems slightly sub-optimal), this two-step localization procedure is definitely the first main component that enables us to handle the boundary crossing probabilities. The other steps for the proof of the Theorem include a concentration of measure argument and a peeling argument, which are more standard.

Bregman divergence. The second main idea is the use of Bregman divergence and its relation to the quadratic norm, which is due to [15]. This enables one to make explicit computations for exponential families of dimension K without too much effort, at the price of losing some "variance" terms (linked to the Hessian of the family). We combine this idea with some key properties of Bregman divergence that enables us to simplify a few steps, notably the concentration step, that we revisited entirely in order to obtain clean bounds valid in finite time and not only asymptotically.

Concentration of measure and boundary effects. One specific difficulty that appeared in the proof was to handle the shape of the parameter set Θ and the fact that $\theta^\star$ should be away from its boundary. The initial asymptotic proof of Lai did not account for this and was not entirely accurate. Going beyond this proved to be quite challenging due to the boundary effects, although the concentration result (Section 5.4, Lemma 15) that we obtain is eventually valid without restriction and the final proof looks deceptively easy. This concentration result is novel.

Cone covering and dimension K. In [16], the author analyzed a boundary crossing problem first in the case of exponential families of dimension 1, and then sketched the analysis for exponential families of dimension K and for the intersection with one cone. However the complete result was nowhere stated explicitly. As a matter of fact, the initial proof in [16] is restricted to a cone, which greatly simplifies the result. In order to obtain the full-blown results, valid in dimension K for the unrestricted event, we introduce a cone covering of the space. This seemingly novel (although not very fancy) idea enables us to get a final result that is only depending on the cone-covering number of the space. It required some careful considerations and simplifications of the initial steps of [16]. Along the way, we made explicit the sketch of proof provided in [16] for the dimension K.

Corollaries and ratios. The final key idea that should be credited to T.L. Lai is about the fine tuning of the final bound resulting from the two changes of measures, the application of concentration and the peeling argument. Indeed these steps lead to a bound by a sum of terms, say $\sum_{i=0}^{I} s_i$, that should be studied and depend on a few free parameters. This corresponds, with our rewriting and modifications, to the statement of Theorem 3. Here the brilliant idea of T.L. Lai, that we separate from the proof of Theorem 3 and use in the proof of Corollaries 1 and 2, is to bound the ratios s_{i+1}/s_i for small values of i and the ratios s_i/s_{i+1} for large values of i separately (instead of resorting, for instance, to a sum-integral comparison lemma). A careful study of these terms enables us to improve the scaling and allow for smaller values of ξ, up to $K/2-1$, while other approaches seem unable to go below $K/2+1$. Nevertheless, in our quest to obtain explicit bounds valid not only asymptotically but also in finite time, this step is quite delicate, since a naive approach easily requires huge values for t before the asymptotic regimes kick-in. By refining the initial proof strategy of [16], we managed to obtain a result valid for all t for the setting of Corollary 1 and for all "reasonably"[2] large t for the more challenging setting of Corollary 2.

---

[2]We require t to be at least about 10^2 times some problem-dependent constant, against a factor that could be e^{15} in the initial analysis.

5. ANALYSIS OF BOUNDARY CROSSING PROBABILITIES: PROOF OF THEOREM 3

In this section, we closely follow the proof technique used in [16] for the proof of Theorem 2 in order to prove the result of Theorem 3. We precise further the constants, remove the cone restriction on the parameter and modify the original proof to be fully nonasymptotic which, using the technique of [16], forces us to make some parts of the proof a little more accurate.

Let us recall that we consider Θ and ρ such that $\theta^\star \in \Theta_\rho \subset \mathring{\Theta}_I$. The proof is divided in four main steps that we briefly present here for clarity.

In Section 5.1, we take care of the random number of pulls of the arm by a peeling argument. Simultaneously, we introduce a covering of the space with cones, which allows for using arguments from the proof of Theorem 2.

In Section 5.2, we proceed with the first change of measure argument: taking advantage of the gap between $\mu^\star$ and $\mu^\star - \epsilon$, we move from a concentration argument around $\theta^\star$ to the one around a shifted point $\theta^\star - \Delta_c$.

In Section 5.3, we localize the empirical parameter $\hat{\theta}_n$ and make use of the second change of measure, this time to a mixture of measures, following [16]. Even though we follow the same high level idea, we modified the original proof in order to better handle the cone covering, and also to make all quantities explicit.

In Section 5.4, we apply a concentration of measure argument. This part requires a specific care since this is the core of the finite-time result. An important complication comes from the "boundary" of the parameter set, which was not explicitly controlled in the original proof of [16]. A very careful analysis enables us to obtain the finite-time concentration result without further restriction.

We finally combine all these steps in Section 5.5.

Notation

$K \in \mathbb{N}$	Dimension of the exponential family
$\Theta \subset \mathbb{R}^K$	Parameter set, see Theorem 3
$\Theta_\rho \subset \mathbb{R}^K$	Enlarged parameter set, see Definition 4
ψ	Log-partition function of the exponential family
$\mathcal{B}^\psi$	Bregman divergence of the exponential family
V_ρ, v_ρ	Largest and smallest eigenvalues of the Hessian of Θ_ρ, see Definition 4
$\theta^\star$	Parameter of the distribution generating the observed samples
$\hat{\theta}_n$	Empirical parameter built from n observations
$\hat{F}_n \in \mathbb{R}^K$	Empirical mean of the $F(X_i)$, $i \le n$, see Section 3.3
f	Threshold function parameterizing the boundary crossing
$\mu^\star \in \mathbb{R}$	Mean of the distribution with parameter $\theta^\star$
$\epsilon > 0$	Shift from the mean
$n \in \mathbb{N}$	Index referring to a number of samples
$p \in [0, 1]$	Angle aperture of the cone
$\eta \in [0, 1]$	Repulsive parameter for cone covering.

5.1. Peeling and Covering

In this section, the intuition we follow is that we want to control the random number of pulls $N_{a^\star}(t) \in [1, t]$ and to this end use a standard peeling argument, considering maximum concentration inequalities on time intervals $[b^i, b^{i+1}]$ for some $b > 1$. Likewise, since the term $\mathcal{K}_{a^\star}(\Pi_{a^\star}(\hat{\nu}_{a^\star,n}), \mu^\star - \epsilon)$ can be viewed as an infimum of some quantity over the parameter set Θ, we use a covering of Θ in order to reduce the control of the desired quantity to that of each cell of the cover. Formally, we show that

Lemma 10 (Peeling and cone covering decomposition)**.** *For all $\beta \in (0, 1)$, $b > 1$ and $\eta \in [0, 1)$, it holds*

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{1\le n\le t} \hat{\theta}_n \in \Theta_\rho \cap \mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\hat{\nu}_{a^\star,n}), \mu^\star - \epsilon\big) \geq f(t/n)/n \Big\}$$
$$\leq \sum_{i=0}^{\lceil \log_b(\beta t+\beta)\rceil -2} \sum_{c=1}^{C_{p,\eta,K}} \mathbb{P}_{\theta^\star}\Big\{ \bigcup_{b^i\le n< b^{i+1}} E_{c,p}(n,t)\Big\} + \sum_{c=1}^{C_{p,\eta,K}} \mathbb{P}_{\theta^\star}\Big\{ \bigcup_{n=b^{\lceil \log_b(\beta t+\beta)\rceil -1}}^{t} E_{c,p}(n,t)\Big\},$$

where the event $E_{c,p}(n, t)$ is defined by

$$E_{c,p}(n,t) \stackrel{\text{def}}{=} \Big\{ \hat{\theta}_n \in \Theta_\rho \cap \hat{F}_n \in \mathcal{C}_p(\theta^\star_c) \cap \mathcal{B}^\psi(\hat{\theta}_n, \theta^\star_c) \geq \frac{f(t/n)}{n}\Big\}. \tag{6}$$

In this definition, $(\theta^\star_c)_{c\le C_{p,\eta,K}}$, constrained to satisfy $\theta^\star_c \notin \mathcal{B}_2(\theta^\star, \eta\rho_\epsilon)$, parameterize a minimal covering of $\nabla\psi(\Theta_\rho \setminus \mathcal{B}_2(\theta^\star, \rho_\epsilon))$ with cones $\mathcal{C}_p(\theta^\star_c) := \mathcal{C}_p(\nabla\psi(\theta^\star_c); \theta^\star - \theta^\star_c)$. That is

$$\nabla\psi(\Theta_\rho \setminus \mathcal{B}_2(\theta^\star, \rho_\epsilon)) \subset \bigcup_{c=1}^{C_{p,\eta,K}} \mathcal{C}_p(\theta^\star_c), \quad \textit{where} \quad \mathcal{C}_p(y; \Delta) = \big\{ y' \in \mathbb{R}^K : \langle y' - y, \Delta\rangle \geq p\|y' - y\|\,\|\Delta\|\big\}.$$

For all $\eta < 1$, $C_{p,\eta,K}$ is of order $(1 - p)^{-K}$ and $C_{p,\eta,1} = 2$, while $C_{p,\eta,K} \to \infty$ when $\eta \to 1$.

Peeling. Let us introduce an increasing sequence $\{n_i\}_{i\in\mathbb{N}}$ such that $n_0 = 1 < n_1 < \ldots < n_{I_t} = t + 1$ for some $I_t \in \mathbb{N}_\star$. Then by a simple union bound it holds for any event E_n

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{1\le n\le t} E_n\Big\} \leq \sum_{i=0}^{I_t - 1} \mathbb{P}_{\theta^\star}\Big\{ \bigcup_{n_i\le n< n_{i+1}} E_n\Big\}.$$

We apply this simple result to the following sequence defined for some $b > 1$ and $\beta \in (0, 1)$ by

$$n_i = \begin{cases} b^i & \text{if } i < I_t \stackrel{\text{def}}{=} \lceil \log_b(\beta t + \beta)\rceil, \\ t+1 & \text{if } i = I_t, \end{cases}$$

(this is indeed a valid sequence since $n_{I_t - 1} \leq b^{\log_b(\beta t+\beta)} = \beta(t + 1) < t + 1 = n_{I_t}$), and to the event

$$E_n \stackrel{\text{def}}{=} \Big\{ \hat{\theta}_n \in \Theta_\rho \cap \mathcal{K}_{a^\star}(\Pi_{a^\star}(\hat{\nu}_{a^\star,n}), \mu^\star - \epsilon) \geq f(t/n)/n\Big\}.$$

Covering. We now make the Kullback–Leibler projection explicit, and remark that in case of a regular family, it holds that

$$\mathcal{K}_{a^\star}(\Pi_{a^\star}(\hat{\nu}_{a^\star,n}), \mu^\star - \epsilon) = \inf\big\{ \mathcal{B}^\psi(\hat{\theta}_n, \theta^\star - \Delta) : \theta^\star - \Delta \in \Theta_{\mathcal{D}}, \mu_{\theta^\star-\Delta} \geq \mu^\star - \epsilon\big\},$$

where $\hat{\theta}_n \in \Theta_{\mathcal{D}}$ is any point such that $\hat{F}_n = \nabla\psi(\hat{\theta}_n)$. This rewriting makes appear explicitly a shift from $\theta^\star$ to another point $\theta^\star - \Delta$. For this reason, it is natural to study the link between $\mathcal{B}^\psi(\hat{\theta}_n, \theta^\star)$ and $\mathcal{B}^\psi(\hat{\theta}_n, \theta^\star - \Delta)$. Immediate computations show that for any Δ such that $\theta^\star - \Delta \in \Theta_{\mathcal{D}}$ it holds

$$\begin{aligned}\mathcal{B}^\psi(\hat{\theta}_n, \theta^\star - \Delta) &= \psi(\theta^\star - \Delta) - \psi(\hat{\theta}_n) - \langle \theta^\star - \Delta - \hat{\theta}_n, \nabla\psi(\hat{\theta}_n)\rangle \\ &= \psi(\theta^\star) - \psi(\hat{\theta}_n) - \langle \theta^\star - \hat{\theta}_n, \nabla\psi(\hat{\theta}_n)\rangle + \psi(\theta^\star - \Delta) - \psi(\theta^\star) + \langle \Delta, \nabla\psi(\hat{\theta}_n)\rangle\end{aligned}$$

$$= \mathcal{B}^\psi(\hat\theta_n, \theta^\star) + \psi(\theta^\star - \Delta) - \psi(\theta^\star) + \langle \Delta, \nabla\psi(\hat\theta_n)\rangle$$
$$= \mathcal{B}^\psi(\hat\theta_n, \theta^\star) \underbrace{- \mathcal{B}^\psi(\theta^\star - \Delta, \theta^\star) - \langle \Delta, \nabla\psi(\theta^\star - \Delta) - \hat F_n\rangle}_{\text{shift}}. \quad (7)$$

