

# Bregman Deviations of Generic Exponential Families

**Sayak Ray Chowdhury**

*Microsoft Research, India*

T-SAYAKR@MICROSOFT.COM

**Patrick Saux**

*Univ. Lille, Inria, CNRS, UMR 9189 - CRISyAL, F-59000, Lille, France*

PATRICK.SAUX@INRIA.FR

**Odalric-Ambrym Maillard**

*Univ. Lille, Inria, CNRS, UMR 9189 - CRISyAL, F-59000, Lille, France*

ODALRIC.MAILLARD@INRIA.FR

**Aditya Gopalan**

*Indian Institute of Science, Bangalore, India*

ADITYA@IISC.AC.IN

**Editors:** Gergely Neu and Lorenzo Rosasco

## Abstract

We revisit the method of mixtures, or Laplace method, to study the concentration phenomenon in generic (possibly multidimensional) exponential families. Using duality properties of the Bregman divergence to construct nonnegative martingales, we establish a generic inequality controlling the deviation between the parameter of the family and a finite sample estimate. This bound is expressed in the local geometry induced by the Bregman pseudo-metric. Moreover, it is time-uniform and involves a quantity extending the classical *information gain* to exponential families, which we call the *Bregman information gain*. For the practitioner, we instantiate this novel bound to several classical families, e.g., Gaussian (including with unknown variance or multivariate), Bernoulli, Exponential, Weibull, Pareto, Poisson and Chi-square, yielding explicit forms of the confidence sets and the Bregman information gain. We further compare the resulting confidence bounds to state-of-the-art time-uniform alternatives and show this novel method yields competitive results. Finally, we apply our result to the design of generalized likelihood ratio tests for change detection, capturing new settings such as variance change in Gaussian families.

**Keywords:** Exponential families, Bregman divergence, concentration bounds.

## 1. Introduction

Concentration inequalities are a powerful set of methods in statistical theory with key applications in machine learning. Often in machine learning applications, a learner estimates some quantity solely based on samples from an unknown distribution and would like to know the magnitude of the estimation error. The typical example is that of the mean  $\mu$  of some real-valued random variable  $X$ , estimated by its empirical mean built from a sample of  $n$  independent and identically distributed (i.i.d.) observations. We refer the interested reader to the monographs of [Boucheron et al. \(2013\)](#); [Raginsky and Sason \(2018\)](#); [Zeitouni and Dembo \(1998\)](#) for standard results and related topics.

In many situations, one may want to further estimate some vector parameter, as in, e.g., linear bandits ([Abbasi-Yadkori et al., 2011](#)) and logistic bandits ([Faury et al., 2020](#)). A closely related problem is to estimate the parameter of a distribution coming from a parametric family ([Chowdhury et al., 2021](#)). *Exponential families* are a flexible way to formalize such distributions over a set  $\mathcal{X}$  by describing densities of the form  $p_\theta(x) \propto h(x) \exp(\langle \theta, F(x) \rangle)$ , for some given feature function  $F : \mathcal{X} \mapsto \mathbb{R}^d$  and base function  $h$  (Section 2). Most classical distributions fall into this category, e.g.,

Gaussian (possibly multivariate, with or without known variance), Exponential, Gamma, Chi-square, Weibull, Pareto, Poisson, Bernoulli and Multinomial distributions ([Amari, 2016](#)).

A number of problems in current-day machine learning involve sequential, active data-sampling strategies ([Cesa-Bianchi and Lugosi, 2006](#)). This includes multi-armed bandits, reinforcement learning, active learning and federated learning, to mention a few application domains. Since the decision to sample a novel observation results from the interaction between the learning algorithm and the environment, and depends on past observations, one needs to design concentration inequalities working with a *random* number of observations typically at a random stopping time ([Durrett, 2019](#)). A natural way to handle this difficulty is to derive time-uniform concentration inequalities, producing sequences of confidence sets valid *uniformly* over all number of observations with high probability, as opposed to being valid for a single number of observation.

A popular method in bandit theory is to combine supermartingale techniques with union bound arguments over a geometric time grid, a technique known as time peeling (or stitching) – see [Bubeck \(2010\)](#); [Cappé et al. \(2013\)](#) for early uses in bandits, as well as [Garivier \(2013\)](#), or more recently [Maillard \(2019b\)](#); see also [Howard et al. \(2020, 2021\)](#) for a recent, complementary survey of the history of this field, and [Kuchibhotla and Zheng \(2021\)](#) for an extension of Bentkus' concentration bounds ([Bentkus, 2004](#)) for bounded distributions using time peeling. The method of mixtures, initiated by [Robbins and Pitman \(1949\)](#); [Robbins \(1970\)](#) and popularized in [Peña et al. \(2008\)](#) is a powerful alternative to peeling for developing time-uniform confidence sets. It has been applied to sub-Gaussian families in [Abbasi-Yadkori et al. \(2011\)](#), leading to a variety of applications ([Chowdhury and Gopalan, 2017](#); [Durand et al., 2018](#); [Kirschner and Krause, 2018](#)). In [Kaufmann and Koolen \(2021\)](#), a generalization to handle one-dimensional exponential families is considered, with applications to Gaussian (known variance) and Gamma distributions (known shape). A fairly different ‘capital process’ construction technique has been recently developed for bounded distributions in [Shafer and Vovk \(2019\)](#), and popularized further in [Waudby-Smith and Ramdas \(2023\)](#).

In this work, we revisit the method of mixtures for parametric exponential families of arbitrary dimension, expressing deviations in the natural (Bregman) divergence of the family. The setting of exponential families is convenient for integration in a Bayesian setup, thanks to the notion of conjugate prior that enables us to reduce computation of tedious integrals to simple parameter updates. Here, we exploit this property to obtain explicit mixtures of martingales. Exponential families are largely used in modern machine learning, yet concentration tools available to the practitioner are comparatively scarce beyond the Gaussian case. To help close this gap, we obtain both sharp and computationally tractable confidence sets, especially in the small sample regime.

**Outline and contributions.** In Section 2, we first recall some background material on exponential families and their associated Bregman divergences. Section 3.1 states our main result (Theorem 3): a *time-uniform* concentration inequality for exponential families. Specifically, we control the Bregman deviations associated with the log-partition function of the family using a novel information-theoretic quantity, the *Bregman information gain*. On a high level, this quantifies the information gain about the parameter of the family after observing i.i.d. samples from it, which is measured in terms of the natural Bregman divergence of the family. To illustrate the utility of this general result, we detail in Section 3.2 how Bregman information gain and deviation inequalities specialize for well-known exponential families, resulting in fully explicit confidence sets (see Table 1). To the best of our knowledge, we are the first to derive an explicit *time-uniform* deviation inequality for two-parameter Gaussian (i.e. both mean and variance are unknown), Chi-square, Weibull, Pareto,

and Poisson distributions. Our result is an adaptation of the *method of mixtures* technique, and a proof sketch is outlined in Section 3.3. In Section 4, we numerically evaluate the high-probability confidence sets built from our method for classical families, and achieve state-of-the-art time-uniform bounds. Finally, in Section 5, we generalize Theorem 3 to obtain a *doubly time-uniform* concentration inequality for generic exponential families, which could be of independent interest (Theorem 5). We present an application of both results in controlling the false alarm probability of the Generalized Likelihood Ratio (GLR) test, used for change detection in the exponential family model.

## 2. Exponential Families and Bregman Divergence

In this section, we introduce exponential families and the link between their Kullback-Leibler (KL) and Bregman divergences, as well as useful properties of these divergences.

**Exponential families.** We consider an exponential family of distributions  $\{p_\theta\}_{\theta \in \Theta}$  over some set  $\mathcal{X}$ , parameterized in some open set  $\Theta \subset \mathbb{R}^d$ , whose density or mass function has the form  $p_\theta(x) = h(x) \exp(\langle \theta, F(x) \rangle - \mathcal{L}(\theta))$ . Here,  $F : \mathcal{X} \rightarrow \mathbb{R}^d$  is the feature function,  $h : \mathcal{X} \rightarrow \mathbb{R}_+$  is the base function, and  $\mathcal{L}$  represents the normalization term (a.k.a. log-partition function, convex w.r.t.  $\theta$ ) given by  $\mathcal{L}(\theta) = \log \int_{\mathcal{X}} h(x) \exp(\langle \theta, F(x) \rangle) dx$ . We denote by  $\Theta_D = \{\theta \in \mathbb{R}^d : \mathcal{L}(\theta) < \infty\}$  the domain of  $\mathcal{L}$  and by  $\Theta_I = \{\theta \in \Theta_D : \det \nabla^2 \mathcal{L}(\theta) > 0\}$  the set on which its Hessian is invertible. We assume that  $\Theta \subset \Theta_I$ , which is tantamount to assuming that the family is minimal, and ensures we only consider non-degenerate distributions ( $\nabla \mathcal{L}$  is one-to-one on its domain). We use notations  $\mathbb{P}_\theta, \mathbb{E}_\theta$  to explicitly refer to the probability and expectation associated to the distribution  $p_\theta$ .

