Boundary Crossing for General Exponential Families

Odalric-Ambrym Maillard

Journal of Machine Learning Research 1:1–34, 2017 Algorithmic Learning Theory 2017

Boundary Crossing for General Exponential Families

Odalric-Ambrym Maillard ODALRIC.MAILLARD@INRIA.FR
INRIA Lille - Nord Europe, Villeneuve d'Ascq, France

Editors: Steve Hanneke and Lev Reyzin

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 $\mathcal{B}^\psi(\widehat{\theta}_n, \theta^\star) \geqslant f(t/n)/n$, where $\theta^\star$ is the parameter of an unknown target distribution, $\widehat{\theta}_n$ 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 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.

1. Multi-armed bandit setup and notations

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 distribution 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 her past observations and observes a stochastic payoff Y_t drawn independently at random according to the distribution ν_{a_t}. Her goal is to maximize her expected cumulated payoff over a possibly unknown number of steps.

Although the term multi-armed bandit problem was probably coined during the 60's in 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 mind of his time, see Robbins (1952) and takes its origin in earlier questions about optimal stopping policies for clinical trials, see Thompson (1933, 1935); Wald (1945). We refer the interested reader to Robbins (2012) 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 Cappé et al. (2013) for key historical aspects about the development of the field. Most notably, they include first the introduction of dynamic allocation indices (aka Gittins indices, Gittins (1979); Lattimore (2016)) 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 (1985) that shows indices 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 by Burnetas and Katehakis (1997) (see also the recent work from Garivier et al. (2016)) as well as the asymptotic analysis by Agrawal (1995) of generic classes of upper-confidence-bound based index policies. Finally Auer et al. (2002) popularized a simple sub-optimal index strategy termed `UCB` and most importantly opened the quest for finite-time, as opposed to asymptotic, performance guarantees. We now recall the necessary formal definitions and notations, closely following Cappé et al. (2013).

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 $a^\star \in \mathrm{Argmax}_{a\in\mathcal{A}}\, \mu_a$. 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 $N_a(T) \stackrel{\text{def}}{=} \sum_{t=1}^{T} \mathbb{I}_{\{a_t=a\}}$.

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

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

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 number of pulls of a given arms are considered. The first series of averages are referred to by using a functional notation for the indexation in the global time: $\widehat{\nu}_a(t)$, while the second series are 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 $m \geqslant 1$, we denote by $\tau_{a,m}$ the round at which arm a was pulled for the m–th time, that is $\tau_{a,m} = \min\big\{t \in \mathbb{N} :\ N_a(t) = m\big\}$. For each round t such that $N_a(t) \geqslant 1$, we then define the following two empirical distributions*

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

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

Lemma 3 [1] *The random variables $X_{a,m} = Y_{\tau_{a,m}}$, where $m = 1, 2, \ldots$, are independent and identically distributed according to ν_a. Moreover, we have the rewriting $\widehat{\nu}_a(t) = \widehat{\nu}_{a,N_a(t)}$.*

1. We refer to Cappé et al. (2013) for a proof of this elementary result.

2. Boundary crossing probabilities for the generic KL-ucb strategy.

The first appearance of the KL-ucb strategy can be traced at least to Lai (1987) although it was not given an explicit name at that time. It seems the strategy was forgotten after the work of Auer et al. (2002) that opened a decade of intensive research on finite-time analysis of bandit strategies and extensions to variants of the problem (Audibert et al. (2009); Audibert and Bubeck (2010), see also Bubeck et al. (2012) for a survey of relevant variants of bandit problems), until the work of Honda and Takemura (2010) 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 by Maillard et al. (2011) for discrete distributions, soon extended to handle exponential families of dimension 1 in the unifying work of Cappé et al. (2013). However, as we will see in this paper, we should all be much in dept of the 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 19 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 providing 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 $\mathfrak{M}_1(\mathbb{R})$ denotes the set of all probability distributions over $\mathbb{R}$. 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 expectation operator is denoted by E while expectations with respect to underlying randomizations are referred to as $\mathbb{E}$.)

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 non-decreasing function f, which is typically such that $f(t) \approx \log(t)$. At each round t, an upper confidence bound $U_a(t)$ is computed on the expectation μ_a of the distribution ν_a of each arm; an arm with highest upper confidence bound is then played.

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

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

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

for each round $t+1$, where $t \geqslant K$, **do**

for each arm a, compute $U_a(t) = \sup\left\{ E(\nu) : \ \nu \in \mathcal{D} \text{ and } \mathtt{KL}\Big(\Pi_{\mathcal{D}}\big(\widehat{\nu}_a(t)\big), \nu\Big) \leqslant \frac{f(t)}{N_a(t)} \right\}$;

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. Under some assumption, one can rewrite this term in an equivalent form more suitable for the analysis:

Lemma 4 (Rewriting, see Cappé et al. (2013)) *Under the assumption that*

Assumption 1 *A known interval $\mathcal{I}\subset\mathbb{R}$ with boundary $\mu^-\leqslant\mu^+$ is such that each model $\mathcal{D}=\mathcal{D}_a$ of probability measures for $a\in\mathcal{A}$ satisfies $\mathcal{D}_a\subset\mathcal{P}(\mathcal{I})$ and $\forall\nu\in\mathcal{D}_a, \mu\in\mathcal{I}\setminus\{\mu^+\}$,*

$$\inf\Big\{\mathtt{KL}(\nu,\nu'):\quad \nu'\in\mathcal{D}_a\ \text{ s.t. }\ E(\nu')>\mu\Big\}=\min\Big\{\mathtt{KL}(\nu,\nu'):\quad \nu'\in\mathcal{D}_a\ \text{ s.t. }\ E(\nu')\geqslant\mu\Big\},$$

then the upper bound used by the `KL-ucb` *algorithm satisfies the following equality*

$$U_a(t) \;=\; \max\Big\{\mu\in\mathcal{I}\setminus\{\mu^+\}:\ \mathcal{K}_a\Big(\Pi_a\left(\widehat{\nu}_a(t)\right),\mu\Big)\leqslant\frac{f(t)}{N_a(t)}\Big\}$$
$$\text{where}\quad \mathcal{K}_a(\nu_a,\mu^\star)\stackrel{\text{def}}{=}\inf_{\nu\in\mathcal{D}_a:\,E(\nu)>\mu^\star}\mathtt{KL}(\nu_a,\nu)\quad\text{and}\quad\Pi_a\stackrel{\text{def}}{=}\Pi_{\mathcal{D}_a}.$$

Likewise, a similar result holds for `KL-ucb+` *but where $f(t)$ is replaced with $f(t/N_a(t))$.*

Remark 5 *For instance, this assumption is valid when $\mathcal{D}_a=\mathcal{P}([0,1])$ and $\mathcal{I}=[0,1]$. Indeed we can replace the strict inequality with an inequality provided that $\mu<1$ by Honda and Takemura (2010), and the infimum is reached by lower semi-continuity of the* `KL` *divergence and convexity and closure of the set $\{\nu'\in\mathcal{P}([0,1])\ \text{ s.t. }\ E(\nu')\geqslant\mu\}$.*

Using boundary-crossing probabilities for regret analysis We continue this warming-up by restating a convenient way to decompose the regret and make appear the *boundary crossing probabilities*. The following lemma is a direct adaptation[2] from Cappé et al. (2013):

Lemma 6 (From Regret to Boundary Crossing Probabilities) *Let $\varepsilon\in\mathbb{R}^+$ be a small constant such that $\varepsilon\in(0,\min_{a\in\mathcal{A}\setminus\{a^\star\}}\Delta_a)$. For $\mu,\gamma\in\mathbb{R}$, let us introduce the following set*

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

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

$$\mathbb{E}\big[N_T(a)\big]\leqslant 2+\inf_{n_0\leqslant T}\Big\{n_0+\sum_{n\geqslant n_0+1}^{T}\mathbb{P}\Big\{\widehat{\nu}_{a,n}\in\mathcal{C}_{\mu^\star-\varepsilon,f(T)/n}\Big\}\Big\}$$
$$+\sum_{t=|\mathcal{A}|}^{T-1}\underbrace{\mathbb{P}\Big\{N_{a^\star}(t)\ \mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\widehat{\nu}_{a^\star,N_{a^\star}(t)}),\ \mu^\star-\varepsilon\big)>f(t)\Big\}}_{\textit{Boundary Crossing Probability}}.$$

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

$$\mathbb{E}\big[N_T(a)\big]\leqslant 2+\inf_{n_0\leqslant T}\Big\{n_0+\sum_{n\geqslant n_0+1}^{T}\mathbb{P}\Big\{\widehat{\nu}_{a,n}\in\mathcal{C}_{\mu^\star-\varepsilon,f(T/n)/n}\Big\}\Big\}$$
$$+\sum_{t=|\mathcal{A}|}^{T-1}\underbrace{\mathbb{P}\Big\{N_{a^\star}(t)\ \mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\widehat{\nu}_{a^\star,N_{a^\star}(t)}),\ \mu^\star-\varepsilon\big)>f(t/N_{a^\star}(t))\Big\}}_{\textit{Boundary Crossing Probability}}.$$

2. the proof is deferred to the appendix; in this lemma, the optimal arm is assumed to be unique.

Lemma 6 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 Cappé et al. (2013) or by resorting to finite-time Sanov-type results such as that of (Dinwoodie, 1992, Theorem 2.1 and comments on page 372), or its variant from (Maillard et al., 2011, Lemma 1). For completeness, the exact result from Dinwoodie (1992) writes

Lemma 7 (Non-asymptotic Sanov's lemma) *Let $\mathcal{C}$ be an open convex subset of $\mathcal{P}(\mathcal{X})$ s.t. $\Lambda(\mathcal{C}) = \inf_{\kappa\in\mathcal{C}} \mathit{KL}(\kappa,\nu)$ is finite. Then, $\forall t \geqslant 1$, $\mathbb{P}_\nu\{\widehat{\nu}_t \in \mathcal{C}\} \leqslant \exp\big(-t\Lambda(\overline{\mathcal{C}})\big)$ where $\overline{\mathcal{C}}$ is the closure of $\mathcal{C}$.*

Scope and focus of this work We present the setting of stochastic multi-armed bandits because it gives a strong and natural motivation for studying boundary crossing probabilities. However, one should understand that one goal of this paper is to give credit to the work of T.L. Lai regarding the neat understanding of boundary crossing probabilities rather than providing 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, see Lerche (2013). Thus, 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 6 is indeed mostly technical and does not seem to require especially illuminating or fancy ideas. We refer to Cappé et al. (2013) for an example of bound in the case of exponential families of dimension 1. Last, since the boundary crossing probability involves properties of a single distribution (that of an optimal arm) and is not specific to bandit problems, we drop $a^\star$ from all notations, thus defining $\mathcal{K}$, Π, $N(t)$, $\widehat{\nu}_{N(t)}$.