With this equality, the Kullback–Leibler projection can be rewritten to make appear an infimum over the shift term only. In order to control the second part of the shift term we localize it thanks to a cone covering of $\nabla\psi(\Theta_{\mathcal{D}})$. More precisely, on the event E_n, we know that $\hat\theta_n \notin \mathcal{B}_2(\theta^\star, \rho_\epsilon)$. Indeed, for all $\theta \in \mathcal{B}_2(\theta^\star, \rho_\epsilon) \cap \Theta_{\mathcal{D}}$ we have $\mu_\theta \geq \mu_\star - \epsilon$, and thus $\mathcal{K}_{a^\star}(\nu_\theta, \mu^\star - \epsilon) = 0$. It is thus natural to build a covering of $\nabla\psi(\Theta_\rho \setminus \mathcal{B}_2(\theta^\star, \rho_\epsilon))$. Formally, for a given $p \in [0, 1]$ and a base point $y \in \mathcal{Y}$, let us introduce the cone

$$\mathcal{C}_p(y; \Delta) = \big\{ y' \in \mathbb{R}^K : \langle \Delta, y' - y\rangle \geq p\|\Delta\|\,\|y' - y\|\big\}.$$

We then associate with each $\theta \in \Theta_\rho$ a cone defined by $\mathcal{C}_p(\theta) = \mathcal{C}_p(\nabla\psi(\theta), \theta^\star - \theta)$. Now for a given p, let $(\theta^\star_c)_{c=1,\ldots,C_{p,\eta,K}}$ be the set of points corresponding to a minimal covering of the set $\nabla\psi(\Theta_\rho \setminus \mathcal{B}_2(\theta^\star, \rho_\epsilon))$, in the sense that

$$\nabla\psi(\Theta_\rho \setminus \mathcal{B}_2(\theta^\star, \rho_\epsilon)) \subset \bigcup_{c=1}^{C_{p,\eta,K}} \mathcal{C}_p(\theta^\star_c) \quad \text{with minimal } C_{p,\eta,K} \in \mathbb{N},$$

constrained to be outside the ball $\mathcal{B}_2(\theta^\star, \eta\rho_\epsilon)$, that is $\theta^\star_c \notin \mathcal{B}_2(\theta^\star, \eta\rho_\epsilon)$ for each c. It can be readily checked that by minimality of the size of the covering $C_{p,\eta,K}$, it must be that $\theta^\star_c \in \Theta_\rho \cap \mathcal{B}_2(\theta^\star, \rho_\epsilon)$. More precisely, when $p < 1$, then $\Delta_c = \theta^\star - \theta^\star_c$ is such that $\rho_\epsilon - \|\Delta_c\|$ is positive and away from 0. Also, by the property of $\mathcal{B}_2(\theta^\star, \rho_\epsilon)$ we have that $\mu_{\theta^\star_c} \geq \mu^\star - \epsilon$, and by the constraint that $\|\Delta_c\| > \eta\rho_\epsilon$.

The size of the covering $C_{p,\eta,K}$ depends on the angle separation p, the ambient dimension K, and the repulsive parameter η. For instance it can be checked that $C_{p,\eta,1} = 2$ for all $p \in (0, 1]$ and $\eta < 1$. In higher dimension, $C_{p,\eta,K}$ typically scales as $(1 - p)^{-K}$ and blows up when $p \to 1$. It also blows up when $\eta \to 1$. It is now natural to introduce the decomposition

$$E_{c,p}(n, t) \stackrel{\text{def}}{=} \left\{ \hat\theta_n \in \Theta_\rho \cap \hat F_n \in \mathcal{C}_p(\theta^\star_c) \cap \mathcal{B}^\psi(\hat\theta_n, \theta^\star_c) \geq \frac{f(t/n)}{n} \right\}. \quad (8)$$

Using this notation, we deduce that for all $\beta \in (0, 1)$, $b > 1$ (remind that $I_t = \lceil \log_b(\beta t + \beta)\rceil$),

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{1\leq n\leq t} \hat\theta_n \in \Theta_\rho \cap \mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\hat\nu_{a^\star,n}), \mu^\star - \epsilon\big) \geq f(t/n)/n \Big\} \leq \sum_{i=0}^{I_t-1} \sum_{c=1}^{C_{p,\eta,K}} \mathbb{P}_{\theta^\star}\Big\{ \bigcup_{n_i\leq n<n_{i+1}} E_{c,p}(n, t)\Big\}.$$

5.2. Change of Measure

In this section, we focus on one event $E_{c,p}(n, t)$. The idea is to take advantage of the gap between $\mu^\star$ and $\mu^\star - \epsilon$, which allows us to shift from $\theta^\star$ to some of the $\theta^\star_c$ from the cover. The key observation is to control the change of measure from $\theta^\star$ to each $\theta^\star_c$. Note that $\theta^\star_c \in (\Theta_\rho \cap \mathcal{B}_2(\theta^\star_c, \rho_\epsilon)) \setminus \mathcal{B}_2(\theta^\star_c, \eta\rho_\epsilon)$ and that $\mu_{\theta^\star_c} \geq \mu^\star - \epsilon$. We show that

Lemma 11 (Change of measure). *If $n \to nf(t/n)$ is nondecreasing, then for any increasing sequence $\{n_i\}_{i\geq 0}$ of nonnegative integers it holds*

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{n=n_i}^{n_{i+1}-1} E_{c,p}(n, t)\Big\} \leq \exp\big(- n_i\alpha^2 - \chi\sqrt{n_i f(t/n_i)}\ \big)\mathbb{P}_{\theta^\star_c}\Big\{ \bigcup_{n=n_i}^{n_{i+1}-1} E_{c,p}(n, t)\Big\},$$

where $\alpha = \alpha(p, \eta, \epsilon) = \eta\rho_\epsilon\sqrt{v_\rho/2}$ and $\chi = p\eta\rho_\epsilon\sqrt{2v_\rho^2/V_\rho}$.

Proof. For any measurable event E, we have by absolute continuity that

$$\mathbb{P}_{\theta^\star}\Big\{E\Big\} = \int_E \frac{d\mathbb{P}_{\theta^\star}}{d\mathbb{P}_{\theta_c^\star}}\, d\mathbb{P}_{\theta_c^\star}.$$

We thus bound the ratio which, in the case of $E = \{\bigcup_{n_i \le n < n_{i+1}} E_{c,p}(n,t)\}$, leads to

$$\begin{aligned}\int_E \frac{d\mathbb{P}_{\theta^\star}}{d\mathbb{P}_{\theta_c^\star}}\, d\mathbb{P}_{\theta_c^\star} &= \int_E \frac{\Pi_{k=1}^n \nu_{\theta^\star}(X_k)}{\Pi_{k=1}^n \nu_{\theta_c^\star}(X_k)}\, d\mathbb{P}_{\theta_c^\star}\\ &= \int_E \exp\big(n\langle \theta^\star - \theta_c^\star, \hat{F}_{a^\star,n}\rangle - n(\psi(\theta^\star) - \psi(\theta_c^\star))\big)\, d\mathbb{P}_{\theta_c^\star}\\ &= \int_E \exp\big(-n\langle \Delta_c, \nabla\psi(\theta_c^\star) - \hat{F}_{a^\star,n}\rangle - n\mathcal{B}^\psi(\theta_c^\star, \theta^\star)\big)\, d\mathbb{P}_{\theta_c^\star},\end{aligned} \tag{9}$$

where $\Delta_c = \theta^\star - \theta_c^\star$. Note that this rewriting gives rise to the same term as the shift term appearing in (7). Now, we remark that since $\theta_c^\star \in \Theta_\rho$ by construction, under the event $E_{c,p}(n,t)$ it holds by convexity of Θ_ρ and elementary Taylor approximation

$$\begin{aligned}-\langle \Delta_c, \nabla\psi(\theta_c^\star) - \hat{F}_n\rangle &\le -p\|\Delta_c\|\,\|\nabla\psi(\theta_c^\star) - \hat{F}_n\| \le -p\|\Delta_c\| v_\rho \|\hat{\theta}_n^\star - \theta_c^\star\|\\ &\le -p\|\Delta_c\|_c v_\rho \sqrt{\frac{2}{V_\rho}\mathcal{B}^\psi(\hat{\theta}_n, \theta_c^\star)} \le -p\eta\rho_\epsilon v_\rho \sqrt{\frac{2f(t/n)}{V_\rho n}},\end{aligned} \tag{10}$$

where we used the fact that $\|\Delta_c\| \ge \eta\rho_\epsilon$. On the other hand, it also holds that

$$-\mathcal{B}^\psi(\theta_c^\star, \theta^\star) \le -\frac{1}{2}v_\rho\|\Delta_c\|^2 \le -\frac{1}{2}v_\rho\eta^2\rho_\epsilon^2. \tag{11}$$

To conclude the proof we plug-in (10) and (11) into (9). Then it remains to use that $n \ge b^i$ and the fact that $n \mapsto nf(t/n)$ is nondecreasing. □

5.3. Localized Change of Measure

In this section, we decompose further the event of interest in $\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n_i \le n < n_{i+1}} E_{c,p}(n,t)\Big\}$ in order to apply some concentration of measure argument. In particular, since by construction

$$\hat{F}_n \in \mathcal{C}_p(\theta_c^\star) \Leftrightarrow \langle \Delta_c, \nabla\psi(\theta_c^\star) - \hat{F}_n\rangle \ge p\|\Delta_c\|\,\|\nabla\psi(\theta_c^\star) - \hat{F}_n\|,$$

it is natural to control $\|\nabla\psi(\theta_c^\star) - \hat{F}_n\|$. This is what we call localization. More precisely, for any sequence $\{\epsilon_{t,i,c}\}_{t,i}$ of positive values we introduce the following decomposition

$$\begin{aligned}\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n_i \le n < n_{i+1}} E_{c,p}(n,t)\Big\} &\le \mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n_i \le n < n_{i+1}} E_{c,p}(n,t) \cap \|\nabla\psi(\theta_c^\star) - \hat{F}_n\| < \epsilon_{t,i,c}\Big\}\\ &\quad + \mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n_i \le n < n_{i+1}} E_{c,p}(n,t) \cap \|\nabla\psi(\theta_c^\star) - \hat{F}_n\| \ge \epsilon_{t,i,c}\Big\}.\end{aligned} \tag{12}$$

We handle the first term in (12) by another change of measure argument that we detail below, and the second term thanks to a concentration of measure argument that we detail in Section 5.4.