**Bregman divergence of an exponential family.** A fundamental property of exponential families is the following form for the KL divergence between two distributions with parameters  $\theta, \theta' \in \Theta_D$ :

$$\text{KL}(p_\theta, p_{\theta'}) = \langle \theta - \theta', \mathbb{E}_\theta(F(X)) \rangle - \mathcal{L}(\theta) + \mathcal{L}(\theta').$$

Here,  $\mathbb{E}_\theta(F(X))$  is called the vector of *expectation parameters* (a.k.a. *dual parameters*), and is equal to  $\nabla \mathcal{L}(\theta)$ . Hence, it holds that  $\text{KL}(p_\theta, p_{\theta'}) = \mathcal{B}_\mathcal{L}(\theta', \theta)$ , where  $\mathcal{B}_\mathcal{L}$  is known as the Bregman divergence (Bregman, 1967) with potential function  $\mathcal{L}$ , defined by

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

**Tail and duality properties.** The canonical Bregman divergence of an exponential family enjoys two fundamental properties. The first one links it to the log-moment generating function of the random variable  $F(X)$ , which makes Bregman divergences well suited to control the tail behavior of random variables appearing in concentration inequalities. The second one highlights duality properties that enables convenient algebraic manipulations. To this end, for any  $\lambda \in \mathbb{R}^d$ , we define the function  $\mathcal{B}_{\mathcal{L}, \theta}(\lambda) = \mathcal{B}_\mathcal{L}(\theta + \lambda, \theta) = \mathcal{L}(\theta + \lambda) - \mathcal{L}(\theta) - \langle \lambda, \nabla \mathcal{L}(\theta) \rangle$ . Also, we introduce the Legendre-Fenchel dual operator  $\star$  associating a function  $G$  to its dual  $G^\star(x) = \sup_\lambda \langle \lambda, x \rangle - G(\lambda)$ .

**Lemma 1 (Properties of Bregman divergences)** *For all  $\theta \in \Theta_D$  and  $\lambda \in \mathbb{R}^d$  such that  $\theta + \lambda \in \Theta_D$ ,*

$$\log \mathbb{E}_\theta [\exp (\langle \lambda, F(X) - \mathbb{E}_\theta[F(X)] \rangle)] = \mathcal{B}_{\mathcal{L}, \theta}(\lambda).$$

*Furthermore, if  $\nabla \mathcal{L}$  is one-to-one, the following Bregman duality relations hold for any  $\theta, \theta' \in \Theta_D$ :*

$$\mathcal{B}_\mathcal{L}(\theta', \theta) = \mathcal{B}_{\mathcal{L}, \theta'}^\star(\nabla \mathcal{L}(\theta) - \nabla \mathcal{L}(\theta')) = \mathcal{B}_{\mathcal{L}^\star}(\nabla \mathcal{L}(\theta), \nabla \mathcal{L}(\theta')).$$

More generally, the following holds for any  $\alpha \in [0, 1]$  :

$$\mathcal{B}_{\mathcal{L}, \theta'}^*(\alpha(\nabla \mathcal{L}(\theta) - \nabla \mathcal{L}(\theta'))) = \mathcal{B}_{\mathcal{L}}(\theta', \theta_\alpha), \text{ where } \theta_\alpha = \nabla \mathcal{L}^{-1}(\alpha \nabla \mathcal{L}(\theta) + (1 - \alpha) \nabla L(\theta')).$$

The second half of this technical lemma is essentially a change of variable formula to move back and forth between two representations of an exponential family: in natural parameters (measured by the Bregman divergence between  $\theta'$  and  $\theta$ ) and in expectation parametrization (measured by the dual Bregman divergence between  $\nabla \mathcal{L}(\theta) = \mathbb{E}_\theta[F(X)]$  and  $\nabla \mathcal{L}(\theta') = \mathbb{E}_{\theta'}[F(X)]$ ). For more background on this, which forms the basis of the *information geometry* field, we refer to [Amari \(2016, Sections 2.1 and 2.7\)](#). This result is at the root of the martingale construction behind Theorem 3 in the next section. For completeness, the proof of this classical lemma is given in Appendix B.

### 3. Time-uniform Bregman Concentration

In this section, we are interested in controlling the deviation between a parameter  $\theta \in \Theta$  and its estimate  $\theta_n$  built from  $n$  observations from distribution  $p_\theta$ . We naturally measure this deviation in terms of the canonical Bregman divergence of the family. Further, we would like to control this deviation not only for a single sample number  $n$ , but *simultaneously for all  $n \in \mathbb{N}$* . Namely, we would like to upper bound quantities of the form  $\mathbb{P}[\exists n \in \mathbb{N} : \mathcal{B}_{\mathcal{L}}(\theta, \theta_n) \geq \dots]$ . Such a control is very useful in contexts when observations are gathered sequentially (either actively or otherwise), and especially when the number of observations  $n$  is unknown beforehand. Classical examples include multi-armed bandits ([Auer et al., 2002](#)), or model-based reinforcement learning ([Jaksch et al., 2010](#)).

#### 3.1. A generic deviation inequality

In this section, we consider the problem of controlling the Bregman-deviations of a parameter estimate of  $\theta$ . To this end, we adapt the method of mixtures (a.k.a. Laplace method) from ([Peña et al., 2008](#)). The method is originally designed in the context of Gaussian distributions, where it yields simple closed-form expressions, even though it can be applied more generally. We state below a generic extension of the method to parametric exponential families and introduce a quantity that measures a form of *information gain* about  $\theta$  after observing  $n$  samples, but expressed in terms of the natural Bregman divergence. For this reason, we call this quantity the *Bregman information gain*.

**Definition 2 (Bregman information gain)** *Let  $X_1, \dots, X_n \sim p_\theta$  be i.i.d. samples generated from  $p_\theta$ , where  $\theta \in \Theta \subset \Theta_I$ , and let  $\theta_0 \in \Theta$  be a reference parameter. For any constant  $c > 0$ , the Bregman information gain about  $\theta$  after observing  $X_1, \dots, X_n$  from  $\theta_0$  is defined as*

$$\gamma_{n,c}(\theta_0) = \log \left( \frac{\int_{\Theta} \exp(-c \mathcal{B}_{\mathcal{L}}(\theta', \theta_0)) d\theta'}{\int_{\Theta} \exp(-(n+c) \mathcal{B}_{\mathcal{L}}(\theta', \theta_{n,c}(\theta_0))) d\theta'} \right),$$

where  $\theta_{n,c}(\theta_0) = (\nabla \mathcal{L})^{-1} \left( \frac{\sum_{t=1}^n F(X_t) + c \nabla \mathcal{L}(\theta_0)}{n+c} \right)$  denotes a parameter estimate of  $\theta$ .<sup>1</sup>

---

1.  $\theta_{n,c}(\theta_0)$  is actually a maximum a posteriori estimate under a conjugate prior on  $\theta$ , depending on the reference point  $\theta_0$ .

**Dependence on  $\theta_0$  and example.** The acute reader can note that the considered parameter estimate  $\theta_{n,c}(\theta_0)$  and Bregman information gain  $\gamma_{n,c}(\theta_0)$  involve a reference parameter  $\theta_0$ . It makes sense to have such a local reference point since the Bregman divergence is typically linked to metrics with *local* (non-constant) curvature. Hence, the (information) geometry seen from the perspective of different points  $\theta_0$  may be different, unlike in the Gaussian case. Specifically, for a Gaussian  $\mathcal{N}(\mu, \sigma^2)$  with known variance  $\sigma^2$ , the Bregman information gain w.r.t. a reference point  $\mu_0$  reads

$$\gamma_{n,c}^{\mathcal{N}}(\mu_0) = \frac{1}{2} \log \frac{2\pi\sigma^2}{c} - \frac{1}{2} \log \frac{2\pi\sigma^2}{n+c} = \frac{1}{2} \log \frac{n+c}{c}.$$

It is independent of the reference parameter  $\mu_0$ , which is a consequence of the fact that the Bregman divergence in this case is proportional to the squared Euclidean distance; in other words, Gaussian distributions with known variance exhibit invariant geometry. However, for other exponential families, the geometries are inherently different, and hence Bregman information gain depend explicitly on local reference points (see Section 3.2 for details). For reference, the classical Gaussian information gain, i.e., the mutual information between a prior  $\mu \sim \mathcal{N}(\mu_0, \sigma^2)$  and the average of an i.i.d. sample  $X_1, \dots, X_n$  drawn from  $\mathcal{N}(\mu, c\sigma^2)$  is  $\frac{1}{2} \log \frac{n+c}{c}$ , which matches the Bregman information gain.

We now present the main result of this paper – a time uniform confidence bound for  $\theta$  connecting the Bregman divergence geometry of the exponential family with its Bregman information gain.

**Theorem 3 (Main result: Laplace method for generic exponential families)** Fix any  $\delta \in (0, 1]$  and  $n \in \mathbb{N}$ . Under the hypothesis of Definition 2, consider the confidence set

$$\Theta_{n,c}(\delta) = \left\{ \theta_0 \in \Theta : (n + c)\mathcal{B}_{\mathcal{L}}(\theta_0, \theta_{n,c}(\theta_0)) \leq \log \frac{1}{\delta} + \gamma_{n,c}(\theta_0) \right\}.$$

The following time-uniform control holds whenever the Bregman information gain is well-defined:

$$\mathbb{P}_{\theta} [\exists n \in \mathbb{N} : \theta \notin \Theta_{n,c}(\delta)] \leq \delta.$$

Note the implicit definition of the confidence set  $\Theta_{n,c}(\delta)$ , where the parameter of interest  $\theta_0$  appears in both arguments of the Bregman divergence, as well as in the Bregman information gain  $\gamma_{n,c}(\theta_0)$ . Because of this, computing this confidence set from the equation in Theorem 3 may seem non-trivial at first glance. However, we show in Section 3.2 how these sets simplify for many classical families, revealing how the computation can be made efficiently. Moreover, we observe that these confidence sets are actually tighter than those of prior work, and as such are especially well suited to be used when  $n$  is small (for large  $n$ , most methods produce essentially equivalent sets). Note that our concentration bound holds uniformly over all  $n$ . Equivalently (Howard et al., 2020, Lemma 3), it also holds that  $\mathbb{P}_{\theta} [\theta \notin \Theta_{\tau,c}(\delta)] \leq \delta$  for any random stopping time  $\tau$ .