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 the observations come from an exponential family of dimension K that satisfies some mild conditions, then for any non-negative ε and some class-dependent but fully explicit constants c, C (depending on ε) it holds that*

$$\mathbb{P}\Big\{N(t)\ \mathcal{K}\big(\Pi(\widehat{\nu}_{N(t)}),\, \mu^\star - \varepsilon\big) > f(t)\Big\} \quad\leqslant\quad \frac{C}{t}\log(t)^{K/2-\xi} e^{-c\sqrt{f(t)}}$$

$$\mathbb{P}\Big\{N(t)\ \mathcal{K}\big(\Pi(\widehat{\nu}_{N(t)}),\, \mu^\star - \varepsilon\big) > f(t/N(t))\Big\} \quad\leqslant\quad \frac{C}{t}\log(tc)^{K/2-\xi-1}\,.$$

The first inequality holds for all t and the second one for large enough $t \geqslant t_c$ for a class dependent but explicit and "reasonably" small t_c. (We dropped $a^\star$ in the notations)

We provide the rigorous statement in Theorem 19 and Corollaries 22, 23 below. This result 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 fact was observed in numerical experiments in Cappé et al. (2013) although not theoretically supported until now.

The remaining of the paper is organized as follows: Section 3 provides the background about exponential families, Section 4 provides the precise statements of the main results, Section B details the proof of Theorem 19 and Section C those of Corollaries 22, 23.

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 $F : \mathcal{X} \to \mathbb{R}^K$ and denote $\mathcal{Y} = F(\mathcal{X}) \subset \mathbb{R}^K$.

Definition 8 (Exponential families) *The exponential family generated by the function F and the reference measure ν_0 on the set $\mathcal{X}$ is*

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

where $\psi(\theta) \stackrel{\text{def}}{=} \log \int_{\mathcal{X}} \exp\Big(\langle\theta, F(x)\rangle\Big)\nu_0(dx)$ is the normalization (log-partition) function of the exponential family. 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 is $\Theta_I \stackrel{\text{def}}{=} \Big\{\theta \in \mathbb{R}^K : 0 < \lambda_{\mathtt{MIN}}(\nabla^2\psi(\theta)) \leqslant \lambda_{\mathtt{MAX}}(\nabla^2\psi(\theta)) < \infty\Big\} \subset \Theta_{\mathcal{D}}$, where $\lambda_{\mathtt{MIN}}(M)$ and $\lambda_{\mathtt{MAX}}(M)$ denote the minimum and maximum eigenvalues of a semi-definite positive matrix M.

Remark 9 *When $\mathcal{X}$ is compact, which is a common 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 assume that the family is regular ($\Theta_{\mathcal{D}}$ has non empty interior) and minimal. 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. This 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 $\theta = \log(\mu/(1-\mu))$. Further, $\Theta_{\mathcal{D}} = \mathbb{R}$ and 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 $\theta = (\frac{\mu}{\sigma^2}, -\frac{1}{2\sigma^2})$. It is immediate to check that $\Theta_{\mathcal{D}} = \mathbb{R} \times (-\infty, 0)$. Degenerate distributions with variance 0 correspond to a parameter 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. For a study of bandits with Gaussian distributions (but specific tools) see Katehakis and Robbins (1995) or more recently Cowan et al. (2015).

3.1. Bregman divergence induced by the exponential family

An interesting property of exponential families is the following straightforward identity:

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

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 $\nabla\psi(\theta)$. Now, we write $\mathtt{KL}(\nu_\theta, \nu_{\theta'}) = \mathcal{B}^\psi(\theta, \theta')$, where we introduced the Bregman divergence with potential function ψ defined 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 simple but powerful properties of the Bregman divergence.

Lemma 10 (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 $\Phi(\eta) = \psi(\theta^\star+\eta) - \psi(\theta^\star)$. Let the Fenchel-Legendre dual of Φ be*

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

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

Lemma 11 (Bregman and Smoothness) *The following inequalities hold true*

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

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

3.2. Dual formulation of the optimization problem

Using Bregman divergences enables 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

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

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

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

For minimal exponential family, this system admits for each θ, μ a unique solution in θ'.

Remark 12 *For $\theta \in \Theta_I$, when the optimal value of λ is $\lambda^\star = 0$, then it means that $\nabla\psi(\theta') = \nabla\psi(\theta)$ and thus $\theta' = \theta$, which is only possible if $\mathbb{E}_{\nu_\theta}(X) \geqslant \mu$. 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 previous optimization problem reduces to a simpler one-dimensional optimization problem, where we optimize over the dual parameter λ. We illustrate this phenomenon on a family of discrete distributions. Let $\mathbb{X} = \{x_1, \dots, x_K, x_\star\}$ be a set of distinct real-values. Without loss of generality, assume that $x_\star > \max_{k \leqslant K} x_k$. The family of distributions p with support in $\mathbb{X}$ is a specific K-dimensional family. Indeed, let F be the feature function with k^{th} component $F_k(x) = \mathbb{I}\{x = x_k\}$, for all $k \in \{1, \dots, K\}$. Then the parameter $\theta = (\theta_k)_{1 \leqslant k \leqslant K}$ of the distribution $p = p_\theta$ has components $\theta_k = \log(\frac{p(x_k)}{p(x_\star)})$ for all $k \neq 0$. Note that $p(x_k) = \exp(\theta_k - \psi(\theta))$ for all $k \neq 0$, and $p(x_0) = \exp(-\psi(\theta))$. It then comes $\psi(\theta) = \log(\sum_{k=1}^K e^{\theta_k} + 1)$, $\nabla\psi(\theta) = (p(x_1), \dots, p(x_K))^\top$ 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 non trivial 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 Honda and Takemura (2010) for discrete distributions. After some tedious but simple steps detailed in Honda and Takemura (2010), one obtains the following easy-to-solve *one-dimensional* optimization problem (see also Cappé et al. (2013)), although the family is of dimension K:

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

3.3. Empirical parameter and definition

We now discuss the well-definedness 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 a specific care to ensure that all the objects are well-defined and 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 the closure $\overline{\nabla\psi(\Theta_\mathcal{D})}$ for all $\{X_i\}_{i \leqslant n} \sim \nu_\theta$ with $\theta \in \Theta_\mathcal{D}$. Thus, for the observed samples $X_1, \dots, X_n \in \mathcal{X}$ coming from $\nu_{\theta^\star}$, the projection $\Pi(\widehat{\nu}_n)$ on the family can be represented by a sequence $\{\widehat{\theta}_{n,m}\}_{m \in \mathbb{N}} \in \Theta_\mathcal{D}$ such that

$$\nabla\psi(\widehat{\theta}_{n,m}) \overset{m}{\to} \widehat{F}_n \;\text{ where }\; \widehat{F}_n \overset{\text{def}}{=} \frac{1}{n}\sum_{i=1}^n F(X_i) \in \mathbb{R}^K .$$

In the sequel, we want to ensure that provided that $\nu_{a^\star} = \nu_{\theta^\star}$ with $\theta^\star \in \mathring{\Theta}_I$ (the interior of Θ_I), then $\widehat{F}_n \in \nabla\psi(\mathring{\Theta}_I)$ holds, which means there is a unique $\widehat{\theta}_n \in \mathring{\Theta}_I$ such that $\nabla\psi(\widehat{\theta}_n) = \widehat{F}_n$, or equivalently $\widehat{\theta}_n = \nabla\psi^{-1}(\widehat{F}_n)$. 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 we can find a unique $\widehat{\theta}_n \in \mathring{\Theta}_I$ such that $\nabla\psi(\widehat{\theta}_n) = \widehat{F}_n$.

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 $n \geqslant 2$, the empirical mean $\widehat{\mu}_n$ is finite and the empirical variance $\widehat{\sigma}_n^2$ is positive, and thus $\theta_n = \nabla\psi^{-1}(\widehat{F}_n)$ 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 a Bernoulli distribution with mean μ. Before $\widehat{F}_n$ 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 $\mathbb{P}(n_0(n) = 0 \text{ or } n_1(n) = 0) = \mu^n + (1-\mu)^n \leqslant 2\max(\mu, 1-\mu)^n$, where $n_x(n)$ denotes the number of observations of symbol $x \in \{0,1\}$ after n samples. For $\mu \geqslant 1/2$, the later quantity is less than $\delta_0 \in (0,1)$ for $n \geqslant \frac{\log(2/\delta_0)}{\log(1/\mu)}$, which depends on the probability level δ_0 and cannot be considered to be especially small when μ is close [3] to 1. That said, even when the parameter $\widehat{\theta}_n$ 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 $\mathbb{P}(n_1(n) = 0) = (1-\mu)^n$. This is less than $\delta_0 \in (0,1)$ for $n \geqslant \frac{\log(1/\delta_0)}{\log(1/(1-\mu))}$, which is essentially a constant. For illustration, when $\delta = 10^{-3}$ and $\mu = 0.9$, this happens for $n \geqslant 3$.

Following the previous discussion, in the sequel we consider that n is always large enough so that $\widehat{\theta}_n = \nabla\psi^{-1}(\widehat{F}_n) \in \mathring{\Theta}_I$ is uniquely defined. We now make formal the separation between the parameter and the boundary. For that purpose introduce the following definition:

Definition 13 (Enlarged parameter set) *Let $\Theta \subset \Theta_{\mathcal{D}}$ and some constant $\rho > 0$. The enlargement of size ρ of Θ in Euclidean norm (aka ρ-neighborhood) is defined by*

$$\Theta_\rho \stackrel{\text{def}}{=} \left\{\theta \in \mathbb{R}^K \,;\, \inf_{\theta' \in \Theta} |\theta - \theta'| < \rho\right\}.$$

For each ρ such that $\Theta_\rho \subset \Theta_I$, we further introduce the quantities

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

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

Lemma 14 (Log-Laplace control) *Let $\Theta \subset \Theta_{\mathcal{D}}$ be a convex set and $\rho > 0$ 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*

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

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 $B(y, \delta) = \{y' \in \mathbb{R}^K : \|y' - y\| \leqslant \delta\}$. Using these notations, the following lemma whose proof is immediate provides conditions for which all future technical considerations are satisfied.

Lemma 15 (Well-defined parameters) *Let $\theta^\star \in \mathring{\Theta}_I$ and $\rho^\star = d(\{\theta^\star\}, \mathbb{R}^K \setminus \Theta_I) > 0$. Then 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 $\quad \Theta_{2\rho} \subset \mathring{\Theta}_I$. Further, for any δ such that $\widehat{F}_n \in B(\nabla\psi(\theta^\star), \delta) \subset \nabla\psi(\Theta_\rho)$, then*

$$\exists \widehat{\theta}_n \in \Theta_\rho \subset \mathring{\Theta}_I \text{ such that } \nabla\psi(\widehat{\theta}_n) = \widehat{F}_n.$$

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

Remark 16 *We recall that when $\mathcal{X}$ is compact and F is continuous, then $\Theta_I = \Theta_{\mathcal{D}} = \mathbb{R}^K$.*

3. This also suggests to replace $\widehat{F}_n$ with a Laplace or a Krichevsky-Trofimov estimate that provide initial bonus to each symbol and, as a result, maps any $\widehat{F}_n$, for $n \geqslant 0$ to a parameter in $\widehat{\theta}_n \in \mathbb{R}$.