Lemma 12 (Change of measure)**.** *For any sequence of positive values $\{\epsilon_{t,i,c}\}_{i \ge 0}$, it holds*

$$\begin{aligned}&\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n_i \le n < n_{i+1}} E_{c,p}(n,t) \cap \|\nabla\psi(\hat{\theta}_n) - \nabla\psi(\theta_c^\star)\| < \epsilon_{t,i,c}\Big\}\\ &\quad \le \alpha_{\rho,p} \exp\Big(-f\Big(\frac{t}{n_{i+1}-1}\Big)\Big) \min\Big\{\rho^2 v_\rho^2, \tilde{\epsilon}_{t,i,c}^2, \frac{(K+2)v_\rho^2}{K(n_{i+1}-1)V_\rho}\Big\}^{-K/2} \tilde{\epsilon}_{t,i,c}^K,\end{aligned}$$

where $\tilde{\epsilon}_{t,i,c} = \min\{\epsilon_{t,i,c}, \mathrm{Diam}\,(\nabla\psi(\Theta_\rho) \cap \mathcal{C}_p(\theta_c^\star))\}$ *and* $\alpha_{\rho,p} = 2\frac{\omega_{p,K-2}}{\omega_{p',K-2}}\big(\frac{V_\rho}{v_\rho^2}\big)^{K/2}\big(\frac{V_\rho}{v_\rho}\big)^K$ *with* $p' > \max\{p, \frac{2}{\sqrt{5}}\}$ *and* $\omega_{p,K} = \int_p^1 \sqrt{1-z^2}^K\, dz$ *for* $K \ge 0$ *and* $w_{p,-1} = 1$.

Let us recall that

$$E_{c,p}(n,t) = \big\{\hat\theta_n \in \Theta_\rho \cap \hat F_n \in \mathcal{C}_p(\theta_c^\star) \cap n\mathcal{B}^\psi(\hat\theta_n, \theta_c^\star) \geq f(t/n)\big\}.$$

The idea is to go from $\theta_c^\star$ to the measure that corresponds to the mixture of all the θ' in the shrink ball $B = \Theta_\rho \cap \nabla\psi^{-1}\big(\mathcal{C}_p(\theta_c^\star) \cap \mathcal{B}_2(\nabla\psi(\theta_c^\star), \epsilon_{t,i,c})\big)$, where $\mathcal{B}_2(y,r) \overset{\text{def}}{=} \big\{y' \in \mathbb{R}^K;\ \|y - y'\| \leq t\big\}$. This makes sense since, on the one hand, $\nabla\psi(\hat\theta_n) \in \mathcal{C}_p(\theta_c^\star)$ under $E_{c,p}(n,t)$, and on the other hand, $\|\nabla\psi(\hat\theta_n) - \nabla\psi(\theta_c^\star)\| \leq \epsilon_{t,i,c}$. For convenience, let us introduce the event of interest

$$\Omega = \Big\{\bigcup_{n_i \leq n < n_{i+1}} E_{c,p}(n,t) \cap \|\nabla\psi(\hat\theta_n) - \nabla\psi(\theta_c^\star)\| \leq \epsilon_{t,i,c}\Big\}.$$

We use the following change of measure

$$d\mathbb{P}_{\theta_c^\star} = \frac{d\mathbb{P}_{\theta_c^\star}}{dQ_B} dQ_B,$$

where $Q_B(\Omega) \overset{\text{def}}{=} \int_{\theta' \in B} \mathbb{P}_{\theta'}\{\Omega\}\, d\theta'$ is the mixture of all distributions with parameter in B. The proof technique consists now in bounding the ratio by some quantity not depending on Ω. We have

$$\begin{aligned}\int_\Omega \frac{d\mathbb{P}_\theta}{dQ_B} dQ_B &= \int_\Omega \bigg[\int_{\theta' \in B} \frac{\Pi_{k=1}^n \nu_{\theta'}(X_k)}{\Pi_{k=1}^n \nu_\theta(X_k)}\, d\theta'\bigg]^{-1} dQ_B \\ &= \int_\Omega \bigg[\int_{\theta' \in B} \exp\big(n\langle\theta' - \theta, \hat F_{a^\star,n}\rangle - n(\psi(\theta') - \psi(\theta))\big)\, d\theta'\bigg]^{-1} dQ_B.\end{aligned}$$

We remark that the term in the exponent can be rewritten in terms of Bregman divergence: by elementary substitution of the definition of the divergence and of $\nabla\psi(\hat\theta_n) = \hat F_{a^\star,n}$, it holds

$$\langle\theta' - \theta, \hat F_{a^\star,n}\rangle - \big(\psi(\theta') - \psi(\theta)\big) = \mathcal{B}^\psi(\hat\theta_n, \theta) - \mathcal{B}^\psi(\hat\theta_n, \theta').$$

Thus the above likelihood ratio simplifies as follows

$$\begin{aligned}\frac{d\mathbb{P}_\theta}{dQ_B} &= \bigg[\int_{\theta' \in B} \exp\big(n\mathcal{B}^\psi(\hat\theta_n, \theta) - n\mathcal{B}^\psi(\hat\theta_n, \theta')\big)\, d\theta'\bigg]^{-1} \\ &\leq \bigg[\int_{\theta' \in B} \exp\big(f(t/n) - n\mathcal{B}^\psi(\hat\theta_n, \theta')\big)\, d\theta'\bigg]^{-1} \\ &= \exp\big(-f(t/n)\big)\bigg[\int_{\theta' \in B} \exp\big(-n\mathcal{B}^\psi(\hat\theta_n, \theta')\big)\, dx\bigg]^{-1},\end{aligned}$$

where both θ' and $\hat\theta_n$ belong to Θ_ρ.

The next step is to consider a set $B' \subset B$ that contains $\hat\theta_n$. For each such set, and the upper bound $\mathcal{B}^\psi(\hat\theta_n, \theta') \leq \frac{V_\rho}{2v_\rho^2}\|\nabla\psi(\hat\theta_n) - \nabla\psi(\theta')\|^2$, we now obtain

$$\begin{aligned}\frac{d\mathbb{P}_\theta}{dQ_B} &\overset{(a)}{\leq} \exp\big(-f(t/n)\big)\bigg[\int_{\theta' \in B'} \exp\Big(-\frac{nV_\rho}{2v_\rho^2}\|\nabla\psi(\hat\theta_n) - \nabla\psi(\theta')\|^2\Big)\, d\theta'\bigg]^{-1} \\ &\overset{(b)}{=} \exp\big(-f(t/n)\big)\bigg[\int_{y \in \nabla\psi(B')} \exp\Big(-\frac{nV_\rho}{2v_\rho^2}\|\nabla\psi(\hat\theta_n) - y\|^2\Big) |\det(\nabla^2\psi^{-1}(y))|\, dy\bigg]^{-1} \\ &\overset{(c)}{\leq} \exp\big(-f(t/n)\big)\bigg[\int_{y \in \nabla\psi(B')} \exp\Big(-\frac{nV_\rho}{2v_\rho^2}\|\nabla\psi(\hat\theta_n) - y\|^2\Big)\, dy\bigg]^{-1} V_\rho^K.\end{aligned}$$

In this derivation, (a) holds by positivity of exp and the inclusion $B' \subset B$, (b) follows by a change of parameter argument, and (c) is obtained by controlling the determinant (in dimension K) of the Hessian, whose highest eigenvalue is V_ρ.

In order to identify a good candidate for the set B' let us now study the set B. The first remark is that $\theta_c^\star$ plays a central role in B: it is not difficult to show that, by construction of B,

$$\nabla\psi^{-1}\big(\nabla\psi(\theta_c^\star) + \mathcal{B}_2(0, \min\{v_\rho\rho, \epsilon_{t,i,c}\}) \cap \mathcal{C}_p(0; \Delta_c)\big) \subset B.$$

Indeed, if θ' belongs to the set on the left-hand side, then it must satisfy, on the one hand, $\nabla\psi(\theta') \in \nabla\psi(\theta_c^\star) + \mathcal{B}_2(0, v_\rho\rho)$. This implies that $\theta' \in \mathcal{B}_2(\theta_c^\star, \rho) \subset \Theta_\rho$ (this last inclusion is by construction of Θ). On the other hand, it satisfies $\nabla\psi(\theta') \in \nabla\psi(\theta_c^\star) + \mathcal{B}_2(0, \epsilon_{t,i,c}) \cap \mathcal{C}_p(0, \Delta_c)$. These two properties show that such a θ' belongs to B.

Thus, a natural candidate B' should satisfy $\nabla\psi(B') \subset \nabla\psi(\theta_c^\star) + \mathcal{B}_2(0, \tilde{r}) \cap \mathcal{C}_p(0; \Delta_c)$, with $\tilde{r} = \min\{v_\rho\rho, \epsilon_{t,i,c}\}$. It is then natural to look for B' in the form $\nabla\psi^{-1}(\nabla\psi(\theta_c^\star) + \mathcal{B}_2(0, \tilde{r}) \cap \mathcal{D})$, where $\mathcal{D} \subset \mathcal{C}_p(0; \Delta_c)$ is a sub-cone of $\mathcal{C}_p(0; \Delta_c)$ with base point 0. In this case, the previous derivation simplifies into

$$\frac{d\mathbb{P}_\theta}{dQ_B} \leq \exp\big(-f(t/n)\big)\left[\int_{y\in\mathcal{B}_2(0,\tilde{r})\cap\mathcal{D}} \exp\big(-C\|y_n - y\|^2\big)\,dy\right]^{-1} V_\rho^K,$$

where $y_n = \nabla\psi(\hat{\theta}_n) - \nabla\psi(\theta_c^\star) \in \mathcal{B}_2(0, \tilde{r}) \cap \mathcal{D}$ and $C = \frac{nV_\rho}{2v_\rho^2}$. The cases of special interest for the set $\mathcal{D}$ are such that the value of the function $g\colon y \mapsto \int_{y'\in\mathcal{B}_2(0,\tilde{r})\cap\mathcal{D}} \exp\big(-C\|y - y'\|^2\big)\,dy'$, for $y \in \mathcal{B}_2(0, \tilde{r}) \cap \mathcal{D}$ is minimal at the base point 0. Indeed this enables one to derive the following bound

$$\begin{aligned}\frac{d\mathbb{P}_\theta}{dQ_B} &\leq \exp\big(-f(t/n)\big)\left[\int_{y\in\mathcal{B}_2(0,\min\{v_\rho\rho,\epsilon_{t,i,c}\})\cap\mathcal{D}} \exp\left(-\frac{nV_\rho}{2v_\rho^2}\|y\|^2\right)dy\right]^{-1} V_\rho^K\\ &\overset{(d)}{=} \exp\big(-f(t/n)\big)\left[\int_{y\in\mathcal{B}_2(0,r_\rho)\cap\mathcal{D}} \exp\big(-n\|y\|^2\big)\,dy\right]^{-1}\left(\frac{V_\rho^2}{2v_\rho^2}\right)^K,\end{aligned}$$

where (d) follows from another change of parameter argument, with $r_\rho = \sqrt{\frac{V_\rho}{2v_\rho^2}}\min\{v_\rho\rho, \epsilon_{t,i,c}\}$ combined with isotropy of the Euclidean norm (the right-hand side of (d) no longer depends on the random direction Δ_n), plus the fact that the sub-cone $\mathcal{D}$ is invariant by rescaling. We recognize here a Gaussian integral on $\mathcal{B}_2(0, r_\rho) \cap \mathcal{D}$ that can be bounded explicitly (see below).

Following this reasoning, we are now ready to specify the set $\mathcal{D}$. Let $\mathcal{D} = \mathcal{C}_{p'}(0; \Delta_n) \subset \mathcal{C}_p(0; \Delta_c)$ be a sub-cone where $p' \geq p$ (remember that the larger p, the more acute is a cone) and Δ_n is chosen such that $\nabla\psi(\hat{\theta}_n) \in \nabla\psi(\theta_c^\star) + \mathcal{D}$ (there always exists such a cone). It thus remains to specify p'. A study of the function g (defined above) on the domain $\mathcal{B}_2(0, \tilde{r}) \cap \mathcal{C}_{p'}(0; \Delta_n)$ reveals that it is minimal at the point 0 provided that p' is not too small, more precisely provided that $p \geq 2/\sqrt{5}$. The intuitive reasons are that the points that contribute most to the integral belong to the set $\mathcal{B}_2(y, r) \cap \mathcal{B}_2(0, \tilde{r}) \cap \mathcal{D}$ for small values of r, that this set has the lowest volume (the map $y \to |\mathcal{B}_2(y, r) \cap \mathcal{B}_2(0, \tilde{r}) \cap \mathcal{D}|$ is minimal) when $y \in \partial\mathcal{B}_2(0, \tilde{r}) \cap \partial\mathcal{D}$ and that $y = 0$ is a minimizer among these points provided that p' is not too small. More formally, the function g rewrites

$$g(y) = \int_{r=0}^{\infty} e^{-Cr^2}|\mathcal{S}_2(y, r) \cap \mathcal{B}_2(0, \tilde{r}) \cap \mathcal{D}|dr,$$

from which we see that a minimal y should be such that the spherical section $|\mathcal{S}_2(y, r) \cap \mathcal{B}_2(0, \tilde{r}) \cap \mathcal{D}|$ is minimal for small values of r (note also that $C = O(n)$). Then, since $B = \mathcal{B}_2(0, \tilde{r}) \cap \mathcal{D}$ is a convex set, the sections $|\mathcal{S}_2(y, r) \cap \mathcal{B}_2(0, \tilde{r}) \cap \mathcal{D}|$ are of minimal size for points $y \in B$ that are extremal, in the sense that y satisfies $B \subset \mathcal{B}_2(y, \mathrm{Diam}(B))$. In order to choose p' and fully specify $\mathcal{D}$, we finally use the following lemma.