**Comparison with prior work.** Similar to this work, Kaufmann and Koolen (2021) extend the method of mixtures technique to derive time-uniform concentration bounds for exponential families. However, their proof technique is fairly different, relying instead on discrete mixtures and stitching, which only works for single parameter families, and involves case-specific calculations that are difficult to generalize beyond Gaussian (with known variance) and Gamma (with known shape). In contrast, our method applies to generic exponential families, including distributions with more than one parameter such as Gaussian when both mean and variance are unknown. Moreover, their discrete prior construction leads to technical constants seemingly unrelated to the exponential family model.

Our prior is naturally induced by the exponential family, leveraging key properties of Bregman divergences, yields more intrinsic quantities (Bregman information gain) and perhaps a more elegant and shorter proof. Hence, our results are not only more general, but also of fundamental interest.

**Asymptotic behavior.** The asymptotic width of  $\Theta_{n,c}(\delta)$  depends on the behavior of  $\gamma_{n,c}(\theta_0)$  as  $n \rightarrow +\infty$ . Standard arguments show that  $\theta_{n,c}(\theta_0) \rightarrow \theta$  and the Taylor expansion of  $\theta' \mapsto \mathcal{B}_{\mathcal{L}}(\theta', \theta)$  around  $\theta' = \theta$  is  $\mathcal{B}_{\mathcal{L}}(\theta', \theta) = \frac{1}{2}(\theta' - \theta)^\top \nabla^2 \mathcal{L}(\theta)(\theta' - \theta) + o(\|\theta' - \theta\|^2)$  (note that  $\nabla^2 \mathcal{L}(\theta)$  is positive definite for  $\theta \in \Theta_I$ ). Laplace's method for integrals (chapter 20 in [Lattimore and Szepesvári \(2019\)](#), [Shun and McCullagh \(1995\)](#)) then gives the following estimate:

$$\int_{\Theta} \exp\left(-(n+c)\mathcal{B}_{\mathcal{L}}(\theta', \theta_{n,c}(\theta_0))\right) d\theta' \approx \int_{\mathbb{R}^d} \exp\left(-\frac{n+c}{2}(\theta - \theta')^\top \nabla^2 \mathcal{L}(\theta)(\theta - \theta')\right) d\theta',$$

which is a simple Gaussian integral, and thus  $\gamma_{n,c}(\theta_0) = \frac{d}{2} \log(1+n/c) + \mathcal{O}(1)$ . This asymptotic scaling is worse than the  $\log \log n$  rate of the law of iterated logarithm. However, this is a standard feature of the method of mixtures compared to stitching ([Maillard, 2019b](#); [Howard et al., 2020](#)), which is compensated by its improved nonasymptotic sharpness, as evidenced in Section 4.

**Dependence on the parameter  $c$ .** The parameter  $c$  is usually chosen to be 1 in the Gaussian case. However, it is useful to study its influence over the bounds. We provide in Section 4 a detailed study of this parameter, revealing first that the confidence bounds are not significantly altered over a large range of values, and then explaining how to pick an optimized value of  $c$ , e.g., for a specific horizon  $n = n_0$ . Note that choosing a variable  $c$  changing with  $n$  is not allowed by the theory, as this would break the martingale property used in our proof (see Section 3.3). Moreover, it would be incompatible with the time-uniform lower bound  $|\Theta_{n,c}(\delta)| = \Omega(\sqrt{\log \log n/n})$  provided by the law of iterated logarithm. Indeed, the Bregman information gain with  $c_n \propto n$  would be asymptotically  $\gamma_{n,c_n}(\theta_0) = \frac{d}{2} \log(1+\mathcal{O}(1)) + \mathcal{O}(1) = \mathcal{O}(1)$ , leading to  $|\Theta_{n,c}(\delta)| = \mathcal{O}(1/\sqrt{n})$  when  $n \rightarrow +\infty$ .

**Application to bandits.** One can apply the technique developed in proving Theorem 3 to build confidence sets in standard  $K$ -armed bandit problems. In such problems, we are given  $K$  distributions  $\{p_{\theta_i}\}_{i=1}^K$  from a generic exponential family (e.g., Gaussian) with parameters  $\{\theta_i\}_{i=1}^K$ , from which we can draw samples, interpreted as rewards we want to maximize. This setting is standard to analyze regret-optimal bandit algorithms ([Cappé et al., 2013](#); [Korda et al., 2013](#); [Baudry et al., 2020](#)). Specifically, at each time  $n$ , we choose an arm  $i_n \in [K]$  based on past observations, and draw a sample  $X_n$  from its distribution  $p_{\theta_{i_n}}$ . The samples are then used to update knowledge about the parameters  $\{\theta_i\}_{i=1}^K$  by building confidence sets for each of them. Let  $N_i(n)$  denote the number of times that we have chosen action  $i$  up to time  $n$ . Also, let  $\theta_{i,N_i(n),c}(\theta_0)$  and  $\gamma_{i,N_i(n),c}(\theta_0)$  denote the parameter estimate and Bregman information gain for arm  $i$ , respectively (similar to Definition 2 with  $n$  replaced by  $N_i(n)$ ). We construct the confidence set for arm  $i$  at time  $n$  as

$$\Theta_{i,n,c}(\delta) = \left\{ \theta_0 \in \Theta : (N_i(n) + c)\mathcal{B}_{\mathcal{L}}(\theta_0, \theta_{i,N_i(n),c}(\theta_0)) \leq \log \frac{1}{\delta} + \gamma_{i,N_i(n),c}(\theta_0) \right\}.$$

Then, similar to Theorem 3, it holds that the true parameter  $\theta_i$  lies in the set  $\Theta_{i,n,c}(\delta)$  for all time-steps  $n \in \mathbb{N}$  with probability at least  $1 - \delta$ . Finally, we take a union bound over  $i \in [K]$  to obtain confidence sets for all arms (with widths inflated by an additive  $\log K$  factor). Such construction is standard and is used in [Abbasi-Yadkori et al. \(2011\)](#) for (sub)-Gaussian families. Possible applications include UCB algorithms for regret minimization and designing GLR stopping rules for tracking algorithms

in pure exploration (Garivier and Kaufmann, 2016) in the context of generic exponential families (see also Section 5 for another application of GLR tests using the parameter estimate  $\theta_{n,c}(\theta_0)$ ).

**Remark 4** We provide in Appendix A.2 a complementary result (Corollary 8) using a Legendre function  $\mathcal{L}_0$  instead of the regularizing parameter  $c$ , which eschews the use of a local reference  $\theta_0$ . However, we argue that such a global regularization is actually less convenient to use except for univariate and multivariate Gaussian distributions that anyway exhibit invariant geometry. Furthermore, in Theorem 7, we prove a more general result that handles the case of a sequence  $(X_t)_{t \in \mathbb{N}}$  of random variables that are not independent, and having possibly different distributions from each other, which we apply to build confidence sets in linear bandits (see Appendix F).

### 3.2. Specification to classical families

In this section, we specify the result of Theorem 3 to some classical exponential families. Interestingly, the literature on time-uniform concentration bounds outside of random variables that are bounded, gamma with fixed shape or Gaussian with known variance is significantly scarce, even though many more distributions are commonly used in machine learning models. We derive below explicit confidence sets for a range of distributions, which we believe will be of interest for the wider machine learning and statistics community. For instance, consider active learning in bandit problems (Carpentier et al., 2011), where one targets upper confidence bounds on the variance (rather than the mean); in the Gaussian case, this can be achieved with Chi-square concentration. Hao et al. (2019) studies the classical UCB algorithm for bandits under a weaker assumption that sub-Gaussianity, involving the Weibull concentration. In differential privacy, concentration of Laplace distribution (symmetrized exponential) is often used to study the utility of differentially private mechanisms (Dwork et al., 2014). Finally, heavy-tailed distributions such as Pareto have recently been of interest to study risk-averse or corruption in bandit problems (Holland and Haress, 2021; Basu et al., 2022).

We now make explicit the Bregman information gains and confidence sets for some illustrative families (more examples and full derivations are provided in Appendix C).