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 $V_\rho \leqslant 1/4$. Now for a family of Gaussian distributions with unknown mean and variance, $\psi(\theta) = -\frac{\theta_1^2}{4\theta_2} + \frac{1}{2}\log\left(\frac{-\pi}{\theta_2}\right)$, where $\theta = (\frac{\mu}{\sigma^2}, -\frac{1}{2\sigma^2})$. Thus, $\nabla\psi(\theta) = (-\frac{\theta_1}{2\theta_2}, \frac{\theta_1^2}{4\theta_2^2} - \frac{1}{2\theta_2})$, and $\nabla^2\psi(\theta) = (-\frac{1}{2\theta_2}, \frac{\theta_1}{2\theta_2^2}; \frac{\theta_1}{2\theta_2^2}, -\frac{\theta_1^2}{2\theta_2^3} + \frac{1}{2\theta_2^2}) = 2\mu\sigma^2(\frac{1}{2\mu}, 1; 1, 2\mu + \frac{\sigma^2}{\mu})$. 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 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 6 for a K-dimensional exponential family $\mathcal{E}(F;\nu_0)$. We first provide an overview of the existing results before detailing our main contribution. As explained in Section 2, the key technical tools that enable to obtain the novel results were already known three decades ago. 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 to solve as a by-product some long-lasting open questions.

4.1. Previous work on boundary-crossing probabilities

Previous results from the bandit literature about boundary-crossing probabilities are restricted to a few specific cases. For instance in Cappé et al. (2013), the authors provide the following control for a distribution $\nu_{\theta^\star}$ with mean $\mu^\star$.

Theorem 17 (Cappe et al, 2013) *For the canonical ($F(x) = x$) exponential family of dimension $K = 1$, and the threshold function $f(x) = \log(x) + \xi\log\log(x)$, then for all $t > 1$*

$$\mathbb{P}_{\theta^\star}\Big\{\bigcup_{n=1}^{t} n\mathcal{K}\big(\Pi(\widehat{\nu}_n),\, \mu^\star\big) > f\big(t\big) \cap \mu^\star > \widehat{\mu}_n\Big\} \leqslant e\lceil f(t)\log(t)\rceil e^{-f(t)}\,.$$

Further, for discrete distributions with S many atoms and $\xi \geqslant 3$ it holds for all $t > 1, \varepsilon > 0$

$$\mathbb{P}_{\theta^\star}\Big\{\bigcup_{n=1}^{t} n\mathcal{K}\big(\Pi(\widehat{\nu}_n),\, \mu^\star - \varepsilon\big) > f\big(t\big)\Big\} \leqslant e^{-f(t)}\Big(3e + 2 + 4\varepsilon^{-2} + 8e\varepsilon^{-4}\Big)\,.$$

In contrast in Lai (1988), the authors 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 Lai (1988) in a way that is suitable for a more direct comparison with other results. The following holds:

Theorem 18 (Lai, 88) *Let us consider an exponential family of dimension K. Define for $\gamma > 0$ the cone $\mathcal{C}_\gamma(\theta) = \{\theta' \in \mathbb{R}^K : \langle\theta', \theta\rangle \geqslant \gamma\|\theta\|\|\theta'\|\}$. Then, for a function f such that $f(x) = \alpha\log(x) + \xi\log\log(x)$ it holds for all $\theta^\dagger \in \Theta$ such that $|\theta^\dagger - \theta^\star|^2 \geqslant \delta_t$, where $\delta_t \to 0$, $t\delta_t \to \infty$ as $t \to \infty$,*

$$\begin{aligned}\mathbb{P}_{\theta^\star}\Big\{\bigcup_{n=1}^{t} \widehat{\theta}_n \in \Theta_\rho \cap n\mathcal{B}^\psi(\widehat{\theta}_n, \theta^\dagger) \geqslant f\Big(\frac{t}{n}\Big) \cap \nabla\psi(\widehat{\theta}_n) - \nabla\psi(\theta^\dagger) \in \mathcal{C}_\gamma(\theta^\dagger - \theta^\star)\Big\} \\ \overset{t\to\infty}{=} O\Big(t^{-\alpha}|\theta^\dagger - \theta^\star|^{-2\alpha}\log^{-\xi-\alpha+K/2}(t|\theta^\dagger - \theta^\star|^2)\Big) \\ = O\Big(e^{-f(t|\theta^\dagger-\theta^\star|^2)}\log^{-\alpha+K/2}(t|\theta^\dagger - \theta^\star|^2)\Big)\,.\end{aligned}$$

Discussion $\mathcal{B}^{\psi}(\widehat{\theta}_n, \theta^\dagger)$ is the direct analog of $\mathcal{K}\big(\Pi(\widehat{\nu}_n), \mu^\star - \varepsilon\big)$ in Theorem 17. However, $f(t/n)$ replaces the larger quantity $f(t)$, which means Theorem 18 controls a larger quantity than Theorem 17, and is thus stronger in this sense. Also, it applies to general exponential families of dimension K. Finally, the right hand side terms of both theorems have different scaling: Since $e\lceil f(t)\log(t)\rceil e^{-f(t)} = O(\frac{\log^{2-\xi}(t)+\xi\log(t)^{1-\xi}\log\log(t)}{t})$, the first part of Theorem 17 requires $\xi > 2$ so that this term is $o(1/t)$, and $\xi > 0$ for the second part. In contrast, Theorem 18 shows that using $f(x) = \log(x) + \xi\log\log(x)$ with $\xi > K/2 - 1$ is enough 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 Cappé et al. (2013) choose in the experiments.

Thus, Theorem 18 improves in three ways over Theorem 17: 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 17 controls a concentration with respect to $\theta^\dagger$, not $\theta^\star$, which takes advantage of the gap that appears when going from $\mu^\star$ to distributions with mean $\mu^\star - \varepsilon$. The proof of Theorem 18 directly takes advantage of this, contrary to that of the first part of Theorem 17.

However, Theorem 18 is asymptotic whereas Theorem 17 holds for finite t. Furthermore, we notice two restrictions on the controlled event. First, $\widehat{\theta}_n \in \Theta_\rho$; we showed in the previous section that this is a minor restriction. Second, the restriction to $\mathcal{C}_\gamma(\theta^\dagger - \theta^\star)$ which simplifies the analysis, is a more dramatic restriction as it cannot be removed trivially. Indeed from the complete statement of (Lai, 1988, Theorem 2), 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 19 in section B.

4.2. Main results and contributions

We provide several results on boundary crossing probabilities, that we prove in details in the next section. We first provide a non-asymptotic 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 19 (Boundary crossing for exponential families) *Let $\varepsilon > 0$ and define $\rho_\varepsilon = \inf\{\|\theta' - \theta\| : \mu_{\theta'} = \mu^\star - \varepsilon, \mu_\theta = \mu^\star\}$. 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 non-increasing and $n \to nf(t/n)$ is non-decreasing. Then, $\forall b > 1, p, q, \eta \in [0,1]$, and $n_i = b^i$ if $i < I_t = \lceil \log_b(qt)\rceil$, $n_{I_t} = t+1$,*

$$\mathbb{P}_{\theta^\star}\Big\{\bigcup_{1\leqslant n\leqslant t} \widehat{\theta}_n \in \Theta_\rho \cap \mathcal{K}(\Pi(\widehat{\nu}_n), \mu^\star - \varepsilon) \geqslant f(t/n)/n\Big\}$$
$$\leqslant C(K,b,\rho,p,\eta)\sum_{i=0}^{I_t-1}\exp\Big(-n_i\rho_\varepsilon^2\alpha^2 - \rho_\varepsilon\chi\sqrt{n_i f(t/n_i)} - f\Big(\frac{t}{n_{i+1}-1}\Big)\Big) f\Big(\frac{t}{n_{i+1}-1}\Big)^{K/2},$$

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

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

Here $C_{p,\eta,K}$ is the cone-covering number of $\nabla\psi\big(\Theta_\rho \setminus \mathcal{B}_2(\theta^\star, \rho_\varepsilon)\big)$ with minimal angular separation p, excluding $\nabla\psi\big(\Theta_\rho \setminus \mathcal{B}_2(\theta^\star, \eta\rho_\varepsilon)\big)$; and $\omega_{p,K} = \int_p^1 \sqrt{1-z^2}^K dz$ if $K \geqslant 0$, 1 else.

Remark 20 *The same result holds by replacing all occurrences of $f(\cdot)$ by the constant $f(t)$.*

Remark 21 *In dimension 1, the result takes a simpler form since $C_{p,\eta,1}=2$ for all $p,\eta\in(0,1)$. 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 $C(1,2,\rho,p,\eta)=2(\frac{1}{8\mu_\rho^3(1-\mu_\rho)^3}+1)$.*

Let $f(x)=\log(x)+\xi\log\log(x)$. We now state two corollaries of Theorem 19, The first one is stated for the case when the boundary is set to $f(t)/n$ 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 non-decreasing only for $x\geqslant e^{-\xi}$. When $x=t$, this requires that $t\geqslant e^{-\xi}$. Now, when $x=t/N(t)$ where $N(t)=t-O(\log(t))$, imposing that f is non-decreasing requires that $\xi\geqslant\log(1-O(\log(t)/t))$ for large t, that is $\xi\geqslant 0$. In the sequel we thus restrict to $t\geqslant e^{-\xi}$ when using the boundary $f(t)$ and to $\xi\geqslant 0$ when using the boundary $f(t/n)$. Finally, we recall that the quantity $\chi=p\eta\sqrt{2v_\rho^2/V_\rho}$ is a function of p,η,ρ, and introduce the notation $\chi_\varepsilon=\rho_\varepsilon\chi$ for convenience.