Lemma 13. *Let $\mathcal{C}_{p'} = \{y'\colon \langle y', \Delta\rangle \geq p'\|y'\|\,\|\Delta\|\}$ be a cone with base point 0 and define $B = \mathcal{B}_2(0, r) \cap \mathcal{C}_{p'}$. If $p' > 2/\sqrt{5}$, the set of extremal points $\{y \in B\colon B \subset \mathcal{B}_2(y, \mathrm{Diam}(B))\}$ reduces to $\{0\}$.*

Proof. First, note that the boundary of the convex set B is supported by the union of the base point 0 and the set $\partial\mathcal{B}_2(0,\tilde{r})\cap\partial\mathcal{D}$. Since this set is a sphere in dimension $K-1$ with radius $\frac{\sqrt{1-p'^2}}{p}\tilde{r}$, all its points are at a distance at most $2\frac{\sqrt{1-p'^2}}{p'}\tilde{r}$ from each other. Now they are also at the distance exactly $\tilde{r}$ from the base point 0. Thus, when $2\frac{\sqrt{1-p'^2}}{p'}\tilde{r}<\tilde{r}$, that is $p'>2/\sqrt{5}$, then 0 is the unique point that satisfies $B\subset\mathcal{B}_2(y,\mathrm{Diam}(B))$. □

We now summarize the previous steps. So far, we have proved the following upper bound

$$\begin{aligned}\mathbb{P}_{\theta_c^\star}\{\Omega\} &\le \max_{n_i\le n<n_{i+1}} \exp\big(-f(t/n)\big)\Big[\int_{y\in\mathcal{B}_2(0,r_\rho)\cap\mathcal{C}_{p'}(0;\mathbf{1})} \exp\big(-n\|y\|^2\big)\,dy\Big]^{-1}\Big(\frac{V_\rho^2}{2v_\rho^2}\Big)^K\int_{\theta'\in B}\mathbb{P}_{\theta'}\{\Omega\}\,d\theta'\\ &\le \exp\big(-f(t/(n_{i+1}-1))\big)\Big[\int_{y\in\mathcal{B}_2(0,r_\rho)\cap\mathcal{C}_{p'}(0;\mathbf{1})}\exp\big(-(n_{i+1}-1)\|y\|^2\big)\,dy\Big]^{-1}\Big(\frac{V_\rho}{2v_\rho^2}\Big)^K V_\rho^K|B|,\end{aligned}$$

where $|B|$ denotes the volume of B, $r_\rho=\sqrt{\frac{V_\rho}{2v_\rho^2}}\min\{v_\rho\rho,\epsilon_{t,i,c}\}$ and $p'>\max\{p,2/\sqrt{5}\}$. We remark that by definition of B it holds

$$\begin{aligned}|B| &\le \sup_{\theta\in\Theta_\rho}\det(\nabla^2\psi^{-1}(\theta))|\mathcal{B}_2(0,\epsilon_{t,i,c})\cap\mathcal{C}_p(0;\mathbf{1})|\\ &\le v_\rho^{-K}|\mathcal{B}_2(0,\epsilon_{t,i,c})\cap\mathcal{C}_p(0;\mathbf{1})|.\end{aligned}$$

Thus it remains to analyze the volume and the Gaussian integral of $\mathcal{B}_2(0,\epsilon_{t,i,c})\cap\mathcal{C}_p(0;\mathbf{1})$. To do so, we use the following result from elementary geometry, whose proof is given in Appendix A.

Lemma 14. *For all $\epsilon,\epsilon'>0$, $p,p'\in[0,1]$ and all $K\ge1$, the following equality and inequality hold*

$$\frac{|\mathcal{B}_2(0,\epsilon)\cap\mathcal{C}_p(0;\mathbf{1})|}{\int_{\mathcal{B}_2(0,\epsilon')\cap\mathcal{C}_{p'}(0;\mathbf{1})}e^{-\|y\|^2/2}\,dy}=\frac{\omega_{p,K-2}}{\omega_{p',K-2}}\frac{\int_0^\epsilon r^{K-1}\,dr}{\int_0^{\epsilon'}e^{-r^2/2}r^{K-1}\,dr}\le 2\frac{\omega_{p,K-2}}{\omega_{p',K-2}}\Big(\frac{\epsilon}{\min\{\epsilon',\sqrt{1+2/K}\}}\Big)^K,$$

where $\omega_{p,K-2}=\int_p^1\sqrt{1-z^2}^{K-2}dz$ for $K\ge2$ and using the convention that $\omega_{p,-1}=1$.

Applying this Lemma, we get for $r_\rho=\sqrt{\frac{V_\rho}{2v_\rho^2}}\min\{v_\rho\rho,\epsilon_{t,i,c}\}$,

$$\begin{aligned}\mathbb{P}_{\theta_c^\star}\{\Omega\} &\le e^{-f(\frac{t}{n_{i+1}-1})}\frac{\big(\frac{V_\rho}{2v_\rho^2}\big)^K\big(\frac{V_\rho}{v_\rho}\big)^K|\mathcal{B}_2(0,\epsilon_{t,i,c})\cap\mathcal{C}_p(0;\mathbf{1})|}{\int_{y\in\mathcal{B}_2(0,r_\rho)\cap\mathcal{C}_{p'}(0;\mathbf{1})}\exp\big(-(n_{i+1}-1)\|y\|^2\big)\,dy}\\ &= e^{-f(\frac{t}{n_{i+1}-1})}\Big(\frac{V_\rho}{v_\rho^2}\Big)^K\Big(\frac{V_\rho}{v_\rho}\Big)^K(n_{i+1}-1)^{K/2}\frac{|\mathcal{B}_2(0,\epsilon_{t,i,c})\cap\mathcal{C}_p(0;\mathbf{1})|}{\int_{y\in\mathcal{B}_2(0,\sqrt{2(n_{i+1}-1)}r_\rho)\cap\mathcal{C}_{p'}(0;\mathbf{1})}\exp\big(-\|y\|^2/2\big)\,dy}\\ &\le 2\frac{\omega_{p,K-2}}{\omega_{p',K-2}}e^{-f(\frac{t}{n_{i+1}-1})}\Big(\frac{V_\rho}{v_\rho^2}\Big)^K\Big(\frac{V_\rho}{v_\rho}\Big)^K(n_{i+1}-1)^{K/2}\Big(\frac{\epsilon_{t,i,c}}{\min\{\sqrt{2(n_{i+1}-1)}r_\rho,\sqrt{1+2/K}\}}\Big)^K\\ &= 2\frac{\omega_{p,K-2}}{\omega_{p',K-2}}e^{-f(\frac{t}{n_{i+1}-1})}\Big(\frac{V_\rho}{v_\rho^2}\Big)^K\Big(\frac{V_\rho}{v_\rho}\Big)^K\Big(\frac{\epsilon_{t,i,c}^2}{\min\{v_\rho^2\rho^2,\epsilon_{t,i,c}^2,\frac{(K+2)v_\rho^2}{KV_\rho(n_{i+1}-1)}\}}\Big)^{K/2}\Big(\frac{V_\rho}{v_\rho^2}\Big)^{-K/2}\\ &= 2\frac{\omega_{p,K-2}}{\omega_{p',K-2}}\Big(\frac{V_\rho}{v_\rho^2}\Big)^{K/2}\Big(\frac{V_\rho}{v_\rho}\Big)^K e^{-f(\frac{t}{n_{i+1}-1})}\min\Big\{v_\rho^2\rho^2,\epsilon_{t,i,c}^2,\frac{(K+2)v_\rho^2}{KV_\rho(n_{i+1}-1)}\Big\}^{-K/2}\epsilon_{t,i,c}^K.\end{aligned}$$

This concludes the proof of Lemma 12. □

5.4. Concentration of Measure

In this section, we focus on the second term in (12), that is we want to control

$$\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n_i\le n<n_{i+1}} E_{c,p}(n,t)\cap\|\nabla\psi(\theta_c^\star)-\hat{F}_n\|\ge\epsilon_{t,i,c}\Big\}.$$

In this term, $\epsilon_{t,i,c}$ should be considered as decreasing fast to 0 with i and slowly increasing with t. Note that by definition $\nabla\psi(\hat\theta_n)=\hat{F}_{a^\star,n}=\frac{1}{n}\sum_{i=1}^n F(X_{a^\star,i})\in\mathbb{R}^K$ is an empirical mean with mean given by $\nabla\psi(\theta_c^\star)\in\mathbb{R}^K$ and covariance matrix $\frac{1}{n}\nabla^2\psi(\theta_c^\star)$. We thus resort to a concentration of measure argument.

Lemma 15 (Concentration of measure). *Let* $\epsilon_c^{\max}=\mathrm{Diam}(\nabla\psi(\Theta_\rho\cap\mathcal{C}_{c,p}))$*, where we introduced the projected cone* $\mathcal{C}_{c,p}=\{\theta\in\Theta\colon\langle\frac{\Delta_c}{\|\Delta_c\|},\frac{\nabla\psi(\theta_c^\star)-\nabla\psi(\theta)}{\|\nabla\psi(\theta_c^\star)-\nabla\psi(\theta)\|}\rangle\ge p\}$*. Then, for all* $\epsilon_{t,i,c}$*, it holds*

$$\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n=n_i}^{n_{i+1}-1} E_{c,p}(n,t)\cap\|\nabla\psi(\hat\theta_n)-\nabla\psi(\theta_c^\star)\|\ge\epsilon_{t,i,c}\Big\}\le\exp\Big(-\frac{n_i^2p\epsilon_{t,i,c}^2}{2V_\rho(n_{i+1}-1)}\Big)\mathbf{1}\{\epsilon_{t,i,c}\le\overline{\epsilon}_c\}.$$

Proof. Note that by definition if $\epsilon_{t,i,c}>\epsilon_c^{\max}$, then

$$\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n_i\le n<n_{i+1}} E_{c,p}(n,t)\cap\|\nabla\psi(\hat\theta_n)-\nabla\psi(\theta_c^\star)\|\ge\epsilon_{t,i,c}\Big\}=0.$$

We thus restrict to the case when $\epsilon_{t,i,c}\le\epsilon_c^{\max}$, or equivalently, replace $\epsilon_{t,i,c}$ by $\tilde\epsilon_{t,i,c}=\min\{\epsilon_{t,i,c},\epsilon_c^{\max}\}$. Now, by definition of the event $E_{c,p}(n,t)$, we have

$$\begin{aligned}&\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n_i\le n<n_{i+1}} E_{c,p}(n,t)\cap\|\nabla\psi(\hat\theta_n)-\nabla\psi(\theta_c^\star)\|\ge\tilde\epsilon_{t,i,c}\Big\}\\
&\quad\le\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n_i\le n<n_{i+1}}\hat\theta_n\in\Theta_\rho\cap\Big\langle\frac{\Delta_c}{\|\Delta_c\|},\nabla\psi(\theta_c^\star)-\nabla\psi(\hat\theta_n)\Big\rangle\ge p\tilde\epsilon_{t,i,c}\Big\}\\
&\quad\le\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n=n_i}^{n_{i+1}-1}\Big\langle\frac{\Delta_c}{\|\Delta_c\|},\sum_{i=1}^n\big(\nabla\psi(\theta_c^\star)-F(X_{a^\star,i})\big)\Big\rangle\ge pn_i\tilde\epsilon_{t,i,c}\Big\}.\end{aligned}$$