**Gaussian (unknown mean and variance).** Let  $X \sim \mathcal{N}(\mu, \sigma^2)$ . Given samples  $X_1, \dots, X_n$ , we define  $S_n = \sum_{t=1}^n X_t$  and  $\hat{\mu}_n = S_n/n$ . Further, for  $\mu, \sigma \in \mathbb{R} \times \mathbb{R}^+$ , we define the normalized sum of squares  $Z_n(\mu, \sigma) = \frac{1}{\sigma^2} \sum_{t=1}^n (X_t - \mu)^2$ . Then, for reference parameters  $\mu_0 \in \mathbb{R}, \sigma_0 \in \mathbb{R}^+, c > 0$ , the Bregman information gain reads

$$\gamma_{n,c}^{\mathcal{N}}(\mu_0, \sigma_0) = \frac{3}{2} \log \left( \frac{n}{n+c} Z_n(\hat{\mu}_n, \sigma_0) + \frac{c}{n+c} Z_n(\mu_0, \sigma_0) + c \right) + f_{n,c},$$

where  $f_{n,c} = \left( \frac{n+c+1}{2} \right) \log(n+c) - \left( \frac{c}{2} + 2 \right) \log c - \frac{n}{2}(1 + \log 2) + \log \Gamma\left(\frac{c+3}{2}\right) - \log \Gamma\left(\frac{n+c+3}{2}\right)$ .

**Bernoulli.** Let  $X \sim \text{Bernoulli}(\mu)$ , with unknown mean  $\mu \in [0, 1]$ . Define, for reference parameter  $\mu_0 \in \mathbb{R}$ ,  $c > 0$ , the estimate  $\mu_{n,c}(\mu_0) = \frac{S_n + c\mu_0}{n+c}$ . Then, the Bregman information gain is given by

$$\gamma_{n,c}^{\text{Bernoulli}}(\mu_0) = c \mathbb{H}(\mu_0) - (n+c) \mathbb{H}(\mu_{n,c}(\mu_0)) + \log \frac{\mathbf{B}(c\mu_0, c(1-\mu_0))}{\mathbf{B}((n+c)\mu_{n,c}(\mu_0), (n+c)(1-\mu_{n,c}(\mu_0)))},$$

where  $\mathbf{B}(\alpha, \beta) = \int_0^1 u^{\alpha-1} (1-u)^{\beta-1} du$  is the Beta function and  $\mathbb{H}(\cdot)$  the Bernoulli entropy function.

**Exponential.** Let  $X \sim \text{Exp}(1/\mu)$ , with unknown mean  $\mu > 0$ . For reference parameters  $\mu_0, c > 0$ , the Bregman information gain reads

$$\gamma_{n,c}^{\text{Exp}}(\mu_0) = \log(S_n/\mu_0 + c) + \log \frac{\Gamma(c)}{\Gamma(n+c)} + (n+c-1) \log(n+c) - c \log c - n .$$

**Pareto.** Let  $X \sim \text{Pareto } (\alpha)$ , with unknown shape  $\alpha > 0$ . Define  $L_n = \sum_{t=1}^n \log X_t$ . Then, for reference parameters  $\alpha_0, c > 0$ , the Bregman information gain is given by

$$\gamma_{n,c}^{\text{Pareto}}(\alpha_0) = \log(\alpha_0 L_n + c) + \log \frac{\Gamma(c)}{\Gamma(n+c)} + (n+c-1) \log(n+c) - c \log c - n .$$

**Chi-square.** Let  $X \sim \chi^2(k)$ , where  $k \in \mathbb{N}$  is unknown. Define  $K_n = \sum_{t=1}^n \log \frac{X_t}{2}$ . For reference points  $k_0 \in \mathbb{N}, c > 0$ , introduce an estimate  $k_{n,c}(k_0)$  satisfying  $\psi_0\left(\frac{k_{n,c}(k_0)}{2}\right) = \frac{K_n + c \psi_0\left(\frac{k_0}{2}\right)}{n+c}$ . In this case, the Bregman information gain is given by

$$\begin{aligned} \gamma_{n,c}^{\chi^2}(k_0) &= \frac{k_{n,c}(k_0)}{2} \left( K_n + c \psi_0\left(\frac{k_0}{2}\right) \right) - (n+c) \log \Gamma\left(\frac{k_{n,c}(k_0)}{2}\right) + c \log \Gamma\left(\frac{k_0}{2}\right) - c \frac{k_0}{2} \psi_0\left(\frac{k_0}{2}\right) \\ &\quad + \log \frac{J(c, c \psi_0\left(\frac{k_0}{2}\right))}{J(n+c, K_n + c \psi_0\left(\frac{k_0}{2}\right))}, \text{ where } J(a, b) := \sum_{k'=1}^{\infty} \exp\left(-a \log \Gamma\left(\frac{k'}{2}\right) + b \frac{k'}{2}\right). \end{aligned}$$

The function  $J(a, b)$  can be estimated using numerical methods (see Section 4). This paves a way to build high-probability confidence sets for Chi-square distribution which has not been adequately captured in prior work. The sums over  $k'$  above derive from the martingale construction of Theorem 3 with discrete mixture (i.e., w.r.t the counting measure). A continuous (i.e., w.r.t the Lebesgue measure) mixture similar to the other families is also possible, as detailed in the appendix in Remark 9.

**Explicit confidence sets.** We now turn to illustrate the confidence sets in the same exponential families as above. They are obtained by specifying the generic form and simplifying the resulting expression. We provide the confidence sets for two-parameter Gaussian (i.e., unknown mean and variance), Bernoulli, Exponential, Pareto and Chi-square distributions, respectively, in Table 1. The technical details of the derivation of the specific forms for each illustrative family is postponed to Appendix C, along with other distributions (Gamma, Poisson, Weibull) in Table 2.

We now provide a set of illustrative numerical experiments to display the confidence envelopes resulting from Theorem 3 in the case of classical exponential families. We plot for a given value of the confidence level  $\delta$  and regularization parameter  $c$  the convex sets  $n \mapsto \bigcap_{n' \leq n} \Theta_{n',c}(\delta)$  (taking the running intersection is standard in sequential testing and was pioneered in [Darling and Robbins \(1967\)](#); it ensures that the upper (resp. lower) confidence envelopes are nonincreasing (resp. nondecreasing) as the sample size grows, thus providing tighter bounds). In dimension  $d = 1$ , we report the extremal points of these intervals, which we call the upper and lower envelopes respectively. We refer to Figure 1 for Pareto, Chi-square and Gaussian (with unknown  $\mu, \sigma$ ) and Appendix D for many other families. Because we exploit the Bregman geometry, our bounds capture a larger setting than typical mean estimation; for instance, we are able to concentrate around the exponent  $\alpha$  of a Pareto distribution even when the distribution is not integrable ( $\alpha < 1$ ). Furthermore, for two-parameter Gaussian, apart from being *anytime*, our confidence sets are convex and bounded in contrast to the one based on Chi-square quantiles with a crude union bound (see Appendix D.1).

Table 1: Bregman confidence sets given by Theorem 3 for representative families

Distribution	Parameters	Confidence Set
Gaussian	$\mu \in \mathbb{R}$ $\sigma \in \mathbb{R}_+$	$\begin{aligned} & \frac{1}{2}Z_n(\mu, \sigma) - \frac{n+c+3}{2} \log \left( \frac{n}{n+c} Z_n(\hat{\mu}_n, \sigma) + \frac{c}{n+c} Z_n(\mu, \sigma) + c \right) \\ & \leq \log \frac{1}{\delta} - \frac{n}{2} \log 2 - \left( \frac{c}{2} + 2 \right) \log c + \frac{1}{2} \log(n+c) + \log \frac{\Gamma(\frac{c+3}{2})}{\Gamma(\frac{n+c+3}{2})} \end{aligned}$
Bernoulli	$\mu \in [0, 1]$	$\begin{aligned} & S_n \log \frac{1}{\mu} + (n-S_n) \log \frac{1}{1-\mu} + \log \frac{\Gamma(S_n+c\mu)\Gamma(n-S_n+c(1-\mu))}{\Gamma(c\mu)\Gamma(c(1-\mu))} \\ & \leq \log \frac{1}{\delta} + \log \frac{\Gamma(n+c)}{\Gamma(c)} \end{aligned}$
Exponential	$\mu \in \mathbb{R}_+$	$\begin{aligned} & \frac{S_n}{\mu} - (n+c+1) \log \left( \frac{S_n}{\mu} + c \right) \\ & \leq \log \frac{1}{\delta} + \log \frac{\Gamma(c)}{\Gamma(n+c)} - \log(n+c) - c \log c \end{aligned}$
Pareto	$\alpha \in \mathbb{R}$	$\begin{aligned} & \alpha L_n - (n+c+1) \log(\alpha L_n + c) \\ & \leq \log \frac{1}{\delta} + \log \frac{\Gamma(c)}{\Gamma(n+c)} - \log(n+c) - c \log c \end{aligned}$
Chi-square	$k \in \mathbb{N}$	$\begin{aligned} & n \log \Gamma\left(\frac{k}{2}\right) - \frac{k}{2} K_n - \log J\left(c, c\psi_0\left(\frac{k}{2}\right)\right) \\ & + \log J\left(n+c, K_n + c\psi_0\left(\frac{k}{2}\right)\right) \leq \log \frac{1}{\delta} \end{aligned}$

### 3.3. Proof Sketch: Theorem 3

We now sketch the proof of Theorem 3, and refer the interested reader to the appendix for details.