Corollary 22 (Boundary crossing for $f(t)$) *Let $f(x)=\log(x)+\xi\log\log(x)$. Using the same notations as in Theorem 19, for all $p,\eta\in[0,1],\rho<\rho^\star$ and all $t\geqslant e^{-\xi}$ such that $f(t)\geqslant 1$ it holds*

$$\mathbb{P}_{\theta^\star}\Big\{\bigcup_{1\leqslant n<t}\widehat{\theta}_n\in\Theta_\rho\cap\mathcal{K}(\Pi(\widehat{\nu}_n),\mu^\star-\varepsilon)\geqslant f(t)/n\Big\}$$
$$\leqslant\frac{C(K,4,\rho,p,\eta)(1+\chi_\varepsilon)}{\chi_\varepsilon t}\left(1+\xi\frac{\log\log(t)}{\log(t)}\right)^{K/2}\log(t)^{-\xi+K/2}e^{-\chi_\varepsilon\sqrt{\log(t)+\xi\log\log(t)}}.$$

Corollary 23 (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 $\xi\geqslant\max(K/2-1,0)$, provided that $t\in[85\chi^{-2},t_\chi]$ where $t_\chi=\chi_\varepsilon^{-2}\frac{\exp(\log(4.5)^2/\chi_\varepsilon^2)}{4\log(4.5)^2}$, it holds

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

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

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

Remark 24 *In Corollary 22, since the asymptotic regime of $\chi_\varepsilon\sqrt{\log(t)}-(K/2-\xi)\log\log(t)$ may take a massive amount of time to kick-in when $\xi<K/2-2\chi_\varepsilon$, we recommend to take $\xi>K/2-2\chi_\varepsilon$. Now, we also note that the value $\xi=K/2-1/2$ is interesting in practice, since then $\log(t)^{K/2-\xi}=\sqrt{\log(t)}<5$ holds for all $t\leqslant 10^9$.*

Remark 25 *The restriction to $t \geqslant 85\chi_\varepsilon^{-2}$ is merely for $\xi \simeq K/2-1$. For instance for $\xi \geqslant K/2$, the restriction becomes $t \geqslant 76\chi_\varepsilon^{-2}$, and it becomes less restrictive for larger ξ. The term t_χ is virtually infinite: when $\chi_\varepsilon = 0.3$, this is already larger than 10^{12}, while $85\chi_\varepsilon^{-2} < 945$.*

Remark 26 *According to this result, the value $K/2-1$ (when it is non-negative) is a critical value for ξ, since the boundary crossing probabilities are not summable in t for $\xi \leqslant K/2-1$, 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 $\xi = K/2-1$. Now in practice, since this asymptotic behavior may take a large time to kick-in, we recommend ξ to be away from $K/2-1$.*

Remark 27 *Controlling probabilities for the threshold $f(t/N(t))$ is more challenging than for $f(t)$. Only the later case was analyzed in Cappé et al. (2013) 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. Interestingly, when $K = 1$, $\max(K/2-1, 0) = 0$, and we indeed observe experimentally a sharp phase transition 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 $\xi = 0$. Instead, a smooth phase transition appears for a negative ξ depending on the problem. Both observations are coherent with the statements of the corollaries, which is remarkable.*

Discussion regarding the proof technique The proof technique significantly differs from the proof from Cappé et al. (2013) and Honda and Takemura (2010), and combines key ideas disseminated in Lai (1988) and Lai (1987) with some non-trivial 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. More precisely:

-**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 of Lai (1988) should be given full credit for this idea. This is in stark contrast with the proof techniques later developed 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 $\theta_c^\star$. Then, which is perhaps more surprising, to to go from this perturbed point to a mixture over a well-chosen ball centered on 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 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 to use Bregman divergence and its relation with the quadratic norm, which is due to Lai (1987). This enables indeed to make explicit computations for exponential families of dimension K without too much effort, at the price of loosing some "variance" terms (linked to the Hessian of the family). We combine this idea with some key properties of Bregman divergence that enable 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 of the proof is 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 challenging due to the boundary effects, although

the concentration result (section B.4, Lemma 33) 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 Lai (1988), the author analyzed a boundary crossing problem first in the case of exponential families of dimension 1, and then sketch the analysis for exponential families of dimension K and for the intersection with one cone. However the complete result was nowhere stated explicitly. Actually, the initial proof from Lai (1988) 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 to get a final result that only depends on the cone-covering number of the space. It required careful considerations and simplifications of the initial steps from Lai (1988). Along the way, we made explicit the sketch of proof provided in Lai (1988) 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 change of measures, the application of concentration and the peeling argument. Indeed these steps lead to a sum of terms, say $\sum_{i=0}^{I} s_i$ that should be studied and depends on a few free parameters. This correspond, with our rewriting and modifications, to the statement of Theorem 19. Here the brilliant idea of T.L. Lai, that we separate from the proof of Theorem 19 and use in the proof of Corollaries 22 and 23 is to bound the ratios of s_{i+1}/s_i for small values of i and the ratio 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 enable to improve the initial scaling and allow for smaller values of ξ, up to $K/2-1$, while alternative 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 requires huge values for t before the asymptotic regimes kick-in. By refining the initial proof strategy of Lai (1988), we obtain a result valid for all t for the setting of Corollary 22 and for all "reasonably"[4] large t in the more challenging setting of Corollary 23.

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 Lai (1988), that we modernized into a fully non-asymptotic statement, with explicit constants; it can be directly used, for instance, for the regret analysis of multi-armed bandits 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, 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 (Honda and Takemura (2010); Maillard et al. (2011); Cappé et al. (2013)). Also, we believe and hope that the novel writing of this proof technique will greatly benefit the community working on boundary crossing probabilities, sequential design of experiments as well as stochastic decision making strategies.

4. 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.

Acknowledgements This work has been supported by the French Agence Nationale de la Recherche (ANR), under grant ANR-16- CE40-0002 (project BADASS), and Inria.

References

Rajeev Agrawal. Sample mean based index policies by $o(\log n)$ regret for the multi-armed bandit problem. *Advances in Applied Probability*, 27(04):1054–1078, 1995.

J-Y. Audibert, R. Munos, and Cs. Szepesvári. Exploration-exploitation trade-off using variance estimates in multi-armed bandits. *Theoretical Computer Science*, 410(19), 2009.

J.Y. Audibert and S. Bubeck. Regret bounds and minimax policies under partial monitoring. *Journal of Machine Learning Research*, 11:2635–2686, 2010.

P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, 47(2):235–256, 2002.

Sébastien Bubeck, Nicolo 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.

A.N. Burnetas and M.N. Katehakis. Optimal adaptive policies for Markov decision processes. *Mathematics of Operations Research*, pages 222–255, 1997.

Olivier Cappé, Aurélien Garivier, Odalric-Ambrym Maillard, Rémi Munos, and Gilles Stoltz. Kullback–Leibler upper confidence bounds for optimal sequential allocation. *Annals of Statistics*, 41(3):1516–1541, 2013.

Wesley Cowan, Junya Honda, and Michael N Katehakis. Normal bandits of unknown means and variances: Asymptotic optimality, finite horizon regret bounds, and a solution to an open problem. *arXiv preprint arXiv:1504.05823*, 2015.

Ian H Dinwoodie. Mesures dominantes et théoreme de sanov. *Annales de l'IHP Probabilités et statistiques*, 28(3):365–373, 1992.

Aurélien Garivier, Pierre Ménard, and Gilles Stoltz. Explore first, exploit next: The true shape of regret in bandit problems. *arXiv preprint arXiv:1602.07182*, 2016.

J.C. Gittins. Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society, Series B*, 41(2):148–177, 1979.

J. Honda and A. Takemura. An asymptotically optimal bandit algorithm for bounded support models. In T. Kalai and M. Mohri, editors, *Conf. Comput. Learning Theory*, Haifa, Israel, 2010.

Michael N Katehakis and Herbert Robbins. Sequential choice from several populations. *Proceedings of the National Academy of Sciences of the United States of America*, 92(19): 8584, 1995.

T.L. Lai and H. Robbins. Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics*, 6(1):4–22, 1985.

Tze Leung Lai. Adaptive treatment allocation and the multi-armed bandit problem. *The Annals of Statistics*, pages 1091–1114, 1987.

Tze Leung Lai. Boundary crossing problems for sample means. *The Annals of Probability*, pages 375–396, 1988.

Tor Lattimore. Regret analysis of the finite-horizon gittins index strategy for multi-armed bandits. In *Conference on Learning Theory*, pages 1214–1245, 2016.

Hans R Lerche. *Boundary crossing of Brownian motion: Its relation to the law of the iterated logarithm and to sequential analysis*, volume 40. Springer Science & Business Media, 2013.

O-A. Maillard, R. Munos, and G. Stoltz. A finite-time analysis of multi-armed bandits problems with kullback-leibler divergences. In *Conf. Comput. Learning Theory*, Budapest, Hungary, 2011.

Herbert Robbins. Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, 58(5):527–535, 1952.

Herbert Robbins. *Herbert Robbins Selected Papers*. Springer, 2012.

William 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.

William R Thompson. On a criterion for the rejection of observations and the distribution of the ratio of deviation to sample standard deviation. *The Annals of Mathematical Statistics*, 6(4):214–219, 1935.

Abraham Wald. Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics*, 16(2):117–186, 1945.

Appendix A. Notations

$K \in \mathbb{N}$	Dimension of the exponential family
$\Theta \subset \mathbb{R}^K$	Parameter set, see Theorem 19
$\Theta_\rho \subset \mathbb{R}^K$	Enlarged parameter set, see Definition 13
ψ	Log-partition function of the exponential family
$\mathcal{B}^\psi$	Bregman divergence of the exponential family
$\mathcal{V}_\rho, v_\rho$	Largest and smallest values of the eigenvalues of the Hessian on Θ_ρ, see Definition 13
$\theta^\star$	Parameter of the distribution generating the observed samples
$\widehat{\theta}_n$	Empirical parameter built from n observations
$\widehat{F}_n \in \mathbb{R}^K$	Empirical mean of the $F(X_i)$, $i \leqslant n$, see Section 3.3
f	Threshold function
$\mu^\star \in \mathbb{R}$	Mean of the distribution with parameter $\theta^\star$
$\varepsilon > 0$	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.

Appendix B. Analysis of boundary crossing probabilities: proof of Theorem 19

In this section[5], we closely follow the proof technique used in Lai (1988) for the proof of Theorem 18, in order to prove the result of Theorem 19. We precise further the constants, remove the cone restriction on the parameter and modify the original proof to be fully non-asymptotic which, using the technique of Lai (1988), 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 B.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 (affine) cones, which enables to later use arguments from proof of Theorem 18.

In Section B.2, we proceed with the first change of measure argument: taking advantage of the gap between $\mu^\star$ and $\mu^\star - \varepsilon$, we move from a concentration argument around $\theta^\star$ to one around a shifted point $\theta^\star - \Delta_c$.

In Section B.3, we localize the empirical parameter $\widehat{\theta}_n$ and make use of the second change of measure, this time to a mixture of measures, following Lai (1988). Even though we follow the same high level idea, we modified the original proof in order to better handle the cone covering, and also make all quantities explicit.

In Section B.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, and was not explicitly controlled in the original proof from Lai (1988). A very careful analysis enables to obtain the finite-time concentration result without further restriction.

We finally combine all these steps in Sections B.5.