Applying the function $x\mapsto\exp(\lambda x)$ on both sides of the inequality, for a deterministic $\lambda>0$, we obtain

$$\begin{aligned}&\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n_i\le n<n_{i+1}} E_{c,p}(n,t)\cap\|\nabla\psi(\hat\theta_n)-\nabla\psi(\theta_c^\star)\|\ge\tilde\epsilon_{t,i,c}\Big\}\\
&\overset{(a)}{\le}\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n=n_i}^{n_{i+1}-1}\exp\Big(\sum_{i=1}^n\Big\langle\frac{\lambda\Delta_c}{\|\Delta_c\|},\nabla\psi(\theta_c^\star)-F(X_{a^\star,i})\Big\rangle\Big)\ge\exp\big(\lambda pn_i\tilde\epsilon_{t,i,c}\big)\Big\}\\
&=\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n=n_i}^{n_{i+1}-1}\exp\Big(\sum_{i=1}^n\Big\langle\frac{\lambda\Delta_c}{\|\Delta_c\|},\nabla\psi(\theta_c^\star)-F(X_{a^\star,i})\Big\rangle-\frac{\lambda^2(n_{i+1}-1)}{2}V_\rho\Big)\\
&\qquad\ge\exp\Big(\lambda pn_i\tilde\epsilon_{t,i,c}-\frac{\lambda^2(n_{i+1}-1)}{2}V_\rho\Big)\Big\}\\
&\le\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n=n_i}^{n_{i+1}-1}\exp\Big(\sum_{i=1}^n\Big\langle\frac{\lambda\Delta_c}{\|\Delta_c\|},\nabla\psi(\theta_c^\star)-F(X_{a^\star,i})\Big\rangle-\frac{\lambda^2 n}{2}V_\rho\Big)\\
&\qquad\ge\exp\Big(\lambda pn_i\tilde\epsilon_{t,i,c}-\frac{\lambda^2(n_{i+1}-1)}{2}V_\rho\Big)\Big\}.\end{aligned}$$

Now we recognize that the sequence $\{W_n(\lambda)\}_{n\geq 0}$, where $W_n(\lambda) = \exp\big(\sum_{i=1}^n \langle \frac{\lambda\Delta_c}{\|\Delta_c\|}, \nabla\psi(\theta_c^\star) - F(X_{a^\star,i})\rangle - n\frac{\lambda^2 V_\rho}{2}\big)$ is a nonnegative supermartingale provided that λ is not too large. Indeed, provided that $\theta_c^\star - \frac{\lambda\Delta_c}{\|\Delta_c\|} \in \Theta_\rho$ it holds

$$\begin{aligned}
&\mathbb{E}_{\theta_c^\star}\Big[\exp\Big(\sum_{i=1}^n \lambda\Big\langle\frac{\Delta_c}{\|\Delta_c\|}, \nabla\psi(\theta_c^\star) - F(X_{a^\star,i})\Big\rangle - \frac{\lambda^2 n V_\rho}{2})\Big)\Big|\mathcal{H}_{n-1}\Big] \\
&\quad\leq \exp\Big(\sum_{i=1}^{n-1} \lambda\Big\langle\frac{\Delta_c}{\|\Delta_c\|}, \nabla\psi(\theta_c^\star) - F(X_{a^\star,i})\Big\rangle - (n-1)\frac{\lambda^2 V_\rho}{2}\Big) \\
&\qquad\times \underbrace{\mathbb{E}_{\theta_c^\star}\Big[\exp\Big(\lambda\Big\langle\frac{\Delta_c}{\|\Delta_c\|}, \nabla\psi(\theta_c^\star) - F(X_{a^\star,n})\Big\rangle - \frac{\lambda^2 V_\rho}{2}\Big)\Big|\mathcal{H}_{n-1}\Big]}_{\leq 1},
\end{aligned}$$

that is $\mathbb{E}\big[W_n(e,\lambda) \mid H_{n-1}\big] \leq W_{n-1}(e,\lambda)$. Thus we apply Doob's maximal inequality for nonnegative supermartingale and deduce that

$$\begin{aligned}
&\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n_i\leq n<n_{i+1}} E_{c,p}(n,t) \cap \|\nabla\psi(\hat\theta_n) - \nabla\psi(\theta_c^\star)\| \geq \tilde\epsilon_{t,i,c}\Big\} \\
&\quad\leq \mathbb{P}_{\theta_c^\star}\Big\{\max_{n_i\leq n<n_{i+1}} W_n(\lambda) \geq \exp\big(\lambda p n_i \tilde\epsilon_{t,i,c} - \lambda^2(n_{i+1}-1)V_\rho/2\big)\Big\} \\
&\quad\leq \mathbb{E}_{\theta_c^\star}[W_{n_i}(\lambda)]\exp\big(-\lambda p n_i \tilde\epsilon_{t,i,c} + \lambda^2(n_{i+1}-1)V_\rho/2\big) \\
&\quad\leq \exp\big(-\lambda p n_i \tilde\epsilon_{t,i,c} + \lambda^2(n_{i+1}-1)V_\rho/2\big).
\end{aligned}$$

Optimizing over λ gives $\lambda = \lambda^\star = \frac{n_i p \tilde\epsilon_{t,i,c}}{(n_{i+1}-1)V_\rho}$, thus the condition becomes $\theta_c^\star - \frac{n_i p \tilde\epsilon_{t,i,c}}{(n_{i+1}-1)V_\rho\|\Delta_c\|}\Delta_c \in \Theta_\rho$. At this point, it is convenient to introduce the quantity

$$\lambda_c = \operatorname{argmax}\Big\{\lambda\colon \theta_c^\star - \lambda\frac{\Delta_c}{\|\Delta_c\|} \in \Theta_\rho \cap \mathcal{C}_{c,p}\Big\}.$$

Indeed, it suffices to show that $\lambda^\star \leq \lambda_c$ to ensure that the condition is satisfied. It is now not difficult to relate λ_c to $\epsilon_c^{\max}$. Indeed, any $\theta_\lambda = \theta_c^\star - \lambda\frac{\Delta_c}{\|\Delta_c\|}$ that maximizes $\|\nabla\psi(\theta_c^\star) - \nabla\psi(\theta_\lambda)\|$ and belongs to Θ_ρ must satisfy

$$\Big\langle\frac{\Delta_c}{\|\Delta_c\|}, \nabla\psi(\theta_c^\star) - \nabla\psi(\theta_\lambda)\Big\rangle \geq p\overline{\epsilon}_c$$

on the one hand, and on the other hand, since $\theta_c^\star, \theta_\lambda \in \Theta_\rho$,

$$\Big\langle\frac{\Delta_c}{\|\Delta_c\|}, \nabla\psi(\theta_c^\star) - \nabla\psi(\theta_\lambda)\Big\rangle \leq V_\rho\Big\|\frac{\Delta_c}{\|\Delta_c\|}\Big\|\,\|\theta_c^\star - \theta_\lambda\| = V_\rho\lambda.$$

Combining these two inequalities, we deduce that $\lambda_c \geq p\epsilon_c^{\max}/V_\rho$. Thus, using that $n_i/(n_{i+1}-1) \leq 1$ and $\tilde\epsilon_{t,i,c} \leq \epsilon_c^{\max}$, we deduce that $\lambda^\star = \frac{n_i p \tilde\epsilon_{t,i,c}}{(n_{i+1}-1)V_\rho} \leq \frac{p\overline{\epsilon}_c}{V_\rho} \leq \lambda_c$ is indeed satisfied. We then get without further restriction

$$\begin{aligned}
&\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n_i\leq n<n_{i+1}} E_{c,p}(n,t) \cap \|\nabla\psi(\hat\theta_n) - \nabla\psi(\theta_c^\star)\| \geq \epsilon_{t,i,c}\Big\} \\
&\qquad\leq \exp\Big(-\frac{n_i^2 p \epsilon_{t,i,c}^2}{2V_\rho(n_{i+1}-1)}\Big)\mathbf{1}\{\epsilon_{t,i,c} \leq \overline{\epsilon}_c\}. \qquad (13)
\end{aligned}$$

□

5.5. Combining the Different Steps

In this part, we recap what we have shown so far. Combining the peeling, change of measure, localization and concentration of measure steps of the four previous sections, we have shown that for all $\{\epsilon_{t,i,c}\}_{t,i}$

$$[1] \stackrel{\text{def}}{=} \mathbb{P}_{\theta^\star}\Big\{ \bigcup_{1\le n\le t} \hat{\theta}_n \in \Theta_\rho \cap \mathcal{K}_{a^\star}(\Pi_{a^\star}(\hat{\nu}_{a^\star,n}), \mu^\star - \epsilon) \ge f(t/n)/n \Big\}$$

$$\le \sum_{c=1}^{C_{p,\eta,K}} \sum_{i=0}^{I_t-1} \underbrace{\exp\big(-n_i\alpha^2 - \chi\sqrt{n_i f(t/n_i)}\big)}_{\text{change of measure}} \Big[\underbrace{\exp\Big(-\frac{n_i^2 p \epsilon_{t,i,c}^2}{2V_\rho(n_{i+1}-1)}\Big)\mathbf{1}\{\epsilon_{t,i,c} \le \overline{\epsilon}_c\}}_{\text{concentration}}$$

$$+ \underbrace{\alpha_{p,K}\exp\Big(-f\Big(\frac{t}{n_{i+1}-1}\Big)\Big)\min\Big\{\rho^2 v_\rho^2, \epsilon_{t,i,c}^2, \frac{(K+2)v_\rho^2}{K(n_{i+1}-1)V_\rho}\Big\}^{-K/2}\epsilon_{t,i,c}^K}_{\text{localization + change of measure}} \Big],$$

where we recall that $\alpha = \alpha(p,\eta,\epsilon) = \eta\rho_\epsilon\sqrt{v_\rho/2}$ and that the definition of n_i is

$$n_i = \begin{cases} b^i & \text{if } i < I_t \stackrel{\text{def}}{=} \lceil \log_b(\beta t + \beta) \rceil, \\ t+1 & \text{if } i = I_t. \end{cases}$$

A simple rewriting leads to the form

$$[1] \le \sum_{c=1}^{C_{p,\eta,K}} \sum_{i=0}^{I_t-1} \exp\big(-n_i\alpha^2 - \chi\sqrt{n_i f(t/n_i)}\ \big) \Big[\alpha_{p,K}\exp\Big(-f\Big(\frac{t}{n_{i+1}-1}\Big)\Big)$$

$$\times \max\Big\{\frac{\epsilon_{t,i,c}}{\rho v_\rho}, 1, \sqrt{\frac{(n_{i+1}-1)V_\rho}{1+2/K}}\frac{\epsilon_{t,i,c}}{v_\rho}\Big\}^K + \exp\Big(-\frac{n_i^2 p \epsilon_{t,i,c}^2}{2V_\rho(n_{i+1}-1)}\Big)\mathbf{1}\{\epsilon_{t,i,c} \le \overline{\epsilon}_c\}\Big],$$

which suggests that we use $\epsilon_{t,i,c} = \sqrt{\frac{2V_\rho(n_{i+1}-1)f(t/(n_{i+1}-1))}{pn_i^2}}$. Replacing this term in the above expression, we obtain

$$[1] \le \sum_{i=0}^{I_t-1} \exp\big(-n_i\alpha^2 - \chi\sqrt{n_i f(t/n_i)} - f(t/(n_{i+1}-1))\big) f(t/(n_{i+1}-1))^{K/2}$$

$$\times C_{p,\eta,K}\Big(\alpha_{p,K}\max\Big\{\frac{2V_\rho}{p\rho^2 v_\rho^2 b^{i-1}}, 1, \frac{b^2 V_\rho^2}{p v_\rho^2(\frac{1}{2}+\frac{1}{K})}\Big\}^{K/2} + 1\Big).$$