**Step 1: Martingale construction.** For any  $\lambda \in \mathbb{R}^d$ , we introduce the quantity

$$M_n^\lambda = \exp(\langle \lambda, n(\mu_n - \mu) \rangle - n\mathcal{B}_{\mathcal{L}, \theta}(\lambda)),$$

where  $\mu_n = \frac{1}{n} \sum_{t=1}^n F(X_t)$  and  $\mu = \mathbb{E}_\theta[F(X)]$ . By Lemma 1,  $M_n^\lambda$  is a martingale such that  $\mathbb{E}[M_n^\lambda] = 1$ . We now introduce the distribution  $q(\theta|\alpha, \beta) = \exp(\langle \theta, \alpha \rangle - \beta \mathcal{L}(\theta)) H(\alpha, \beta)$  where  $H(\cdot, \cdot)$  is the normalization function, and define the mixture martingale  $M_n = \int M_n^\lambda q(\theta + \lambda|\alpha, \beta) d\lambda$ . Note that  $M_n$  also satisfies  $\mathbb{E}[M_n] = 1$ .

**Step 2. Choice of parameters and duality properties.** Choosing  $\alpha = c\nabla \mathcal{L}(\theta)$  and  $\beta = c$  yields

$$M_n = G(\theta, c) \int \exp(\langle \lambda, n(\mu_n - \mu) \rangle - (n+c)\mathcal{B}_{\mathcal{L}, \theta}(\lambda)) d\lambda,$$

where  $G(\theta, c) = 1 / \int \exp(-c\mathcal{B}_{\mathcal{L}}(\theta', \theta)) d\theta'$ . Now, we consider the function  $\mathcal{B}_{\mathcal{L}, \theta}^*(x) = \max_\lambda \langle \lambda, x \rangle - \mathcal{B}_{\mathcal{L}, \theta}(\lambda)$ . Note that its maximal point  $\lambda_x^*$  satisfies  $x + \nabla \mathcal{L}(\theta) = \nabla \mathcal{L}(\theta + \lambda_x^*)$  for every  $x$ . In particular, for the choice  $x = \frac{n}{n+c}(\mu_n - \mu)$ , we have  $\frac{n}{n+c}\mu_n + \frac{c}{n+c}\nabla \mathcal{L}(\theta) = \nabla \mathcal{L}(\theta + \lambda_x^*)$ , which justifies introducing the following regularized estimate

$$\theta_{n,c}(\theta) = (\nabla \mathcal{L})^{-1} \left( \frac{n}{n+c}\mu_n + \frac{c}{n+c}\nabla \mathcal{L}(\theta) \right),$$

provided that  $\nabla \mathcal{L}$  is one-to-one in  $\Theta$ .

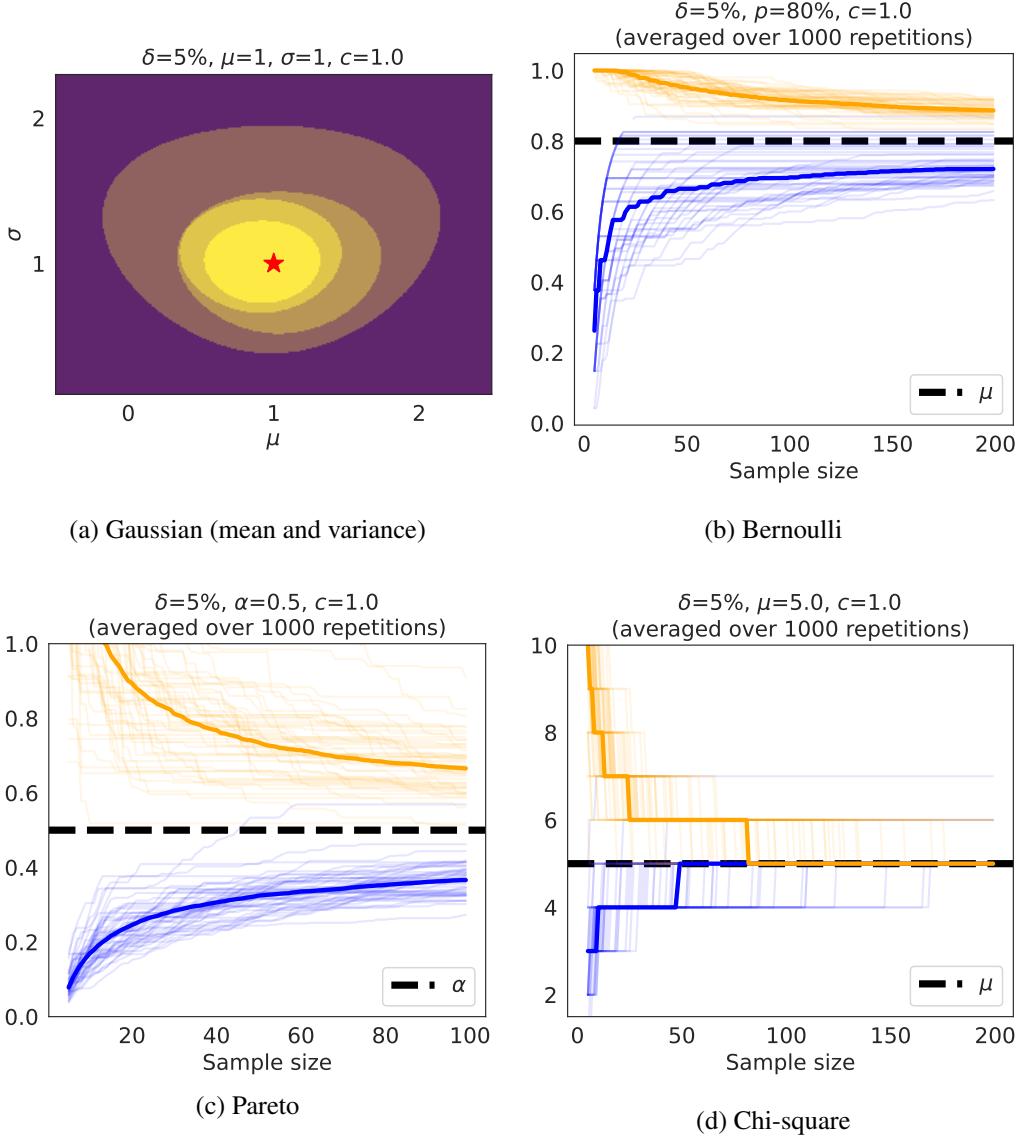


Figure 1: Example of Gaussian time-uniform joint confidence sets for  $(\mu, \sigma) = (1, 1)$  with sample sizes  $n \in \{10, 25, 50, 100\}$  observations (smaller confidence sets correspond to larger sample sizes), and examples of time-uniform confidence envelopes for Bernoulli(0.8), Pareto(0.5) and  $\chi^2(5)$  on several realizations as a function of the number of observations  $n$ . Thick lines indicate the median curve over 1000 replicates.

**Step 3. Martingale rewriting and conclusion.** Note that the duality property also yields  $\langle \lambda, x \rangle - \mathcal{B}_{\mathcal{L},\theta}(\lambda) - \langle \lambda_x^*, x \rangle + \mathcal{B}_{\mathcal{L},\theta}(\lambda_x^*) = \langle \lambda - \lambda_x^*, \nabla \mathcal{L}(\theta + \lambda_x^*) \rangle + \mathcal{L}(\theta + \lambda_x^*) - \mathcal{L}(\theta + \lambda)$  for each  $x$ . Using this property, we obtain

$$M_n = \exp((n+c)\mathcal{B}_{\mathcal{L},\theta}^*(x)) \frac{G(\theta, c)}{G(\theta_{n,c}(\theta), n+c)}.$$

After applying Markov's inequality, we obtain the following inequality:

$$\mathbb{P}_\theta \left[ \mathcal{B}_{\mathcal{L},\theta}^* \left( \frac{n}{n+c} (\mu_n - \mu) \right) \geq \frac{1}{n+c} \log \frac{G(\theta_{n,c}(\theta), n+c)}{\delta G(\theta, c)} \right] \leq \delta.$$

Now, we use duality property of Bregman divergence to obtain  $\mathcal{B}_{\mathcal{L},\theta}^* \left( \frac{n}{n+c} (\mu_n - \mu) \right) = \mathcal{B}_{\mathcal{L}}(\theta, \theta_{n,c}(\theta))$ , which yields the desired form. Finally, by properties of the martingale  $M_n$ , the deviation bound also holds for any random stopping time  $N$ , and hence we obtain the time uniform bound by employing a stopping time construction similar to that of Peña et al. (2008); Abbasi-Yadkori et al. (2011).

#### 4. Numerical Experiments

In this section, we provide illustrative numerical results to show the resulting confidence bands built from Theorem 3 are promising alternatives to existing competitors (detailed in Appendix D). All the confidence sets presented here are implemented in the open source *concentration-lib* Python package (<https://pypi.org/project/concentration-lib/>).

**Comparison.** We compare our confidence envelopes to state-of-the-art time-uniform bounds in Figure 2 and Appendix D. Interestingly, our upper and lower bounds are not necessarily symmetrical, and in the Bernoulli case, they fit within the distribution support without clipping, contrary to most other methods, thus adapting to the local geometry of the family. Our bounds are comparable to those of Kaufmann and Koolen (2021), if slightly tighter; however, we emphasize again that our scope is wider and captures more distributions.