5. The content of this supplementary material is made available on ArXiv.

B.1. Peeling and covering

In this section, we want to control the random number of pulls $N(t) \in [1,t]$ and to this end we 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 $\mathcal{K}(\Pi(\widehat{\nu}_{,n}), \mu^\star - \varepsilon)$ can be seen as an infimum of some quantity over the set of parameters 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 28 (Peeling and cone covering decomposition) *For all $\beta, \eta \in (0,1), b > 1$*

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{1\leqslant n\leqslant t} \widehat{\theta}_n \in \Theta_\rho \cap \mathcal{K}(\Pi(\widehat{\nu}_n), \mu^\star - \varepsilon) \geqslant f(t/n)/n \Big\}$$
$$\leqslant \sum_{i=0}^{\lceil \log_b(\beta t+\beta)\rceil - 2} \sum_{c=1}^{C_{p,\eta,K}} \mathbb{P}_{\theta^\star}\Big\{ \bigcup_{n=b^i}^{b^{i+1}-1} E_{c,p}(n,t) \Big\} + \sum_{c=1}^{C_{p,\eta,K}} \mathbb{P}_{\theta^\star}\Big\{ \bigcup_{n=b^{\lceil \log_b(\beta t+\beta)\rceil -1}}^{t} E_{c,p}(n,t) \Big\},$$

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

$$E_{c,p}(n,t) \stackrel{\text{def}}{=} \Big\{ \widehat{\theta}_n \in \Theta_\rho \cap \widehat{F}_n \in \mathcal{C}_p(\theta_c^\star) \cap \mathcal{B}^\psi(\widehat{\theta}_n, \theta_c^\star) \geqslant \frac{f(t/n)}{n} \Big\}. \tag{4}$$

The points $(\theta_c^\star)_{c\leqslant C_{p,\eta,K}}$ such that $\theta_c^\star \notin \mathcal{B}_2(\theta^\star, \eta\rho_\varepsilon)$, parameterize a minimal covering of $\nabla\psi(\Theta_\rho \backslash \mathcal{B}_2(\theta^\star, \rho_\varepsilon))$ with cones $\mathcal{C}_p(\theta_c^\star) := \mathcal{C}_p(\nabla\psi(\theta_c^\star); \theta^\star - \theta_c^\star)$ (That is $\nabla\psi(\Theta_\rho \backslash \mathcal{B}_2(\theta^\star, \rho_\varepsilon)) \subset \bigcup_{c=1}^{C_{p,\eta,K}} \mathcal{C}_p(\theta_c^\star)$), where $\mathcal{C}_p(y;\Delta) = \Big\{ y' \in \mathbb{R}^K : \langle y' - y, \Delta\rangle \geqslant p\|y'-y\|\|\Delta\| \Big\}$. For all $\eta < 1$, $C_{p,\eta,K}$ is of order $(1-p)^{-K}$; $C_{p,\eta,1} = 2$ and $C_{p,\eta,K} \to \infty$ when $\eta \to 1$.

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

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

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 $n_{I_t - 1} \leqslant b^{\log_b(\beta t+\beta)} = \beta(t+1) < t+1 = n_{I_t}$), and to the event

$$E_n \stackrel{\text{def}}{=} \Big\{ \widehat{\theta}_n \in \Theta_\rho \cap \mathcal{K}(\Pi(\widehat{\nu}_n), \mu^\star - \varepsilon) \geqslant f(t/n)/n \Big\}.$$

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

$$\mathcal{K}(\Pi(\widehat{\nu}_n), \mu^\star - \varepsilon) = \inf\Big\{ \mathcal{B}^\psi(\widehat{\theta}_n, \theta^\star - \Delta) : \theta^\star - \Delta \in \Theta_{\mathcal{D}}, \mu_{\theta^\star - \Delta} \geqslant \mu^\star - \varepsilon \Big\},$$

where $\widehat{\theta}_n \in \Theta_{\mathcal{D}}$ is any point such that $\widehat{F}_n = \nabla\psi(\widehat{\theta}_n)$. 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 $\mathcal{B}^\psi(\widehat{\theta}_n, \theta^\star)$ and $\mathcal{B}^\psi(\widehat{\theta}_n, \theta^\star - \Delta)$. Immediate computations show that for any Δ such that $\theta^\star - \Delta \in \Theta_{\mathcal{D}}$ it holds

$$\begin{aligned}
\mathcal{B}^\psi(\widehat{\theta}_n, \theta^\star - \Delta) &= \psi(\theta^\star - \Delta) - \psi(\widehat{\theta}_n) - \langle \theta^\star - \Delta - \widehat{\theta}_n, \nabla\psi(\widehat{\theta}_n)\rangle \\
&= \psi(\theta^\star) - \psi(\widehat{\theta}_n) - \langle \theta^\star - \widehat{\theta}_n, \nabla\psi(\widehat{\theta}_n)\rangle + \psi(\theta^\star - \Delta) - \psi(\theta^\star) + \langle \Delta, \nabla\psi(\widehat{\theta}_n)\rangle \\
&= \mathcal{B}^\psi(\widehat{\theta}_n, \theta^\star) + \psi(\theta^\star - \Delta) - \psi(\theta^\star) + \langle \Delta, \nabla\psi(\widehat{\theta}_n)\rangle \\
&= \mathcal{B}^\psi(\widehat{\theta}_n, \theta^\star) \underbrace{- \mathcal{B}^\psi(\theta^\star - \Delta, \theta^\star) - \langle \Delta, \nabla\psi(\theta^\star - \Delta) - \widehat{F}_n\rangle}_{\text{shift}} . \qquad (5)
\end{aligned}$$

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 $\widehat{\theta}_n \notin \mathcal{B}_2(\theta^\star, \rho_\varepsilon)$. Indeed, for all $\theta \in \mathcal{B}_2(\theta^\star, \rho_\varepsilon) \cap \Theta_{\mathcal{D}}$, $\mu_\theta \geqslant \mu_\star - \varepsilon$, and thus $\mathcal{K}(\nu_\theta, \mu^\star - \varepsilon) = 0$. It is thus natural to build a covering of $\nabla\psi(\Theta_\rho \setminus \mathcal{B}_2(\theta^\star, \rho_\varepsilon))$. Formally, for a given $p \in [0,1]$ and a base point $y \in \mathcal{Y}$, let us introduce the cone

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

We then associate to 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 $(\theta_c^\star)_{c=1,\dots,C_{p,\eta,K}}$ be a set of points corresponding to a minimal covering of the set $\nabla\psi(\Theta_\rho \setminus \mathcal{B}_2(\theta^\star, \rho_\varepsilon))$, in the sense that

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

constrained to be outside the ball $\mathcal{B}_2(\theta^\star, \eta\rho_\varepsilon)$, that is $\theta_c^\star \notin \mathcal{B}_2(\theta^\star, \eta\rho_\varepsilon)$ 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 $\theta_c^\star \in \Theta_\rho \cap \mathcal{B}_2(\theta^\star, \rho_\varepsilon)$. More precisely, when $p < 1$, then $\Delta_c = \theta^\star - \theta_c^\star$ is such that $\rho_\varepsilon - \|\Delta_c\|$ is positive and away from 0. Also, we have by property of $\mathcal{B}_2(\theta^\star, \rho_\varepsilon)$ that $\mu_{\theta_c^\star} \geqslant \mu^\star - \varepsilon$, and by the constraint that $\|\Delta_c\| > \eta\rho_\varepsilon$.

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

$$E_{c,p}(n,t) \stackrel{\text{def}}{=} \left\{ \widehat{\theta}_n \in \Theta_\rho \cap \widehat{F}_n \in \mathcal{C}_p(\theta_c^\star) \cap \mathcal{B}^\psi(\widehat{\theta}_n, \theta_c^\star) \geqslant \frac{f(t/n)}{n} \right\}. \qquad (6)$$

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

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{1 \leqslant n \leqslant t} \widehat{\theta}_n \in \Theta_\rho \cap \mathcal{K}(\Pi(\widehat{\nu}_n), \mu^\star - \varepsilon) \geqslant f(t/n)/n \Big\} \leqslant \sum_{i=0}^{I_t - 1} \sum_{c=1}^{C_{p,\eta,K}} \mathbb{P}_{\theta^\star}\Big\{ \bigcup_{n_i \leqslant n < n_{i+1}} E_{c,p}(n,t) \Big\}.$$

B.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 $\mu^\star - \varepsilon$, that allows to shift from $\theta^\star$ to some of the $\theta_c^\star$ from the cover. The key observation is to control the change of measure from $\theta^\star$ to each $\theta_c^\star$. Note that $\theta_c^\star \in (\Theta_\rho \cap \mathcal{B}_2(\theta_c^\star, \rho_\varepsilon)) \setminus \mathcal{B}_2(\theta_c^\star, \eta\rho_\varepsilon)$ and that $\mu_{\theta_c^\star} \geqslant \mu^\star - \varepsilon$. We show that

Lemma 29 (Change of measure) *If $n \to nf(t/n)$ is non-decreasing, then for any increasing sequence $\{n_i\}_{i\geqslant 0}$ of non-negative integers it holds*

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

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

Proof : For any event measurable 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 $E = \{\bigcup_{n_i\leqslant n<n_{i+1}} E_{c,p}(n,t)\}$, leads to

$$\begin{aligned}
\int_E \frac{d\mathbb{P}_{\theta^\star}}{d\mathbb{P}_{\theta_c^\star}} d\mathbb{P}_{\theta_c^\star} &= \int_E \frac{\Pi_{k=1}^n \nu_{\theta^\star}(X_k)}{\Pi_{k=1}^n \nu_{\theta_c^\star}(X_k)} d\mathbb{P}_{\theta_c^\star} \\
&= \int_E \exp\Big(n\langle \theta^\star - \theta_c^\star, \widehat{F}_n\rangle - n\Big(\psi(\theta^\star) - \psi(\theta_c^\star)\Big)\Big) d\mathbb{P}_{\theta_c^\star} \\
&= \int_E \exp\Big(-n\langle \Delta_c, \nabla\psi(\theta_c^\star) - \widehat{F}_n\rangle - n\mathcal{B}^\psi(\theta_c^\star, \theta^\star)\Big) d\mathbb{P}_{\theta_c^\star}\,, \qquad (7)
\end{aligned}$$

where $\Delta_c = \theta^\star - \theta_c^\star$. Note that this rewriting makes appear the same term as the shift term appearing in (5). Now, we remark that since $\theta_c^\star \in \Theta_\rho$ by construction, then under the event $E_{c,p}(n,t)$ it holds by convexity of Θ_ρ and elementary Taylor approximation

$$\begin{aligned}
-\langle \Delta_c, \nabla\psi(\theta_c^\star) - \widehat{F}_n\rangle &\leqslant -p\|\Delta_c\|\|\nabla\psi(\theta_c^\star) - \widehat{F}_n\| \\
&\leqslant -p\|\Delta_c\| v_\rho \|\widehat{\theta}_n^\star - \theta_c^\star\| \\
&\leqslant -p\|\Delta_c\|_c v_\rho \sqrt{\frac{2}{V_\rho}\mathcal{B}^\psi(\widehat{\theta}_n, \theta_c^\star)} \\
&\leqslant -p\eta\rho_\varepsilon v_\rho \sqrt{\frac{2f(t/n)}{V_\rho n}}\,. \qquad (8)
\end{aligned}$$