At this point, using the somewhat crude lower bound $b^i \ge 1$ it is convenient to introduce the constant

$$C(K,\rho,p,b,\eta) = C_{p,\eta,K}\Big(\alpha_{p,K}\max\Big\{\frac{2bV_\rho}{p\rho^2 v_\rho^2}, 1, \frac{b^2 V_\rho^2}{p v_\rho^2(\frac{1}{2}+\frac{1}{K})}\Big\}^{K/2} + 1\Big),$$

which leads to the final bound

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{1\le n\le t} \hat{\theta}_n \in \Theta_\rho \cap \mathcal{K}_{a^\star}(\Pi_{a^\star}(\hat{\nu}_{a^\star,n}), \mu^\star - \epsilon) \ge f(t/n)/n \Big\}$$

$$\le C(K,\rho,p,b,\eta)\sum_{i=0}^{I_t-1} \exp\big(-n_i\alpha^2 - \chi\sqrt{n_i f(t/n_i)} - f(t/(n_{i+1}-1))\big) f(t/(n_{i+1}-1))^{K/2}.$$

6. FINE-TUNED UPPER BOUNDS

In this section, we study the behavior of the bound obtained in Theorem 3 as a function of t for a specific choice of function f, namely $f(x) = \log(x) + \xi \log\log x$, and prove Corollaries 1 and 2 using a fine-tuning of the remaining free quantities. This tuning is not completely trivial, as a naive tuning yields the condition that $\xi > K/2 + 1$ to ensure that the final bound is $o(1/t)$, while proceeding with some more care enables us to show that $\xi > K/2 - 1$ is enough. Let us remind that f is nondecreasing only for $x \geq e^{-\xi}$. We thus restrict to $t \geq e^{-\xi}$ in Corollary 1 that uses the threshold $f(t)$, and to $\xi \geq 0$ in Corollary 2 that uses the threshold function $f(t/n)$. In the sequel, we use the short-hand notation C in order to replace $C(K, \rho, p, b, \eta)$.

6.1. Proof of Corollary 1

As a warm-up, we start by the boundary crossing probability involving $f(t)$ instead of $f(t/n)$. Indeed, controlling the boundary crossing probability with term $f(t/n)$ is more challenging. Although we focused so far on the boundary crossing probability with term $f(t/n)$, the previous proof directly applies to the case when $f(t)$ is considered. In particular, the result of Theorem 3 holds also when all the terms $f(t/n)$, $f(t/b^i)$, $f(t/b^{i+1})$ are replaced with $f(t)$.

With the choice $f(x) = \log(x) + \xi \log\log x$, which is nonincreasing on the set of x such that $\xi > -\log(x)$, Theorem 3 specifies for all $b > 1, p, q, \eta \in (0, 1)$ to

$$\begin{aligned}
&\mathbb{P}_{\theta^\star}\Big\{\bigcup_{1\leq n<t} \hat{\theta}_n \in \Theta_\rho \cap \mathcal{K}_{a^\star}(\Pi_{a^\star}(\hat{\nu}_{a^\star,n}), \mu^\star - \epsilon) \geq f(t)/n\Big\} \\
&\leq C \sum_{i=0}^{\lceil \log_b(qt)\rceil - 1} \exp\big(-\alpha^2 b^i - \chi\sqrt{b^i f(t)}\big) e^{-f(t)} f(t)^{K/2} \\
&= \frac{C}{t}\Bigg[\sum_{i=0}^{\lceil \log_b(qt)\rceil - 1} \underbrace{e^{-\alpha^2 b^i - \chi\sqrt{b^i f(t)}}}_{s_i}\Bigg] \log(t)^{K/2-\xi}\left(1 + \xi\frac{\log\log(t)}{\log(t)}\right)^{K/2}.
\end{aligned}$$

In order to study the sum $S = \sum_{i=0}^{\lceil \log_b(qt)\rceil - 1} s_i$, we provide two strategies. First, a direct upper bound gives $S \leq \lceil \log_b(qt)\rceil \leq \log_b(qt) + 1$. Thus, setting $q = 1$ and $b = 2$, we obtain

$$\begin{aligned}
&\mathbb{P}_{\theta^\star}\Big\{\bigcup_{1\leq n<t} \hat{\theta}_n \in \Theta_\rho \cap \mathcal{K}_{a^\star}(\Pi_{a^\star}(\hat{\nu}_{a^\star,n}), \mu^\star - \epsilon) \geq f(t)/n\Big\} \\
&\leq \frac{C}{t}\Bigg(1 + \underbrace{\xi\frac{\log\log(t)}{\log(t)}}_{o(1)}\Bigg)^{K/2} \log(t)^{-\xi+K/2}(\log_2(t) + 1).
\end{aligned}$$

This term is thus $o(1/t)$ whenever $\xi > K/2 + 1$ and $O(1/t)$ when $\xi = K/2 + 1$. We now show that a more careful analysis leads to a similar behavior even for smaller values of ξ. Indeed, let us note that for all $i \geq 0$, it holds by definition

$$\frac{s_{i+1}}{s_i} = \exp\big[-\chi b^{i/2}(b^{1/2} - 1) f(t)^{1/2} - \alpha^2 b^i(b-1)\big] \leq \exp\big[-\chi(b^{1/2} - 1) f(t)^{1/2}\big].$$

Since $f(t) \geq 1$ if we set $b = \lceil (1 + \frac{\log(1+\chi)}{\chi})^2\rceil$, which belongs to $(1, 4]$ for all $\chi \geq 0$, we obtain that $s_{i+1}/s_i \leq \frac{1}{1+\chi}$. Hence we deduce that

$$S \leq s_0 \sum_{i=0}^{\infty}(1+\chi)^{-i} = s_0\frac{1+\chi}{\chi} = \frac{1+\chi}{\chi}\exp(-\alpha^2 - \chi\sqrt{f(t)}).$$

Thus S is asymptotically $o(1)$, and we deduce that $\mathbb{P}_{\theta^\star}\big\{\bigcup_{1\leq n<t} \hat{\theta}_n \in \Theta_\rho \cap \mathcal{K}_{a^\star}(\Pi_{a^\star}(\hat{\nu}_{a^\star,n}), \mu^\star - \epsilon) \geq f(t)/n\big\} = o(1/t)$ beyond the condition $\xi > K/2 + 1$. It is interesting to note that due to the term $-\chi\sqrt{f(t)}$ in the exponent, and owing to the fact that $\alpha\sqrt{\log(t)} - \beta\log\log(t) \to \infty$ for all positive α and all β, we actually have the stronger property that $S\log(t)^{-\xi+K/2} = o(1)$ for all ξ (using $\alpha = \chi$ and $\beta = K/2 - \xi$). However, since this asymptotic regime may take a massive amount of time to kick-in when $\alpha/\beta < 1/2$, we do not advise to take ξ smaller than $K/2 - 2\chi$. All in all, we obtain, for $C = C(K, b, \rho, p, \eta)$ with $b = \lceil (1 + \frac{\log(1+\chi)}{\chi})^2 \rceil \le 4$,

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{1\le n<t} \hat\theta_n \in \Theta_\rho \cap \mathcal{K}_{a^\star}(\Pi_{a^\star}(\hat\nu_{a^\star,n}), \mu^\star - \epsilon) \ge f(t)/n \Big\}$$
$$\le \frac{C(1+\chi)}{t\chi}\bigg(1 + \xi\frac{\log\log(t)}{\log(t)}\bigg)^{K/2} \log(t)^{-\xi+K/2} \exp\big(-\chi\sqrt{\log(t) + \xi\log\log(t)}\ \big).$$

6.2. Proof of Corollary 2

Let us now focus on the proof of Corollary 2 involving the threshold $f(t/n)$. We consider the choice $f(x) = \log(x) + \xi\log\log x$, which is nonincreasing on the set of x such that $\xi > -\log(x)$. When $x = t/n$ and n is about $t - O(\log(t))$, ensuring this monotonicity property means that we require ξ to dominate $\log(1 - O(\log(t)/t))$, that is $\xi \ge 0$. Now, following the result of Theorem 3, we obtain for all $b > 1, p, q, \eta \in (0, 1)$,

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{1\le n<t} \hat\theta_n \in \Theta_\rho \cap \mathcal{K}_{a^\star}(\Pi_{a^\star}(\hat\nu_{a^\star,n}), \mu^\star - \epsilon) \ge f(t/n)/n \Big\}$$
$$\le C\exp\bigg(-\frac{\alpha^2 q}{b}t - \chi\sqrt{\frac{tqf(b/q)}{b}}\bigg)$$
$$+ C\sum_{i=0}^{\lceil\log_b(qt)\rceil-2} \exp\big(-\alpha^2 b^i - \chi\sqrt{b^i f(t/b^i)} - f(t/(b^{i+1}-1))\big) f\Big(\frac{t}{b^{i+1}-1}\Big)^{K/2}$$
$$= Ce^{-\frac{\alpha^2 qt}{b} - \sqrt{\frac{\chi^2 tqf(b/q)}{b}}} + C\sum_{i=0}^{\lceil\log_b(qt)\rceil-2} \underbrace{e^{-\alpha^2 b^i - \chi\sqrt{b^i f(t/b^i)}}\Big(\frac{b^{i+1}-1}{t}\Big)\log\Big(\frac{t}{b^{i+1}-1}\Big)^{K/2-\xi}}_{s_i}$$
$$\times\Bigg(1 + \underbrace{\xi\frac{\log\log\big(\frac{t}{b^{i+1}-1}\big)}{\log\big(\frac{t}{(b^{i+1}-1)}\big)}}_{o(1)}\Bigg)^{K/2}. \tag{14}$$

We thus study the sum $S = \sum_{i=0}^{\lceil\log_b(qt)\rceil-2} s_i$. To this end, let us first study the term s_i. Since $i \mapsto \log(t/b^{i+1})$ is a decreasing function of i, it holds for any index $i_0 \in \mathbb{N}$ that

$$s_i \le \begin{cases} \big(\frac{b^{i+1}}{t}\big)\log\big(\frac{t}{b-1}\big)^{-\xi+K/2} & \text{if } \xi \le K/2, \quad i \le i_0, \\ \big(\frac{b^{i+1}}{t}\big)\log\big(\frac{t}{b^{i_0+1}-1}\big)^{-\xi+K/2} & \text{if } \xi \ge K/2, \quad i \le i_0, \\ \exp(-\chi\sqrt{b^i f(t/b^i)}\ )\big(\frac{b^{i+1}}{t}\big)\log\big(\frac{t}{b^{i_0+1}-1}\big)^{-\xi+K/2} & \text{if } \xi \le K/2, \quad i \ge i_0, \\ \exp(-\chi\sqrt{b^i f(t/b^i)}\ )\big(\frac{b^{i+1}}{t}\big)\log\big(\frac{1}{q}\big)^{-\xi+K/2} & \text{if } \xi \ge K/2, \quad i \ge i_0. \end{cases}$$

Small values of i. We start by handling the terms corresponding to small values of $i \le i_0$ for some i_0 to be chosen. In that case, we note that $r_i = \frac{b^{i+1}}{t}$ satisfies $r_{i-1}/r_i = 1/b < 1$ and thus

$$\sum_{i=0}^{i_0} s_i \le s_{i_0}\sum_{i=0}^{\infty}(1/b)^i = \frac{bs_{i_0}}{b-1},$$

from which we deduce that

$$\sum_{i=0}^{i_0} s_i \leq \begin{cases} \big(\frac{bb^{i_0+1}}{t(b-1)}\big)\log\big(\frac{t}{b^{i_0+1}}\big)^{K/2-\xi} & \text{if } \xi \geq K/2,\\ \big(\frac{bb^{i_0+1}}{t(b-1)}\big)\log\big(\frac{t}{b-1}\big)^{K/2-\xi} & \text{if } \xi \leq K/2.\end{cases}$$

Following [16], in order to ensure that this quantity is summable in t, it is convenient to define i_0 as

$$i_0 = \lfloor \log_b(t_0)\rfloor, \quad \text{where} \quad t_0 = \frac{1}{c\log(ct)^\eta},$$

for $\eta > K/2-\xi$ and a positive constant c. Indeed in that case when $i_0 \geq 0$ we obtain the bounds[3]

$$\sum_{i=0}^{i_0} s_i \leq \frac{b^2}{(b-1)ct\log(tc)^\eta} \times \begin{cases}\log(tc\log(tc)^\eta/b)^{K/2-\xi} & \text{if } \xi \geq K/2,\\ \log(t/(b-1))^{K/2-\xi} & \text{if } \xi \leq K/2.\end{cases}$$

We easily see that this is $o(1/t)$ both when $\xi > K/2$ and when $\xi \leq K/2$ by construction of η. Note that η can further be chosen to be equal to 0 when $\xi > K/2$. The value of c is fixed by looking at what happens for larger values of $i \geq i_0$. We note that the initial proof in [16] uses the value $\eta = 1$.