**Numerical complexity.** Some of our formulas (Poisson, Chi-square) require the evaluation of integrals for which no closed-form expression exist. However, these can be estimated up to arbitrary precision by numerical routines, see Remarks 10 and 11 for details. Similarly, the use of special functions (digamma, ratio of Gamma) may lead to numerical instabilities or overflows for large  $n$ ; we recommend instead to use the log-Gamma function and an efficient implementation of the log-sum-exp operator. We report results for  $n \leq 200$  as we think the small sample regime is where our bounds shine (most reasonable methods produce similar confidence sets for large  $n$ ). Finally, most confidence sets reported in Figure 2 are implicitly defined (after reorganizing terms) as level sets  $\{F(\theta) \leq 0\}$  for some function  $F$ . For one-dimensional families, we use a root search routine for a fast estimation of the boundaries of such sets; in higher dimension, we evaluate  $F$  on a uniform grid (e.g, in Figure 1a, we use a  $1024 \times 1024$  grid over  $[-2, 4] \times [0.1, 4]$ ).

**Influence of the regularization parameter  $c$ .** Here, we fix  $\delta = 0.05$  (note that all bounds exhibit the typical  $\log \frac{1}{\delta}$  dependency). The parameter  $c$  is fixed to  $c = 1$  and can be interpreted as a number of virtual prior samples, thus introducing a bias. The choice  $c = 1$  is a typical value considered in literature on Gaussian concentration. To see the influence of the parameter  $c$  on the bounds, we provide a second set of experiments, in which we compute confidence envelopes for  $c$  varying between 0.1 and 30 (Figure 2d). We find that  $c$  has limited influence on the width of the confidence interval. To get more insight on the tuning of  $c$ , we perform additional experiments in Appendix D, where we optimize the confidence width w.r.t. to  $c$  for varying sample sizes  $n$ . The resulting choice  $c_n^* = \text{argmin}_c |\Theta_{n,c}(\delta)|$  exhibits a linear trend  $c_n^* \approx 0.12n$ , which seems consistent across the tested distributions. Of note, the constant 0.12 is reminiscent of the optimal tuning of the sub-Gaussian Laplace method (Howard et al., 2020). This is consistent with the observation in Section 3.1 that

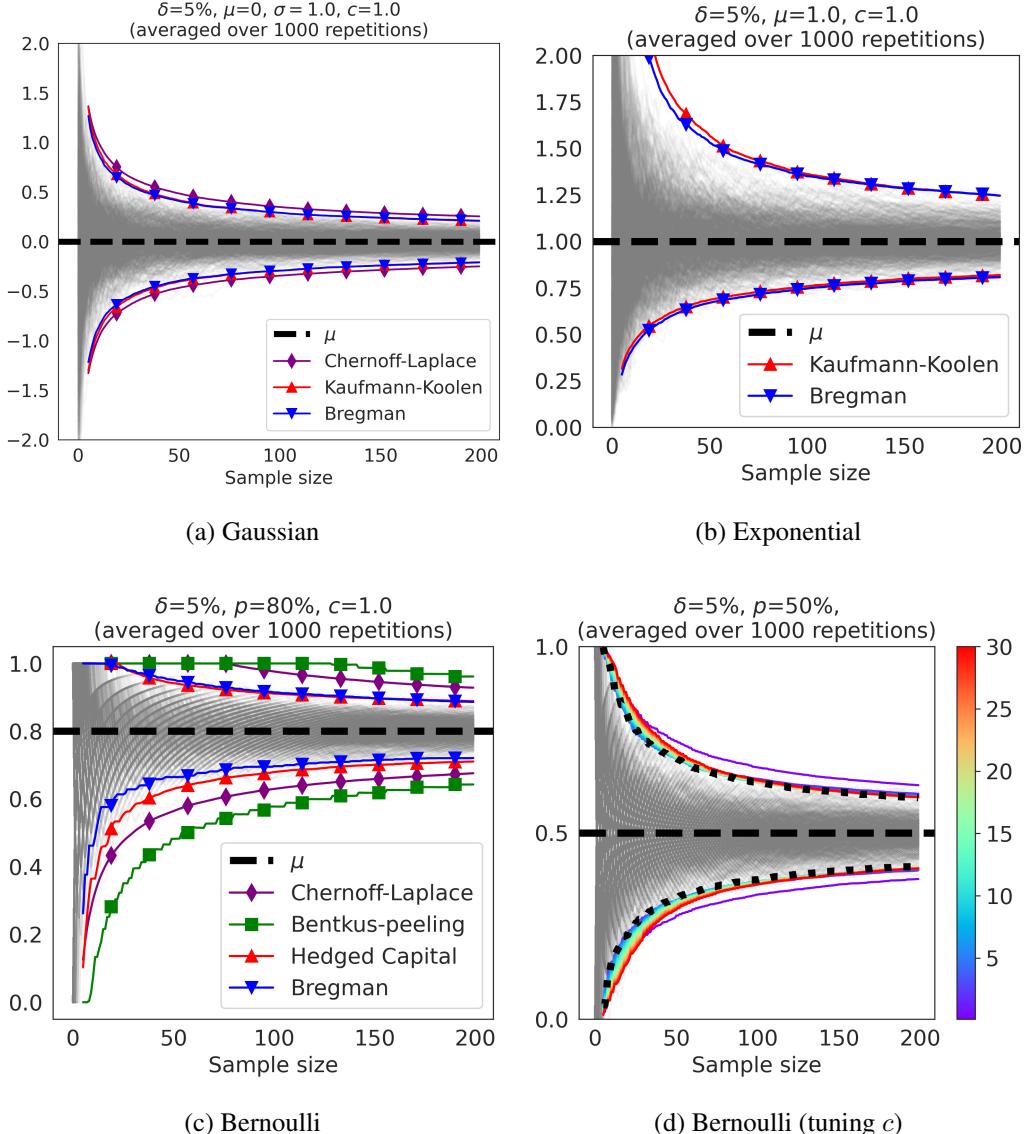


Figure 2: Comparison of median confidence envelopes around the mean for  $\mathcal{N}(0, 1)$ , Bernoulli(0.8) and Exp(1), as a function of the sample size  $n$ , over 1000 independent replicates. Grey lines are trajectories of empirical means  $\hat{\mu}_n$ . Bottom right: confidence envelopes for varying regularization  $c \in [0.1, 30]$  for Bernoulli(0.5) distribution. Dotted black line: heuristic  $c_n \approx 0.12n$ .

the confidence width  $|\Theta_{n,c}(\delta)|$  is equivalent to the Gaussian confidence width when  $n \rightarrow +\infty$ . We report in Figure 2d this heuristic for  $c$ . However, we restate that sample a size-dependent regularization parameter  $c$  is *not* supported by the theory (as it would violate the law of iterated logarithms). Now, this does not prevent from choosing a specific horizon of interest  $n_0$  and set  $c = 0.12n_0$  to promote sharpness around  $n = zn_0$ .

## 5. Application: Generalized Likelihood Ratio Test in Exponential Families

In this section, we apply our result to control the false alarm rate of Generalized Likelihood Ratio (GLR) tests when detecting a change of measure from a sequence of observations. The GLR test in the exponential family model  $\mathcal{E}$  is defined, for a threshold  $\alpha > 0$ , as

$$\tau(\alpha; \mathcal{E}) = \min\{t \in \mathbb{N} : \max_{s \in [0, t)} G_{1:s:t}^{\mathcal{E}} \geq \alpha\},$$

where the GLR is defined to be  $G_{1:s:t}^{\mathcal{E}} = \inf_{\theta'} \sup_{\theta_1, \theta_2} \log \left( \frac{\prod_{t'=1}^s p_{\theta_1}(X_{t'}) \prod_{t'=s+1}^t p_{\theta_2}(X_{t'})}{\prod_{t'=1}^t p_{\theta'}(X_{t'})} \right)$ .

Hereafter, we consider that all the observations  $(X_{t'})_{t' \in \mathbb{N}}$  come from a distribution with same parameter  $\theta \in \Theta$ . Then, the false alarm rate of the GLR test can be bounded using

$$\mathbb{P}_{\theta}(\tau(\alpha; \mathcal{E}) < \infty) = \mathbb{P}_{\theta}(\exists (t, s) \in \mathbb{N}^2, s < t : G_{1:s:t}^{\mathcal{E}} \geq \alpha).$$

Observe that solving for the inner supremums in the expression of GLR, we obtain

$$G_{1:s:t}^{\mathcal{E}} = \inf_{\theta'} s \mathcal{B}_{\mathcal{L}}(\theta', \theta_{1:s}) + (t - s) \mathcal{B}_{\mathcal{L}}(\theta', \theta_{s+1:t}),$$

where we introduce parameter estimates  $\theta_{1:s}$  and  $\theta_{s+1:t}$  such that  $\nabla \mathcal{L}(\theta_{1:s}) = \frac{1}{s} \sum_{t'=1}^s F(X_{t'})$  and  $\nabla \mathcal{L}(\theta_{s+1:t}) = \frac{1}{t-s} \sum_{t'=s+1}^t F(X_{t'})$ , respectively. Hence, the false alarm rate can be controlled as

$$\mathbb{P}_{\theta}(\tau(\alpha; \mathcal{E}) < \infty) \leq \mathbb{P}_{\theta}\left(\exists s \in \mathbb{N} : s \mathcal{B}_{\mathcal{L}}(\theta, \theta_{1:s}) \geq \alpha_1\right) + \mathbb{P}_{\theta}\left(\exists (t, s) \in \mathbb{N}^2, s < t : (t - s) \mathcal{B}_{\mathcal{L}}(\theta, \theta_{s+1:t}) \geq \alpha_2\right)$$

for appropriate terms  $\alpha_1, \alpha_2$  such that  $\alpha = \alpha_1 + \alpha_2$ . The control of the first term comes naturally from our time-uniform deviation result (Theorem 3) using a regularized version of the estimate  $\theta_{1:s}$ . For the second term, we need to study the concentration of  $\mathcal{B}_{\mathcal{L}}(\theta, \theta_{s+1:t})$  uniformly over  $s$  and  $t$ .