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

$$-\mathcal{B}^\psi(\theta_c^\star, \theta^\star) \leqslant -\frac{1}{2}v_\rho\|\Delta_c\|^2 \leqslant -\frac{1}{2}v_\rho\eta^2\rho_\varepsilon^2\,. \qquad (9)$$

To conclude the proof we plug-in (8) and (9) into (7). Then, it remains to use that $n \geqslant b^i$ together with the fact that $n \mapsto nf(t/n)$ is non decreasing. □

B.3. Localized change of measure

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

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

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

$$\begin{aligned}\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n_i\leqslant n<n_{i+1}} E_{c,p}(n,t)\Big\} \;\leqslant\; & \mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n_i\leqslant n<n_{i+1}} E_{c,p}(n,t)\cap \|\nabla\psi(\theta_c^\star) - \widehat{F}_n\| < \varepsilon_{t,i,c}\Big\} \\ & + \mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n_i\leqslant n<n_{i+1}} E_{c,p}(n,t)\cap \|\nabla\psi(\theta_c^\star)-\widehat{F}_n\| \geqslant \varepsilon_{t,i,c}\Big\}\,. \quad (10)\end{aligned}$$

We handle the first term in (10) 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 B.4. We will show more precisely that

Lemma 30 (Change of measure) *For any sequence of positive values $\{\varepsilon_{t,i,c}\}_{i\geqslant 0}$, it holds*

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

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

Let us recall that $E_{c,p}(n,t) = \{\widehat{\theta}_n\in\Theta_\rho \cap \widehat{F}_n\in\mathcal{C}_p(\theta_c^\star)\cap n\mathcal{B}^\psi(\widehat{\theta}_n,\theta_c^\star)\geqslant f(t/n)\}$.

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 $B = \Theta_\rho\cap\nabla\psi^{-1}\big(\mathcal{C}_p(\theta_c^\star)\cap\mathcal{B}_2(\nabla\psi(\theta_c^\star),\varepsilon_{t,i,c})\big)$ where $\mathcal{B}_2(y,r)\stackrel{\text{def}}{=}\Big\{y'\in\mathbb{R}^K\,;\,\|y-y'\|\leqslant r\Big\}$. This makes sense since, on the one hand, under $E_{c,p}(n,t)$, $\nabla\psi(\widehat{\theta}_n)\in\mathcal{C}_p(\theta_c^\star)$, and on the other hand, $\|\nabla\psi(\widehat{\theta}_n) - \nabla\psi(\theta_c^\star)\|\leqslant\varepsilon_{t,i,c}$. For convenience, let us introduce the event of interest

$$\Omega = \Big\{\bigcup_{n_i\leqslant n<n_{i+1}} E_{c,p}(n,t)\cap\|\nabla\psi(\widehat{\theta}_n) - \nabla\psi(\theta_c^\star)\|\leqslant\varepsilon_{t,i,c}\Big\}\,.$$

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 $Q_B(\Omega) \stackrel{\text{def}}{=} \int_{\theta' \in B} \mathbb{P}_{\theta'}\Big\{\Omega\Big\} d\theta'$ 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 the event Ω.

$$\begin{aligned}
\int_\Omega \frac{d\mathbb{P}_\theta}{dQ_B} dQ_B &= \int_\Omega \left[\int_{\theta' \in B} \frac{\Pi_{k=1}^n \nu_{\theta'}(X_k)}{\Pi_{k=1}^n \nu_\theta(X_k)} d\theta' \right]^{-1} dQ_B \\
&= \int_\Omega \left[\int_{\theta' \in B} \exp\Big(n\langle \theta' - \theta, \widehat{F}_n \rangle - n\big(\psi(\theta') - \psi(\theta)\big) \Big) d\theta' \right]^{-1} dQ_B \,.
\end{aligned}$$

It is now convenient to 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 $\nabla\psi(\widehat{\theta}_n) = \widehat{F}_n$, it holds

$$\langle \theta' - \theta, \widehat{F}_n \rangle - \big(\psi(\theta') - \psi(\theta)\big) = \mathcal{B}^\psi(\widehat{\theta}_n, \theta) - \mathcal{B}^\psi(\widehat{\theta}_n, \theta') \,.$$

Thus, the previous likelihood ratio simplifies as follows

$$\begin{aligned}
\frac{d\mathbb{P}_\theta}{dQ_B} &= \left[\int_{\theta' \in B} \exp\Big(n\mathcal{B}^\psi(\widehat{\theta}_n, \theta) - n\mathcal{B}^\psi(\widehat{\theta}_n, \theta') \Big) d\theta' \right]^{-1} \\
&\leqslant \left[\int_{\theta' \in B} \exp\Big(f(t/n) - n\mathcal{B}^\psi(\widehat{\theta}_n, \theta') d\theta' \right]^{-1} \\
&= \exp\Big(- f(t/n) \Big) \left[\int_{\theta' \in B} \exp\Big(- n\mathcal{B}^\psi(\widehat{\theta}_n, \theta') \Big) dx \right]^{-1} ,
\end{aligned}$$

where we note that both θ' and $\widehat{\theta}_n$ belong to Θ_ρ.

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

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

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. A first remark is that $\theta_c^\star$ plays a central role in B: It is not difficult to show that, by construction of B,

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

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 $\nabla\psi(\theta')\in\nabla\psi(\theta_c^\star)+\mathcal{B}_2(0,\varepsilon_{t,i,c})\cap\mathcal{C}_p(0,\Delta_c)$. 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 $\tilde r=\min\{v_\rho\rho,\varepsilon_{t,i,c}\}$. 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

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

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

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

where (d) follows from another change of parameter argument, with $r_\rho=\sqrt{\frac{V_\rho}{2v_\rho^2}}\min\{v_\rho\rho,\varepsilon_{t,i,c}\}$ 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 $p'\geqslant p$ (remember that the larger p, the more acute is a cone) and Δ_n is chosen such that $\nabla\psi(\widehat\theta_n)\in\nabla\psi(\theta_c^\star)+\mathcal{D}$ (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 point 0 provided that p' is not too small, more precisely provided that $p'>2/\sqrt5$. 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 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 amongst these point provided that p' is

not too small. More formally, the function g rewrites

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

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

Proof : First, let us 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 distance at most $2\frac{\sqrt{1-p'^2}}{p'}\tilde{r}$ from each other. Now they are also at 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

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

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

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

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

Lemma 32 *For all $\varepsilon, \varepsilon' > 0$, $p, p' \in [0,1]$ and all $K \geqslant 1$ the following equality and inequality hold*

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

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

Applying this Lemma, and introducing $r_\rho = \sqrt{\frac{V_\rho}{2v_\rho^2}}\min\{v_\rho\rho, \varepsilon_{t,i,c}\}$, we obtain that $\mathbb{P}_{\theta_c^\star}\Big\{\Omega\Big\}$ is

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

This concludes the proof of Lemma 30.

B.4. Concentration of measure

In this section, we focus on the second term in (10), that is $\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n=n_i}^{n_{i+1}-1} E_{c,p}(n,t)\cap\|\nabla\psi(\theta_c^\star) - \widehat{F}_n\| \geqslant \varepsilon_{t,i,c}\Big\}$. In this term, $\varepsilon_{t,i,c}$ should be considered as decreasing fast to 0 with i, and slowly increasing with t. Note that by definition $\nabla\psi(\widehat{\theta}_n) = \widehat{F}_n = \frac{1}{n}\sum_{i=1}^n F(X_i) \in \mathbb{R}^K$ 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 33 (Concentration of measure) *Let $\varepsilon_c^{\max} = Diam(\nabla\psi(\Theta_\rho\cap\mathcal{C}_{c,p}))$ where we introduced the projected cone $\mathcal{C}_{c,p} = \{\theta\in\Theta : \langle \frac{\Delta_c}{\|\Delta_c\|}, \frac{\nabla\psi(\theta_c^\star)-\nabla\psi(\theta)}{\|\nabla\psi(\theta_c^\star)-\nabla\psi(\theta)\|}\rangle \geqslant p\}$. Then, for all $\varepsilon_{t,i,c}$, it holds*

$$\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n=n_i}^{n_{i+1}-1} E_{c,p}(n,t)\cap\|\nabla\psi(\widehat{\theta}_n)-\nabla\psi(\theta_c^\star)\| \geqslant \varepsilon_{t,i,c}\Big\} \leqslant \exp\left(-\frac{n_i^2 p\varepsilon_{t,i,c}^2}{2V_\rho(n_{i+1}-1)}\right)\mathbb{I}\{\varepsilon_{t,i,c}\leqslant\overline{\varepsilon}_c\}.$$

Proof : Note that by definition if $\varepsilon_{t,i,c} > \varepsilon_c^{\max}$, then

$$\mathbb{P}_{\theta_c^\star}\Big\{ \bigcup_{n_i \leqslant n < n_{i+1}} E_{c,p}(n,t) \cap \|\nabla\psi(\widehat{\theta}_n) - \nabla\psi(\theta_c^\star)\| \geqslant \varepsilon_{t,i,c} \Big\} = 0\,.$$

We thus restrict to the case when $\varepsilon_{t,i,c} \leqslant \varepsilon_c^{\max}$, or equivalently, replace $\varepsilon_{t,i,c}$ by the quantity $\tilde{\varepsilon}_{t,i,c} = \min\{\varepsilon_{t,i,c}, \varepsilon_c^{\max}\}$. Now, by definition of the event $E_{c,p}(n,t)$, we have the rewriting

$$\begin{aligned}
&\mathbb{P}_{\theta_c^\star}\Big\{ \bigcup_{n_i \leqslant n < n_{i+1}} E_{c,p}(n,t) \cap \|\nabla\psi(\widehat{\theta}_n) - \nabla\psi(\theta_c^\star)\| \geqslant \tilde{\varepsilon}_{t,i,c} \Big\} \\
&\leqslant \mathbb{P}_{\theta_c^\star}\Big\{ \bigcup_{n_i \leqslant n < n_{i+1}} \widehat{\theta}_n \in \Theta_\rho \cap \langle \frac{\Delta_c}{\|\Delta_c\|}, \nabla\psi(\theta_c^\star) - \nabla\psi(\widehat{\theta}_n)\rangle \geqslant p\tilde{\varepsilon}_{t,i,c} \Big\} \\
&\leqslant \mathbb{P}_{\theta_c^\star}\Big\{ \bigcup_{n=n_i}^{n_{i+1}-1} \langle \frac{\Delta_c}{\|\Delta_c\|}, \sum_{i=1}^n \big(\nabla\psi(\theta_c^\star) - F(X_i)\big)\rangle \geqslant p n_i \tilde{\varepsilon}_{t,i,c} \Big\}\,.
\end{aligned}$$