Large values of i. We now consider the terms of the sum S corresponding to large values $i > i_0$ and thus focus on the term $s'_i = \exp(-\chi\sqrt{b^i\log(t/b^i)})b^{i+1}$, or, better, on the ratio

$$\frac{s'_{i+1}}{s'_i} = b\exp\Big[-\chi b^{i/2}\Big(b^{1/2}\log\Big(\frac{t}{b^i b}\Big)^{1/2} - \log\Big(\frac{t}{b^i}\Big)^{1/2}\Big)\Big].$$

Remarking that this ratio is a nonincreasing function of i, we upper bound it by replacing i with either i_0+1 or 0. Using that $b^{i_0+1} \leq t_0$ we thus obtain

$$\frac{s'_{i+1}}{s'_i} \leq \begin{cases} b\exp\Big[-\sqrt{\frac{\chi^2}{c}}\Big(\sqrt{\frac{b\log(tc\log(tc)^\eta/b)}{\log(tc)^\eta}} - \sqrt{\frac{\log(tc\log(tc)^\eta)}{\log(tc)^\eta}}\Big)\Big] & \text{if } i_0 \geq 0,\\ b\exp\Big[-\chi\Big(\sqrt{b\log(t/b)} - \sqrt{\log(t)}\Big)\Big] & \text{else.}\end{cases}$$

Since we would like this ratio to be less than 1 for all (large enough) t, we readily see from this expression that this excludes the cases when $\eta > 1$: the term in the inner brackets converges to 0 in such cases, and thus the ratio is asymptotically upper bounded by $b > 1$. Thus we assume that $\eta \leq 1$, that is $\xi \geq K/2-1$.

For the critical value $\eta = 1$ it is then natural to study the term $\sqrt{\frac{b\log(x\log(x)/b)}{\log(x)}} - \sqrt{\frac{\log(x\log(x))}{\log(x)}}$. First, when $b = 4$, this quantity is larger than $1/2$ for $x \geq 8.2$. Then, it can be checked that $4\exp(-\frac{1}{2}\sqrt{\chi^2/c}) < 1$ if $c > \chi^2/(2\log(4))^2$. These two conditions show that $\frac{s'_{i+1}}{s'_i} < 1$ for

$$t \geq 8.2(2\log(4))^2\chi^{-2} \simeq 63\chi^{-2}.$$

Now, in order to get the ratio $\frac{s'_{i+1}}{s'_i}$ away from 1, we target the bound $\frac{s'_{i+1}}{s'_i} < b/(b+1)$. This can be achieved by requiring that $t \geq 8.2(2\log(5))^2\chi^{-2} \simeq 85\chi^{-2}$ and setting $c = \chi^2/(2\log(5))^2$. Eventually, we obtain for $b = 4$ and $t \geq 85\chi^{-2}$ the bound

$$\sum_{i=i_0+1}^{I_t-2} s'_i \leq s'_{i_0+1}\sum_{i=i_0+1}^{I_t-2}(b/(b+1))^{i-i_0-1} \leq s'_{i_0+1}(b+1)$$
$$\leq (b+1)\exp\big[-\chi\sqrt{bt_0\log(t/bt_0)}\,\big]b^2t_0 \leq b^2(b+1)t_0.$$

Remark 12. Another notable value is $\eta = 0$. A similar study to the previous one shows that for $b = 3.5$, the term $\sqrt{b\log(x/b)} - \sqrt{\log(x)}$ is larger than $1/2$ for $x > 12$, which entails that $\frac{s'_{i+1}}{s'_i} < b/(b+1)$ provided that $t \geq 12(2\log(3.5))^2\chi^{-2} \simeq 76\chi^{-2}$.

[3]This is also valid when $i_0 < 0$ since the sum is equal to 0 in that case.

Plugging-in the definition of t_0 and since $b^{i_0+1} \leq bt_0$, we obtain for $i_0 \geq 0$, $b = 4$, and $c = \chi^2/(2\log(5))^2$,

$$\sum_{i=i_0+1}^{I_t-2} s_i \leq \begin{cases} \frac{b^2(b+1)}{tc\log(tc)} \log(1/q)^{K/2-\xi} & \text{if } \xi \geq K/2, \\ \frac{b^2(b+1)}{tc\log(tc)} \log\left(t\frac{c\log(tc)}{b-c\log(tc)}\right)^{K/2-\xi} & \text{if } \xi \in [K/2-1, K/2]. \end{cases} \tag{15}$$

It remains to handle the case when $i_0 < 0$. Note that this case only happens for t large enough so that $t > c^{-1}e^{\frac{1}{bc}}$. This quantity may be huge, since $1/bc = \log(5)^2\chi^{-2}$ becomes large when χ is close to 0. In that case, we directly control $\sum_{i=0}^{I_t-2} s_i$. We control the ratio s'_{i+1}/s'_i by $b/(b+1/2)$ provided that

$$\sqrt{b\log(t/b)} - \sqrt{\log(t)} > \frac{\log(b+1/2)}{\chi}, \qquad \text{where} \quad b = 4.$$

Thus, if we define t_χ to be the smallest such t, then when $t > c^{-1}e^{\frac{1}{bc}}$ and provided that $t \geq t_\chi$, the bound of (15) remains valid for the sum S, up to replacing $b^2(b+1)$ with $2b^2(b+1/2)$ and $\log\left(t\frac{c\log(tc)}{b-c\log(tc)}\right)$ with $\log(t/(b-1))$. The constraint $t \geq t_\chi$ is satisfied as soon as $4\log(5)^2\chi^{-2}e^{\chi^{-2}\log(5)^2} \geq t_\chi$, which is generally satisfied for χ not too large.

Final control on S. We can now control the term S by combining the two bounds for large and small i. We get for $c = \chi^2/(2\log(4.5))^2$ and $b = 4$, and provided that $t \geq 85\chi^{-2}$ and $t \leq \chi^{-2}\frac{\exp\left(\chi^{-2}\log(4.5)^2\right)}{4\log(4.5)^2}$, the following bound

$$S \leq \frac{b}{ct\log(tc)} \begin{cases} \frac{b}{(b-1)}\log(tc\log(tc)/b)^{K/2-\xi} + b(b+1)\log(1/q)^{K/2-\xi} \\ \qquad \text{if } \xi \geq K/2, \\ \frac{b}{(b-1)}\log(t/(b-1))^{K/2-\xi} + b(b+1)\log\left(t\frac{c\log(tc)}{b-c\log(tc)}\right)^{K/2-\xi} \\ \qquad \text{if } \xi \in [K/2-1, K/2]. \end{cases} \tag{16}$$

Further, for larger values of t, $t \geq \chi^{-2}\frac{\exp\left(\chi^{-2}\log(4.5)^2\right)}{4\log(4.5)^2}$, we have

$$S \leq \frac{2b^2(b+1/2)}{ct\log(tc)} \begin{cases} \log(1/q)^{K/2-\xi} & \text{if } \xi \geq K/2, \\ \log(t/(b-1))^{K/2-\xi} & \text{if } \xi \in [K/2-1, K/2]. \end{cases} \tag{17}$$

Concluding step. In this final step, we combine equation (14) with the bounds (16), (17) on S. We obtain that for all $p, q, \eta \in (0,1)$

$$\mathbb{P}_{\theta^\star}\Big\{\bigcup_{1\leq n<t} \hat{\theta}_n \in \Theta_\rho \cap \mathcal{K}_{a^\star}(\Pi_{a^\star}(\hat{\nu}_{a^\star,n}), \mu^\star - \epsilon) \geq f(t/n)/n\Big\}$$
$$\leq C(K,\rho,p,b,\eta)\Big(e^{-\frac{\alpha^2 qt}{b} - \sqrt{\frac{\chi^2 tqf(b/q)}{b}}} + S(1+\xi)^{K/2}\Big),$$

where we recall the definition of the constants $\alpha = \eta\rho_\epsilon\sqrt{v_\rho/2}$, $\chi = p\eta\rho_\epsilon\sqrt{2v_\rho^2/V_\rho}$.

When $\xi \in [K/2-1, K/2]$, one can choose $q = 1$. When $\xi \geq K/2$, there is a trade-off in q, since the first term (the exponential) is decreasing with q while the second term is increasing with q. For instance, choosing $q = \exp(-\kappa^{-1/\eta})$, where $\eta = \xi - K/2$ and $\kappa > 0$, leads to $\log(1/q)^{K/2-\xi} = \kappa$. When $b = 4$, simply choosing $q = 0.8$ gives the final bound after some cosmetic simplifications.

APPENDIX: TECHNICAL DETAILS

Lemma 9 (Dimension 1). *Consider a canonical one-dimensional family rm(that is $K = 1$ and $F_1(x) = x \in \mathbb{R}$). Then, for all f such that $f(t/n)/n$ is nonincreasing in n,*

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{m \le n < M} \mathcal{B}^\psi(\hat\theta_n, \theta^\star) \ge f(t/n)/n \Big\} \le \exp\Big(-\frac{m}{M} f(t/M) \Big).$$

Proof. Observe that

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{m \le n < M} \mathcal{B}^\psi(\hat\theta_n, \theta^\star) \ge f(t/n)/n \Big\} = \mathbb{P}_{\theta^\star}\Big\{ \bigcup_{m \le n < M} \Phi^\star(\hat F_n) \ge f(t/n)/n \Big\}$$
$$\le \mathbb{P}_{\theta^\star}\Big\{ \bigcup_{m \le n < M} \Phi^\star(\hat F_n) \ge f(t/M)/M \Big\}.$$

At this point note that if $\Phi^\star(F) < f(t/M)/M$ for all $F = \nabla\psi(\theta)$ with mean $\mu_\theta \le \mu^\star - \epsilon$, then the probability of interest is 0 and we are done. In the other case, there exists an F_M such that $\Phi^\star(F_M) = f(t/M)/M$. We thus proceed with this case as follows