**Doubly time-uniform concentration.** For any  $s < t$ ,  $c > 0$  and reference parameter  $\theta_0 \in \Theta$ , we define the regularized estimate  $\theta_{s+1:t,c}(\theta_0) = (\nabla \mathcal{L})^{-1} \left( \frac{1}{t-s+c} \left( \sum_{t'=s+1}^t F(X_{t'}) + c \nabla \mathcal{L}(\theta_0) \right) \right)$  built from  $(t-s)$  observations  $X_{s+1}, \dots, X_t$ . Similarly, we introduce the corresponding Bregman information gain  $\gamma_{s+1:t,c}(\theta_0)$  as in Definition 2. The following result extends the Laplace method for exponential families (Theorem 3) to control the Bregman deviation of  $\theta_{s+1:t,c}(\theta_0)$  around  $\theta_0$ .

**Theorem 5 (Doubly time-uniform concentration)** *Let  $\delta \in (0, 1]$  and  $g: \mathbb{N} \rightarrow \mathbb{R}_+^*$  such that  $\sum_{t=1}^{\infty} 1/g(t) \leq 1$  (e.g.,  $g(t) = \kappa(1+t) \log^2(1+t)$  where  $\kappa = 2.10975$ ). Then, under the assumptions of Definition 2, it holds that*

$$\mathbb{P}_{\theta} \left[ \exists t \in \mathbb{N}, \exists s < t : (t - s + c) \mathcal{B}_{\mathcal{L}}(\theta, \theta_{s+1:t,c}(\theta)) \geq \log \frac{g(t)}{\delta} + \gamma_{s+1:t,c}(\theta) \right] \leq \delta.$$

**Remark 6** *Maillard (2019a) proves a doubly uniform concentration inequality for means of (sub)-Gaussian random variables and conjectures that it could be extended to other types of changes, such as changes of variance in a Gaussian family. Theorem 5 can be seen as a generalization of this result to generic exponential families. This naturally comes with additional challenges due to the intrinsic local geometry of exponential families. The proof of this result is deferred to Appendix E.*

**A regularized GLR test.** For a given false alarm (i.e., false positive) probability  $\delta \in (0, 1]$  and parameter  $c > 0$ , we define the regularized GLR test  $\tau_c^\delta(\mathcal{E})$  as follows:

$$\begin{aligned}\Theta_{s,c}(\delta) &= \{\theta \in \Theta : (s + c)\mathcal{B}_{\mathcal{L}}(\theta, \theta_{1:s,c}(\theta)) \leq \log(1/\delta) + \gamma_{1:s,c}(\theta)\} , \\ \Theta_{s+1:t,c}(\delta) &= \{\theta \in \Theta : (t - s + c)\mathcal{B}_{\mathcal{L}}(\theta, \theta_{s+1:t,c}(\theta)) \leq \log(g(t)/\delta) + \gamma_{s+1:t,c}(\theta)\} , \\ \tau_c^\delta(\mathcal{E}) &= \min \{t \in \mathbb{N} : \exists s < t, \Theta_{s,c}(\delta/2) \cap \Theta_{s+1:t,c}(\delta/2) = \emptyset\} .\end{aligned}$$

Note that, under the measure  $\mathbb{P}_\theta$ ,  $\theta \in \Theta_{s,c}(\delta/2)$  holds with probability at least  $1 - \delta/2$  by time-uniform concentration over  $s$  (Theorem 3), and  $\theta \in \Theta_{s+1:t,c}(\delta/2)$  holds with probability at least  $1 - \delta/2$  by doubly time-uniform concentration over both  $s$  and  $t$  (Theorem 5). Then, by a union-bound, this test is guaranteed to have a false alarm probability controlled by  $\delta$ . The factor  $g(t)$  is essentially a function growing slightly faster than linearly; other common choices include  $g(t) \propto t^\eta$  for some  $\eta > 1$ , though it leads to a looser bound due to the faster growth. Furthermore, if one is interested in detecting changes up to a known horizon  $T$ , one can replace the function  $g(t)$  by the slightly tighter factor  $g(t) = \kappa(1+t) \log^2(1+t)$  with  $\kappa = \sum_{t=1}^T 1/g(t)$  and still guarantee a false alarm rate under  $\delta$ . In Appendix E.3, we report simulations of this GLR test to detect changes of variance in the family of centered Gaussian (Figure 14), which answers an open question from [Maillard \(2019a\)](#).

## 6. Conclusion

We apply the method of mixture to derive a time-uniform deviation inequality for generic parametric exponential families expressed in terms of their Bregman divergences, highlighting the role of a quantity, the *Bregman information gain*, that is related to the geometry of the family. We specialize this general result to build confidence sets for classical examples. Our method compares favorably to the state-of-the art for Gaussian, Bernoulli and exponential distributions, and extends to other families like Chi-square, Pareto and two-parameter Gaussian, where no known time-uniform bound exists. Our method is also general enough to design GLR tests for change detection. An interesting direction for future work would be to consider the case where the log-partition function (and thus the Bregman divergence) is misspecified or unknown and has to be learned alongside the parameter.

## Acknowledgments

The authors acknowledge the funding of the French National Research Agency, the French Ministry of Higher Education and Research, Inria, the MEL and the I-Site ULNE regarding project R-PILOTE-19-004-APPRENF and Bandits For Health (B4H). We thank the anonymous reviewers for their careful reading of the paper and their suggestions for improvements. Experiments presented in this paper were carried out using the Grid'5000 testbed, supported by a scientific interest group hosted by Inria and including CNRS, RENATER and several universities as well as other organizations (see <https://www.grid5000.fr>).

## References

- Yasin Abbasi-Yadkori, Dávid Pál, and Csaba Szepesvári. Improved algorithms for linear stochastic bandits. In *Advances in Neural Information Processing Systems*, pages 2312–2320, 2011.
- Shun-ichi Amari. *Information geometry and its applications*, volume 194. Springer, 2016.
- P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time Analysis of the Multi-armed Bandit Problem. *Machine Learning*, 47(2):235–256, 2002.
- Debabrota Basu, Odalric-Ambrym Maillard, and Timothée Mathieu. Bandits corrupted by nature: Lower bounds on regret and robust optimistic algorithm. *arXiv preprint arXiv:2203.03186*, 2022.
- Dorian Baudry, Emilie Kaufmann, and Odalric-Ambrym Maillard. Sub-sampling for efficient non-parametric bandit exploration. *Advances in Neural Information Processing Systems*, 33:5468–5478, 2020.
- Vidmantas Bentkus. On hoeffding’s inequalities. *The Annals of Probability*, 32(2):1650–1673, 2004.
- Daniel Berend and Aryeh Kontorovich. On the concentration of the missing mass. *Electronic Communications in Probability*, 18:1–7, 2013.
- Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.
- L.M. Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR Computational Mathematics and Mathematical Physics*, 7(3):200–217, 1967. ISSN 0041-5553.
- Sébastien Bubeck. *Bandits games and clustering foundations*. PhD thesis, 2010.
- Olivier Cappé, Aurélien Garivier, Odalric-Ambrym Maillard, Rémi Munos, and Gilles Stoltz. Kullback-leibler upper confidence bounds for optimal sequential allocation. *The Annals of Statistics*, pages 1516–1541, 2013.
- Alexandra Carpentier, Alessandro Lazaric, Mohammad Ghavamzadeh, Rémi Munos, and Peter Auer. Upper-confidence-bound algorithms for active learning in multi-armed bandits. In *International Conference on Algorithmic Learning Theory*, pages 189–203. Springer, 2011.
- Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.
- Sayak Ray Chowdhury and Aditya Gopalan. On kernelized multi-armed bandits. In *International Conference on Machine Learning*, pages 844–853. PMLR, 2017.
- Sayak Ray Chowdhury, Aditya Gopalan, and Odalric-Ambrym Maillard. Reinforcement learning in parametric mdps with exponential families. In *International Conference on Artificial Intelligence and Statistics*, pages 1855–1863. PMLR, 2021.
- DA Darling and Herbert Robbins. Iterated logarithm inequalities. *Proceedings of the National Academy of Sciences*, 57(5):1188–1192, 1967.