Applying on both side of the inequality the function $x \mapsto \exp(\lambda x)$, for a deterministic $\lambda > 0$, it comes

$$\begin{aligned}
&\mathbb{P}_{\theta_c^\star}\Big\{ \bigcup_{n_i \leqslant n < n_{i+1}} E_{c,p}(n,t) \cap \|\nabla\psi(\widehat{\theta}_n) - \nabla\psi(\theta_c^\star)\| \geqslant \tilde{\varepsilon}_{t,i,c} \Big\} \\
&\overset{(a)}{\leqslant} \mathbb{P}_{\theta_c^\star}\Big\{ \bigcup_{n=n_i}^{n_{i+1}-1} \exp\Big(\sum_{i=1}^n \langle \frac{\lambda\Delta_c}{\|\Delta_c\|}, \nabla\psi(\theta_c^\star) - F(X_i)\rangle \Big) \geqslant \exp\Big(\lambda p n_i \tilde{\varepsilon}_{t,i,c}\Big) \Big\} \\
&= \mathbb{P}_{\theta_c^\star}\Big\{ \bigcup_{n=n_i}^{n_{i+1}-1} \exp\Big(\sum_{i=1}^n \langle \frac{\lambda\Delta_c}{\|\Delta_c\|}, \nabla\psi(\theta_c^\star) - F(X_i)\rangle - \frac{\lambda^2(n_{i+1}-1)}{2} V_\rho \Big) \geqslant \\
&\qquad\qquad \exp\Big(\lambda p n_i \tilde{\varepsilon}_{t,i,c} - \frac{\lambda^2(n_{i+1}-1)}{2} V_\rho\Big) \Big\} \\
&\leqslant \mathbb{P}_{\theta_c^\star}\Big\{ \bigcup_{n=n_i}^{n_{i+1}-1} \exp\Big(\sum_{i=1}^n \langle \frac{\lambda\Delta_c}{\|\Delta_c\|}, \nabla\psi(\theta_c^\star) - F(X_i)\rangle - \frac{\lambda^2 n}{2} V_\rho \Big) \geqslant \\
&\qquad\qquad \exp\Big(\lambda p n_i \tilde{\varepsilon}_{t,i,c} - \frac{\lambda^2(n_{i+1}-1)}{2} V_\rho\Big) \Big\}\,.
\end{aligned}$$

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

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

that is $\mathbb{E}\Big[W_n(e,\lambda)\Big|H_{n-1}\Big] \leqslant W_{n-1}(e,\lambda)$. Thus, we apply Doob's maximal inequality for non-negative super-martingale and deduce that

$$\begin{aligned}
&\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n_i\leqslant n<n_{i+1}} E_{c,p}(n,t)\cap\|\nabla\psi(\widehat{\theta}_n)-\nabla\psi(\theta_c^\star)\|\geqslant\tilde{\varepsilon}_{t,i,c}\Big\}\\
&\leqslant\ \mathbb{P}_{\theta_c^\star}\Big\{\max_{n_i\leqslant n<n_{i+1}} W_n(\lambda)\geqslant\exp\Big(\lambda pn_i\tilde{\varepsilon}_{t,i,c}-\lambda^2(n_{i+1}-1)V_\rho/2\Big)\Big\}\\
&\leqslant\ \mathbb{E}_{\theta_c^\star}[W_{n_i}(\lambda)]\exp\Big(-\lambda pn_i\tilde{\varepsilon}_{t,i,c}+\lambda^2(n_{i+1}-1)V_\rho/2\Big)\\
&\leqslant\ \exp\Big(-\lambda pn_i\tilde{\varepsilon}_{t,i,c}+\lambda^2(n_{i+1}-1)V_\rho/2\Big).
\end{aligned}$$

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

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

Indeed, it suffices to show that $\lambda^\star\leqslant\lambda_c$ to ensure that the condition is satisfied. It is now not difficult to relate λ_c to $\varepsilon_c^{\max}$: 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

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

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

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

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

$$\mathbb{P}_{\theta_c^\star}\Big\{\bigcup_{n_i\leqslant n<n_{i+1}} E_{c,p}(n,t)\cap\|\nabla\psi(\widehat{\theta}_n)-\nabla\psi(\theta_c^\star)\|\geqslant\varepsilon_{t,i,c}\Big\}\leqslant\exp\Big(-\frac{n_i^2p\varepsilon_{t,i,c}^2}{2V_\rho(n_{i+1}-1)}\Big)\mathbb{I}\{\varepsilon_{t,i,c}\leqslant\overline{\varepsilon}_c\}.$$

□

B.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 $\{\varepsilon_{t,i,c}\}_{t,i}$, then

$$\begin{aligned}
[1]&\stackrel{\mathrm{def}}{=}\mathbb{P}_{\theta^\star}\Big\{\bigcup_{1\leqslant n\leqslant t}\widehat{\theta}_n\in\Theta_\rho\cap\mathcal{K}(\Pi(\widehat{\nu}_n),\mu^\star-\varepsilon)\geqslant f(t/n)/n\Big\}\\
&\leqslant\sum_{c=1}^{C_{p,\eta,K}}\sum_{i=0}^{I_t-1}\underbrace{\exp\Big(-n_i\alpha^2-\chi\sqrt{n_if(t/n_i)}\Big)}_{\text{change of measure}}\Big[\underbrace{\exp\Big(-\frac{n_i^2p\varepsilon_{t,i,c}^2}{2V_\rho(n_{i+1}-1)}\Big)\mathbb{I}\{\varepsilon_{t,i,c}\leqslant\overline{\varepsilon}_c\}}_{\text{concentration}}\\
&\quad+\underbrace{\alpha_{p,K}\exp\Big(-f\Big(\frac{t}{n_{i+1}-1}\Big)\Big)\min\Big\{\rho^2v_\rho^2,\varepsilon_{t,i,c}^2,\frac{(K+2)v_\rho^2}{K(n_{i+1}-1)V_\rho}\Big\}^{-K/2}\varepsilon_{t,i,c}^K}_{\text{localization + change of measure}}\Big],
\end{aligned}$$

where we recall that $\alpha = \alpha(p,\eta,\varepsilon) = \eta\rho_\varepsilon\sqrt{v_\rho/2}$ 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

$$\begin{aligned}[1] \quad \leqslant \quad & \sum_{c=1}^{C_{p,\eta,K}} \sum_{i=0}^{I_t-1} \exp\Big(- n_i\alpha^2 - \chi\sqrt{n_i f(t/n_i)}\Big)\Big[\alpha_{p,K}\exp\Big(-f\Big(\frac{t}{n_{i+1}-1}\Big)\Big) \times \\ & \max\Big\{\frac{\varepsilon_{t,i,c}}{\rho v_\rho}, 1, \sqrt{\frac{(n_{i+1}-1)V_\rho}{1+2/K}}\frac{\varepsilon_{t,i,c}}{v_\rho}\Big\}^K + \exp\Big(-\frac{n_i^2 p\varepsilon_{t,i,c}^2}{2V_\rho(n_{i+1}-1)}\Big)\mathbb{I}\{\varepsilon_{t,i,c} \leqslant \overline{\varepsilon}_c\}\Big]\,,\end{aligned}$$

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

$$\begin{aligned}[1] \quad \leqslant \quad & \sum_{i=0}^{I_t-1} \exp\Big(- n_i\alpha^2 - \chi\sqrt{n_i f(t/n_i)} - f(t/(n_{i+1}-1))\Big) f(t/(n_{i+1}-1))^{K/2} \times \\ & C_{p,\eta,K}\Big(\alpha_{p,K}\max\Big\{\frac{2V_\rho}{p\rho^2 v_\rho^2 b^{i-1}}, 1, \frac{b^2V_\rho^2}{pv_\rho^2(\frac{1}{2}+\frac{1}{K})}\Big\}^{K/2} + 1\Big)\,.\end{aligned}$$

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

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

which leads to the final bound

$$\begin{aligned} & \mathbb{P}_{\theta^\star}\Big\{\bigcup_{1\leqslant n\leqslant t} \widehat{\theta}_n \in \Theta_\rho \cap \mathcal{K}(\Pi(\widehat{\nu}_n), \mu^\star - \varepsilon) \geqslant f(t/n)/n\Big\} \\ & \leqslant \quad C(K,\rho,p,b,\eta)\sum_{i=0}^{I_t-1}\exp\Big(- n_i\alpha^2 - \chi\sqrt{n_i f(t/n_i)} - f(t/(n_{i+1}-1))\Big) f(t/(n_{i+1}-1))^{K/2}\,.\end{aligned}$$

Appendix C. Fine-tuned upper bounds

In this section, we study the behavior of the bound obtained in Theorem 19 as a function of t, for a specific choice of function f, namely $f(x) = \log(x) + \xi\log\log x$, and prove corollary 22 and corollary 23, 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 to show that $\xi > K/2-1$ is enough. Let us recall that f is non-decreasing only for $x \geqslant e^{-\xi}$. We thus restrict to $t \geqslant e^{-\xi}$ in corollary 22 that uses the threshold $f(t)$, and to $\xi \geqslant 0$ in corollary 23 that uses the threshold function $f(t/n)$. In the sequel, we use the short-hand notation C in order to replace $C(K,\rho,p,b,\eta)$.