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{m \le n < M} \mathcal{B}^\psi(\hat\theta_n, \theta^\star) \ge f(t/n)/n \Big\} \le \mathbb{P}_{\theta^\star}\Big\{ \bigcup_{m \le n < M} \Phi^\star(\hat F_n) \ge \Phi^\star(F_M) \Big\}$$
$$\stackrel{(a)}{\le} \mathbb{P}_{\theta^\star}\Big\{ \bigcup_{m \le n < M} \hat F_n \le F_M \Big\} \stackrel{(b)}{\le} \mathbb{P}_{\theta^\star}\Big\{ \bigcup_{m \le n < M} \exp\Big(\lambda \sum_{i=1}^n F(X_i) \Big) \ge \exp(n\lambda F_M) \Big\}$$
$$\le \mathbb{P}_{\theta^\star}\Big\{ \bigcup_{m \le n < M} \exp\Big(\sum_{i=1}^n \Big(\lambda F(X_i) - \Phi(\lambda) \Big) \Big) \ge \exp\big(n[\lambda F_M - \Phi(\lambda)] \big) \Big\}$$
$$\stackrel{(c)}{\le} \mathbb{P}_{\theta^\star}\Big\{ \max_{m \le n < M} \exp\Big(\sum_{i=1}^n \Big(\lambda F(X_i) - \Phi(\lambda) \Big) \Big) \ge \exp\big(m[\lambda F_M - \Phi(\lambda)] \big) \Big\},$$

where (a) holds by (4), (b) holds for all $\lambda < 0$, and (c) for all $\lambda < 0$ such that $\lambda F_M - \Phi(\lambda) > 0$. Now, the process defined by $W_{\lambda,0} = 1$ and $W_{\lambda,n} = \exp\big(\sum_{i=1}^n (\lambda F(X_i) - \Phi(\lambda)) \big)$ is a nonnegative supermartingale, since it holds

$$\mathbb{E}_{\theta^\star}\Big[\exp\Big(\sum_{i=1}^n (\lambda F(X_i) - \Phi(\lambda)) \Big) \Big| \mathcal{H}_{n-1} \Big] = W_{\lambda,n-1} \mathbb{E}_{\theta^\star}\big[\exp\big(\lambda F(X_n) - \Phi(\lambda) \big) \big| \mathcal{H}_{n-1} \big]$$
$$\le W_{\lambda,n-1} \exp\big(\Phi(\lambda) - \Phi(\lambda) \big) \le 1.$$

Thus we deduce that for all $\lambda < 0$ such that $\lambda F_M - \Phi(\lambda) > 0$

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{m \le n < M} \mathcal{B}^\psi(\hat\theta_n, \theta^\star) \ge f(t/n)/n \Big\} \le \exp\big(-m[\lambda F_M - \Phi(\lambda)] \big).$$

Since by (5) this is satisfied by the optimal λ for $\Phi^\star(F_M)$, we obtain

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{m \le n < M} \mathcal{B}^\psi(\hat\theta_n, \theta^\star) \ge f(t/n)/n \Big\} \le \exp\big(-m\Phi^\star(F_M) \big) = \exp\Big(-\frac{m}{M} f(t/M) \Big).$$

□

Lemma 14. *For all $\epsilon, \epsilon' > 0$, $p, p' \in [0,1]$ and all $K \ge 1$ the following equality holds:*

$$\frac{|\mathcal{B}_2(0,\epsilon) \cap \mathcal{C}_p(0;\mathbf{1})|}{\int_{\mathcal{B}_2(0,\epsilon') \cap \mathcal{C}_{p'}(0;\mathbf{1})} e^{-\|y\|^2/2}\, dy} = \frac{\omega_{p,K-2}}{\omega_{p',K-2}} \frac{\int_0^\epsilon r^{K-1}\, dr}{\int_0^{\epsilon'} e^{-r^2/2} r^{K-1}\, dr},$$

where $\omega_{p,K-2} = \int_p^1 \sqrt{1-z^2}^{K-2} dz$ *for* $K \geq 2$ *and using the convention that* $\omega_{p,-1} = 1$. *Further,*

$$\frac{|\mathcal{B}_2(0,\epsilon) \cap \mathcal{C}_p(0;\mathbf{1})|}{\int_{\mathcal{B}_2(0,\epsilon') \cap \mathcal{C}_{p'}(0;\mathbf{1})} e^{-\|y\|^2/2}\, dy} \leq 2\frac{\omega_{p,K-2}}{\omega_{p',K-2}}\Big(\frac{\epsilon}{\min\{\epsilon', \sqrt{1+2/K}\}}\Big)^K.$$

Proof. First of all, remark that for $K \geq 2$ it holds

$$\begin{aligned}
|\mathcal{B}_2(0,\epsilon) \cap \mathcal{C}_p(0;\mathbf{1})| &= \int_0^\epsilon |\{y \in \mathbb{R}^K : \langle y, \mathbf{1}\rangle \geq rp, \|y\| = r\}|\, dr \\
&= \int_0^\epsilon \int_{rp}^r |\{y \in \mathbb{R}^K : y_1 = z, \|y\| = r\}|\, dz dr = \int_0^\epsilon \int_{rp}^r |\{y \in \mathbb{R}^{K-1} : \|y\| = \sqrt{r^2 - z^2}\}|\, dz dr \\
&= \int_0^\epsilon r^{K-1} \int_p^1 \sqrt{1-z^2}^{K-2} |\mathcal{S}_{K-1}|\, dz dr.
\end{aligned}$$

where $\mathcal{S}_{K-1} \subset \mathbb{R}^{K-1}$ is the $(K-2)$-dimensional unit sphere of $\mathbb{R}^{K-1}$. Let us recall that when $K = 2$, we get $|\mathcal{S}_{K-1}| = 2$. For convenience, let us denote $\omega_{p,K-2} = \int_p^1 \sqrt{1-z^2}^{K-2} dz$. Then, for $K \geq 2$,

$$|\mathcal{B}_2(0,\epsilon) \cap \mathcal{C}_p(0;\mathbf{1})| = |\mathcal{S}_{K-1}| \int_0^\epsilon r^{K-1} \omega_{p,K-2}\, dr.$$

For $K = 1$, $|\mathcal{B}_2(0,\epsilon) \cap \mathcal{C}_p(0;\mathbf{1})| = \epsilon$. Likewise, we obtain, following the same steps that

$$\int_{\mathcal{B}_2(0,\epsilon) \cap \mathcal{C}_p(0;\mathbf{1})} e^{-\|y\|^2/2}\, dy = |\mathcal{S}_{K-1}| \int_0^\epsilon e^{-r^2/2} r^{K-1} \omega_{p,K-2}\, dr.$$

We obtain the first part of the lemma by combining the two previous equalities. For the second part, we use the inequality $e^{-x} \geq 1 - x$, which gives

$$\int_0^\epsilon e^{-r^2/2} r^{K-1}\, dr \geq \int_0^\epsilon r^{K-1} - \frac{1}{2} r^{K+1}\, dr = \epsilon^K \Big(\frac{1}{K} - \frac{\epsilon^2}{2(K+2)}\Big).$$

Thus, whenever $\epsilon^2 < (K+2)/K$, we obtain

$$\int_0^\epsilon e^{-r^2/2} r^{K-1}\, dr \geq \frac{\epsilon^K}{2K}.$$

On the other hand, if $\epsilon^2 \geq (K+2)/K$, then

$$\int_0^\epsilon e^{-r^2/2} r^{K-1}\, dr \geq \int_0^{(K+2)/K} e^{-r^2/2} r^{K-1}\, dr \geq \frac{\sqrt{1+2/K}^K}{2K}.$$

Thus, in all cases, the integral is larger than $\frac{\min\{\epsilon, \sqrt{1+2/K}\}^K}{2K}$, and we conclude by simple algebra. □

CONCLUSION

In this work, that should be considered as a tribute to the contributions of T.L. Lai, we shed light on a beautiful and seemingly forgotten result from [16], that we modernized into a fully nonasymptotic statement, with explicit constants that can be directly used, for instance, for the regret analysis of multi-armed bandit strategies. Interestingly, the final results, whose roots are thirty-years old, show that the existing analysis of `KL-ucb` that was only stated for exponential families of dimension 1 and discrete distributions lead to a sub-optimal constraints on the tuning of the threshold function f, and can be extended to work with exponential families of arbitrary dimension K and even for the thresholding term of the `KL-ucb+` strategy, whose analysis was left open.

This proof technique is mostly based on a change-of-measure argument, like the lower bounds for the analysis of sequential decision-making strategies and in stark contrast with other key results in the literature [13, 17, 8]. We believe and hope that the novel writing of this proof technique that we provided here will greatly benefit the community working on boundary crossing probabilities, sequential design of experiments as well as stochastic decision-making strategies.

ACKNOWLEDGMENTS

The author acknowledges the support of the French Agence Nationale de la Recherche (ANR) under grant ANR-16-CE40-0002 (project BADASS), the French Ministry of Higher Education and Research, CPER Nord-Pas de Calais/FEDER DATA Advanced data science and technologies 2015-2020 and Inria Lille — Nord Europe.

REFERENCES

1. Rajeev Agrawal, "Sample Mean Based Index Policies by $o(\log n)$ Regret for the Multi-Armed Bandit Problem", Adv. in Appl. Probab. **27** (04), 1054–1078 (1995).
2. J.-Y. Audibert, R. Munos, and Cs. Szepesvári, "Exploration-Exploitation Trade-Off Using Variance Estimates in Multi-Armed Bandits", Theoret. Comp. Sci. **410** (19), (2009).
3. J.-Y. Audibert and S. Bubeck, "Regret Bounds and Minimax Policies under Partial Monitoring", J. Machine Learning Res. **11**, 2635–2686 (2010).
4. P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-Time Analysis of the Multiarmed Bandit Problem", Machine Learning **47** (2), 235–256 (2002).
5. L. M. Bregman, "The Relaxation Method of Finding the Common Point of Convex Sets and Its Application to the Solution of Problems in Convex Programming", USSR Comput. Math. and Math. Phys. (Elsevier) **7** (3), 200–217 (1967).
6. S. Bubeck, N. Cesa-Bianchi, et al., "Regret Analysis of Stochastic and Nonstochastic Multi-Armed Bandit Problems", Foundations and Trends® in Machine Learning **5** (1), 1–122 (2012).
7. A. N. Burnetas and M. N. Katehakis, "Optimal Adaptive Policies for Markov Decision Processes", in *Mathematics of Operations Research* (1997), pp. 222–255.
8. O. Cappé, A. Garivier, O.-A. Maillard, R. Munos, and G. Stoltz, "Kullback–Leibler Upper Confidence Bounds for Optimal Sequential Allocation", Ann. Statist. **41** (3), 1516–1541 (2013).
9. Y. S. Chow and H. Teicher, *Probability Theory*, 2nd. ed. (Springer, 1988).
10. I. H. Dinwoodie, "Mesures dominantes et théoreme de sanov", in *Annales de l'IHP Probabilités et statistiques* (1992), Vol. 28, pp. 365–373.
11. A. Garivier, P. Ménard, and G. Stoltz, *Explore first, Exploit Next: The True Shape of Regret in Bandit Problems* arXiv preprint arXiv:1602.07182 (2016).
12. J. C. Gittins, "Bandit Processes and Dynamic Allocation Indices", J. Roy. Statist. Soc., Ser. B **41** (2), 148–177 (1979).
13. J. Honda and A. Takemura, "An Asymptotically Optimal Bandit Algorithm for Bounded Support Models", in *Conf. Comput. Learning Theory*, Ed. by T. Kalai and M. Mohri (Haifa, Israel, 2010).
14. T. L. Lai and H. Robbins, "Asymptotically Efficient Adaptive Allocation Rules", Advances in Appl. Math. **6** (1), 4–22 (1985).
15. T. L. Lai, "Adaptive Treatment Allocation and the Multi-Armed Bandit Problem", Ann. Statist, 1091–1114 (1987).
16. T. L. Lai, "Boundary Crossing Problems for Sample Means", Ann. Probab., 375–396 (1988).
17. O.-A. Maillard, R. Munos, and G. Stoltz, "A Finite-Time Analysis of Multi-Armed Bandits Problems with Kullback–Leibler Divergences", in *Proc. 24th Conference On Learning Theory* (Budapest, Hungary), 497–514 (2011).
18. H. Robbins, "Some Aspects of the Sequential Design of Experiments", Bull. Amer. Math. Soc. **58** (5), 527–535 (1952).
19. H. Robbins, *Herbert Robbins Selected Papers* (Springer, 2012).
20. W. R. Thompson, "On the Likelihood That One Unknown Probability Exceeds Another in View of the Evidence of Two Samples", Biometrika **25** (3/4), 285–294 (1933).
21. W. R. Thompson, "On a Criterion for the Rejection of Observations and the Distribution of the Ratio of Deviation to Sample Standard Deviation", Ann. Math. Statist. **6** (4), 214–219 (1935).
22. A. Wald, "Sequential Tests of Statistical Hypotheses", Ann. Math. Statist. **16** (2), 117–186 (1945).