C.1. Proof of Corollary 22

As a warming-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 19 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 non-increasing on the set of x such that $\xi > -\log(x)$, Theorem 19 specifies for all $b > 1, p, q, \eta \in (0,1)$, to

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

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 $S \leqslant \lceil \log_b(qt) \rceil \leqslant \log_b(qt) + 1$. Thus, setting $q = 1$ and $b = 2$ we obtain

$$\begin{aligned}
&\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{1\leqslant n<t} \widehat{\theta}_n \in \Theta_\rho \cap \mathcal{K}(\Pi(\widehat{\nu}_n), \mu^\star - \varepsilon) \geqslant f(t)/n \Big\} \\
&\leqslant \frac{C}{t} \Big(1 + \underbrace{\xi \frac{\log\log(t)}{\log(t)}}_{o(1)} \Big)^{K/2} \log(t)^{-\xi + K/2} (\log_2(t) + 1) .
\end{aligned}$$

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 $i \geqslant 0$, it holds by definition

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

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

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

Thus, S is asymptotically $o(1)$, and we deduce that $\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{1\leqslant n<t} \widehat{\theta}_n \in \Theta_\rho \cap \mathcal{K}(\Pi(\widehat{\nu}_n), \mu^\star - \varepsilon) \geqslant f(t)/n \Big\} = o(1/t)$ 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 $b = \lceil (1 + \frac{\log(1+\chi)}{\chi})^2 \rceil \leqslant 4$,

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{1 \leqslant n < t} \widehat{\theta}_n \in \Theta_\rho \cap \mathcal{K}(\Pi(\widehat{\nu}_n), \mu^\star - \varepsilon) \geqslant f(t)/n \Big\}$$
$$\leqslant \frac{C(1+\chi)}{t\chi} \left(1 + \xi \frac{\log\log(t)}{\log(t)}\right)^{K/2} \log(t)^{-\xi + K/2} \exp(-\chi\sqrt{\log(t) + \xi \log\log(t)})\,.$$

C.2. Proof of Corollary 23

Let us now focus on the proof of Corollary 23 involving the threshold $f(t/n)$. We consider the choice $f(x) = \log(x) + \xi \log\log x$, which is non-increasing 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 $\xi \geqslant 0$. Now, following the result of Theorem 19, we thus obtain for all $b > 1, p, q, \eta \in (0, 1)$,

$$\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{1 \leqslant n < t} \widehat{\theta}_n \in \Theta_\rho \cap \mathcal{K}(\Pi(\widehat{\nu}_n), \mu^\star - \varepsilon) \geqslant f(t/n)/n \Big\} \leqslant C \exp\left(-\frac{\alpha^2 q}{b} t - \chi\sqrt{\frac{tqf(b/q)}{b}}\right)$$
$$+ C \sum_{i=0}^{\lceil \log_b(qt) \rceil - 2} \exp\left(-\alpha^2 b^i - \chi\sqrt{b^i f(t/b^i)} - f(t/(b^{i+1} - 1))\right) f\left(\frac{t}{b^{i+1} - 1}\right)^{\frac{K}{2}}$$
$$= C \exp\left(-\frac{\alpha^2 q}{b} t - \chi\sqrt{\frac{tqf(b/q)}{b}}\right)$$
$$+ C \sum_{i=0}^{\lceil \log_b(qt) \rceil - 2} \underbrace{e^{-\alpha^2 b^i - \chi\sqrt{b^i f(t/b^i)}} \left(\frac{b^{i+1} - 1}{t}\right) \log\left(\frac{t}{b^{i+1} - 1}\right)^{K/2 - \xi}}_{s_i} \Bigg(1 + \underbrace{\xi \frac{\log\log\left(\frac{t}{b^{i+1} - 1}\right)}{\log\left(\frac{t}{(b^{i+1} - 1)}\right)}}_{o(1)}\Bigg)^{\frac{K}{2}}. \tag{11}$$

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

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

Small values of i We start by handling the terms corresponding to small values of $i \leqslant i_0$, 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

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

from which we deduce that

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

Following Lai (1988), 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 $i_0 \geqslant 0$ we obtain the bounds[6]

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

We easily see that this is $o(1/t)$ both when $\xi > K/2$ and when $\xi \leqslant K/2$, 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 $i \geqslant i_0$. We note that the initial proof of Lai (1988) 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 $s'_i = \exp(-\chi\sqrt{b^i\log(t/b^i)})b^{i+1}$, and better on the following ratio

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

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

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

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 ineer brackets converges to 0 in such cases, and thus the ratio is asymptotically upper bounded by $b > 1$. Thus we impose that $\eta \leqslant 1$, that is $\xi \geqslant K/2 - 1$.

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 $x \geqslant 8.2$. 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

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

that, provided that $t \geqslant 8.2(2\log(4))^2\chi^{-2} \simeq 63\chi^{-2}$, then $\frac{s'_{i+1}}{s'_i} < 1$. Now, in order to get a ratio $\frac{s'_{i+1}}{s'_i}$ that is away from 1, we target the bound $\frac{s'_{i+1}}{s'_i} < b/(b+1)$. This can be achieved by requiring that $t \geqslant 8.2(2\log(5))^2\chi^{-2} \simeq 85\chi^{-2}$ by setting $c = \chi^2/(2\log(5))^2$. Eventually, we obtain for $b = 4$ and $t \geqslant 85\chi^{-2}$ the bound

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

Remark 34 *Another notable value is $\eta = 0$. A similar study than 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 $t \geqslant 12(2\log(3.5))^2\chi^{-2} \simeq 76\chi^{-2}$.*

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

$$\sum_{i=i_0+1}^{I_t-2} s_i \leqslant \begin{cases}\frac{b^2(b+1)}{tc\log(tc)}\log(1/q)^{K/2-\xi} & \text{if } \xi \geqslant K/2\\ \frac{b^2(b+1)}{tc\log(tc)}\log(t\frac{c\log(tc)}{b-c\log(tc)})^{K/2-\xi} & \text{if } \xi \in [K/2-1, K/2]\,.\end{cases} \tag{12}$$

It remains to handle the case when $i_0 < 0$. Note that this case only happens for t large enough so that $t > c^{-1}e^{\frac{1}{bc}}$. The later quantity may be huge since $1/bc = \log(5)^2\chi^{-2}$ is possibly 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

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

Thus, if we define t_χ to be the smallest such t, then when $t > c^{-1}e^{\frac{1}{bc}}$ and provided that $t \geqslant t_\chi$, the bound of (12) remains valid for the sum S, up to replacing $b^2(b+1)$ with $2b^2(b+1/2)$ and $\log(t\frac{c\log(tc)}{b-c\log(tc)})$ with $\log(t/(b-1))$. The later constraint $t \geqslant t_\chi$ is satisfied as soon as $4\log(5)^2\chi^{-2}e^{\chi^{-2}\log(5)^2} \geqslant t_\chi$ 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 $t \geqslant 85\chi^{-2}$ and $t \leqslant \chi^{-2}\frac{\exp\left(\chi^{-2}\log(4.5)^2\right)}{4\log(4.5)^2}$, the following bound

$$S \leqslant \frac{b}{ct\log(tc)}\begin{cases}\frac{b}{(b-1)}\log(tc\log(tc)/b)^{\frac{K}{2}-\xi} + b(b+1)\log(1/q)^{\frac{K}{2}-\xi} & \text{if } \xi \geqslant K/2\\ \frac{b}{(b-1)}\log(t/(b-1))^{\frac{K}{2}-\xi} + b(b+1)\log(t\frac{c\log(tc)}{b-c\log(tc)})^{\frac{K}{2}-\xi} & \text{if } \xi \in [K/2-1, K/2]\end{cases}. \tag{13}$$

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

$$S \leqslant \frac{2b^2(b+1/2)}{ct\log(tc)}\begin{cases}\log(1/q)^{K/2-\xi} & \text{if } \xi \geqslant K/2\\ \log(t/(b-1))^{K/2-\xi} & \text{if } \xi \in [K/2-1, K/2]\,.\end{cases} \tag{14}$$

Concluding step In this final step, we now gather equation (11) together with the previous bounds (13), (14) on S. We obtain that for all $p, q, \eta \in (0,1)$

$$\begin{aligned}
&\mathbb{P}_{\theta^\star}\Big\{ \bigcup_{1\leqslant n<t} \widehat{\theta}_n \in \Theta_\rho \cap \mathcal{K}(\Pi(\widehat{\nu}_n), \mu^\star - \varepsilon) \geqslant f(t/n)/n \Big\} \\
&\leqslant \; C(K,\rho,p,b,\eta)\Big(e^{-\frac{\alpha^2 qt}{b} - \sqrt{\frac{\chi^2 tqf(b/q)}{b}}} + S(1+\xi)^{K/2} \Big) .
\end{aligned}$$

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

When $\xi \in [K/2-1, K/2]$, one can then choose $q = 1$. When $\xi \geqslant K/2$, 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 D. Technical details

Lemma 32 *For all $\varepsilon, \varepsilon' > 0$, $p, p' \in [0,1]$ and all $K \geqslant 1$ the following equality holds*

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

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

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

Proof of Lemma 32: First of all, let us remark that provided that $K \geqslant 2$, then

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

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 $K \geqslant 2$,

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

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

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

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

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

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

$$\int_0^\varepsilon e^{-r^2/2} r^{K-1} dr \geqslant \frac{\varepsilon^K}{2K}.$$

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

$$\int_0^\varepsilon e^{-r^2/2} r^{K-1} dr \quad \geqslant \quad \int_0^{(K+2)/K} e^{-r^2/2} r^{K-1} dr \geqslant \frac{\sqrt{1+2/K}^K}{2K} .$$

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

Proof of Lemma 6: The first part of this lemma for `KL-ucb` is proved in Cappé et al. (2013). 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: We start by introducing a small $\varepsilon > 0$ that satisfies $\varepsilon < \min\{ \mu^\star - \mu_a \, , \, a \in \mathcal{A} \setminus \{a^\star\} \}$, and then consider the following inclusion of events:

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

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

$$\Big\{ \mu^\star - \varepsilon < U_a(t) \Big\} \subseteq \Big\{ \exists \nu' \in \mathcal{D} \colon E(\nu') > \mu^\star - \varepsilon \text{ and } N_a(t) \, \mathcal{K}_a\big(\Pi_a(\widehat{\nu}_{a,N_a(t)}), \, \mu^\star - \varepsilon \big) \leqslant f(t/N_a(t)) \Big\} ,$$
$$\text{and} \quad \Big\{ \mu^\star - \varepsilon \geqslant U_{a^\star}(t) \Big\} \subseteq \Big\{ \exists \nu' \in \mathcal{D} \colon N_{a^\star}(t) \, \mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\widehat{\nu}_{a^\star,N_{a^\star}(t)}), \, \mu^\star - \varepsilon \big) > f(t/N_{a^\star}(t)) \Big\} ,$$

since $\mathcal{K}_a$ is a non-decreasing function in its second argument and $\mathcal{K}_a\big(\nu, E(\nu)\big) = 0$ for all distributions ν. Therefore, this simple remark leads us to the following decomposition

$$\mathbb{E}\big[N_T(a)\big] \leqslant 1 + \sum_{t=|\mathcal{A}|}^{T-1} \mathbb{P}\Big\{ N_{a^\star}(t) \; \mathcal{K}_{a^\star}\big(\Pi_{a^\star}(\widehat{\nu}_{a^\star,N_{a^\star}(t)}), \, \mu^\star - \varepsilon \big) > f(t/N_{a^\star}(t)) \Big\}$$
$$+ \sum_{t=|\mathcal{A}|}^{T-1} \mathbb{P}\Big\{ N_a(t) \; \mathcal{K}_a\big(\Pi_a(\widehat{\nu}_{a,N_a(t)}), \, \mu^\star - \varepsilon \big) \leqslant f(t/N_a(t)) \;\text{ and }\; a_{t+1} = a \Big\} .$$

The remaining steps of the proof of the result from Cappé et al. (2013), equation (10) can now be straightforwardly modified to work with $f(t/N_a(t))$ instead of $f(t)$, thus concluding this proof. □