- Audrey Durand, Odalric-Ambrym Maillard, and Joelle Pineau. Streaming kernel regression with provably adaptive mean, variance, and regularization. *Journal of Machine Learning Research*, 19(1):650–683, 2018.
- Rick Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019.
- Cynthia Dwork, Aaron Roth, et al. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407, 2014.
- Louis Faurie, Marc Abeille, Clément Calauzènes, and Olivier Fercoq. Improved optimistic algorithms for logistic bandits. In *International Conference on Machine Learning*, pages 3052–3060. PMLR, 2020.
- Aurélien Garivier and Emilie Kaufmann. Optimal best arm identification with fixed confidence. In *Conference on Learning Theory*, pages 998–1027. PMLR, 2016.
- Aurélien Garivier. Informational confidence bounds for self-normalized averages and applications. In *2013 IEEE Information Theory Workshop (ITW)*, pages 1–5, 2013. doi: 10.1109/ITW.2013.6691311.
- Botao Hao, Yasin Abbasi Yadkori, Zheng Wen, and Guang Cheng. Bootstrapping upper confidence bound. *Advances in Neural Information Processing Systems*, 32:12123–12133, 2019.
- Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. ISSN 01621459.
- Matthew Holland and El Mehdi Haress. Learning with risk-averse feedback under potentially heavy tails. In *International Conference on Artificial Intelligence and Statistics*, pages 892–900. PMLR, 2021.
- Steven R Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform chernoff bounds via nonnegative supermartingales. *Probability Surveys*, 17:257–317, 2020.
- Steven R Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform, nonparametric, nonasymptotic confidence sequences. *The Annals of Statistics*, 2021.
- Thomas Jaksch, Ronald Ortner, and Peter Auer. Near-optimal regret bounds for reinforcement learning. *Journal of Machine Learning Research*, 11(4), 2010.
- Emilie Kaufmann and Wouter M Koolen. Mixture martingales revisited with applications to sequential tests and confidence intervals. *Journal of Machine Learning Research*, 22:246–1, 2021.
- Michael Kearns and Lawrence Saul. Large deviation methods for approximate probabilistic inference. In *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*, pages 311–319, 1998.
- Johannes Kirschner and Andreas Krause. Information directed sampling and bandits with heteroscedastic noise. In *Conference On Learning Theory*, pages 358–384. PMLR, 2018.
- Nathaniel Korda, Emilie Kaufmann, and Remi Munos. Thompson sampling for 1-dimensional exponential family bandits. *Advances in neural information processing systems*, 26, 2013.

- Arun K Kuchibhotla and Qinqing Zheng. Near-optimal confidence sequences for bounded random variables. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 5827–5837. PMLR, 18–24 Jul 2021.
- T. Lattimore and C. Szepesvári. *Bandit Algorithms*. Cambridge University Press, 2019.
- Beatrice Laurent and Pascal Massart. Adaptive estimation of a quadratic functional by model selection. *Annals of Statistics*, pages 1302–1338, 2000.
- O.-A. Maillard. Sequential change-point detection: Laplace concentration of scan statistics and non-asymptotic delay bounds. In *Algorithmic Learning Theory (ALT)*, 2019a.
- Odalric-Ambrym Maillard. *Mathematics of Statistical Sequential Decision Making*. Habilitation à diriger des recherches, Université de Lille Nord de France, February 2019b.
- Victor H Peña, Tze Leung Lai, and Qi-Man Shao. *Self-normalized processes: Limit theory and Statistical Applications*. Springer Science & Business Media, 2008.
- Maxim Raginsky and Igal Sason. *Concentration of Measure Inequalities in Information Theory, Communications, and Coding*. Now Publishers, 2018.
- Herbert Robbins. Statistical methods related to the law of the iterated logarithm. *The Annals of Mathematical Statistics*, 41(5):1397–1409, 1970.
- Herbert Robbins and EJG Pitman. Application of the method of mixtures to quadratic forms in normal variates. *The annals of mathematical statistics*, pages 552–560, 1949.
- Glenn Shafer and Vladimir Vovk. *Game-Theoretic Foundations for Probability and Finance*, volume 455. John Wiley & Sons, 2019.
- Zhenming Shun and Peter McCullagh. Laplace approximation of high dimensional integrals. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(4):749–760, 1995.
- Jean Ville. Etude critique de la notion de collectif. *Bull. Amer. Math. Soc*, 45(11):824, 1939.
- Ian Waudby-Smith and Aaditya Ramdas. Estimating means of bounded random variables by betting. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2023.
- A Dembo O Zeitouni and A Dembo. Large deviations techniques and applications. *Applications of Mathematics*, 38, 1998.

## Appendix A. Proof of the Main Result About Bregman Deviations

In this section, we detail the proof of the main results regarding Bregman deviation inequalities, stated with its two variants Theorem 3 and 7. We start in Section A.1 with the proof of Theorem 3 when considering a numerical constant  $c$  to build the regularizer. In Section A.2, we then provide the proof for Theorem 7 when considering instead a Legendre function  $\mathcal{L}_0$  as a regularizer, and also extend to the non i.i.d. case.

### A.1. Proof of Bregman concentration using mixture parameter $c$

The proof of this first result follows an adaptation of the method of mixture, combined with properties of Bregman divergences recalled in Lemma 1.

**Theorem 3 (Main result: Laplace method for generic exponential families)** *Fix any  $\delta \in (0, 1]$  and  $n \in \mathbb{N}$ . Under the hypothesis of Definition 2, consider the confidence set*

$$\Theta_{n,c}(\delta) = \left\{ \theta_0 \in \Theta : (n + c)\mathcal{B}_{\mathcal{L}}(\theta_0, \theta_{n,c}(\theta_0)) \leq \log \frac{1}{\delta} + \gamma_{n,c}(\theta_0) \right\}.$$

The following time-uniform control holds whenever the Bregman information gain is well-defined:

$$\mathbb{P}_{\theta} [\exists n \in \mathbb{N} : \theta \notin \Theta_{n,c}(\delta)] \leq \delta.$$

#### Proof of Theorem 3:

##### Step 1. Martingale and mixture martingale construction.

Let us note that  $\mathbb{E}_{\theta}[F(X)] = \nabla \mathcal{L}(\theta)$  and  $\log \mathbb{E}_{\theta}[\exp \langle \lambda, F(X) \rangle] = \mathcal{L}(\theta + \lambda) - \mathcal{L}(\theta)$ . Hence, we deduce that

$$\log \mathbb{E}_{\theta}[\exp \langle \lambda, F(X) - \mathbb{E}_{\theta}[F(X)] \rangle] = \mathcal{B}_{\mathcal{L},\theta}(\lambda) \stackrel{\text{def}}{=} \mathcal{L}(\theta + \lambda) - \mathcal{L}(\theta) - \langle \lambda, \nabla \mathcal{L}(\theta) \rangle.$$

Now, let  $\mu_n = \frac{1}{n} \sum_{t=1}^n F(X_t)$  and  $\mu = \mathbb{E}_{\theta}[F(X)]$ . For any  $\lambda \in \mathbb{R}^d$ , the following quantity

$$M_n^{\lambda} = \exp \left( \langle \lambda, n(\mu_n - \mu) \rangle - n\mathcal{B}_{\mathcal{L},\theta}(\lambda) \right),$$

is thus a martingale such that  $\mathbb{E}[M_n^{\lambda}] = 1$ .

We now introduce the distribution  $q(\theta|\alpha, \beta) = H(\alpha, \beta) \cdot \exp(\langle \theta, \alpha \rangle - \beta \mathcal{L}(\theta))$  where  $H$  is the normalization term. We further introduce the following quantity

$$M_n = \int_{\Lambda_{\theta}} M_n^{\lambda} q(\theta + \lambda | \alpha, \beta) d\lambda,$$

where  $\Lambda_{\theta} = \{\lambda : \theta + \lambda \in \Theta\}$ . This also satisfies  $\mathbb{E}[M_n] = 1$ . Further, we have the rewriting

$$M_n = H(\alpha, \beta) \int_{\Lambda_{\theta}} \exp \left( \langle \lambda, n(\mu_n - \mu) \rangle - n\mathcal{B}_{\mathcal{L},\theta}(\lambda) + \langle \lambda + \theta, \alpha \rangle - \beta \mathcal{L}(\theta + \lambda) \right) d\lambda.$$

**Step 2. Choice of parameters and duality properties.** Considering in particular the choice  $\alpha = c\nabla\mathcal{L}(\theta)$  and  $\beta = c$ , we get

$$\begin{aligned} M_n &= H(c\nabla\mathcal{L}(\theta), c) \int_{\Lambda_\theta} \exp \left( \langle \lambda, n(\mu_n - \mu) \rangle - n\mathcal{B}_{\mathcal{L},\theta}(\lambda) + c\langle \lambda + \theta, \nabla\mathcal{L}(\theta) \rangle - c\mathcal{L}(\theta + \lambda) \right) d\lambda \\ &= G(\theta, c) \int_{\Lambda_\theta} \exp \left( \langle \lambda, n(\mu_n - \mu) \rangle - (n+c)\mathcal{B}_{\mathcal{L},\theta}(\lambda) \right) d\lambda, \end{aligned}$$

where we introduce the following quantity

$$\begin{aligned} G(\theta, c) = H(c\nabla\mathcal{L}(\theta), c) \exp \left( c\langle \theta, \nabla\mathcal{L}(\theta) \rangle - c\mathcal{L}(\theta) \right) &= \frac{\exp \left( c\langle \theta, \nabla\mathcal{L}(\theta) \rangle - c\mathcal{L}(\theta) \right)}{\int_{\Theta} \exp \left( c\langle \theta', \nabla\mathcal{L}(\theta) \rangle - c\mathcal{L}(\theta') \right) d\theta'} \\ &= \left[ \int_{\Theta} \exp (-c\mathcal{B}_{\mathcal{L}}(\theta', \theta)) d\theta' \right]^{-1}. \end{aligned}$$

At this point, let us introduce  $x = \frac{n}{n+c}(\mu_n - \mu)$ . We also consider the Legendre-Fenchel dual function  $\mathcal{B}_{\mathcal{L},\theta}^*(x) = \max_{\lambda} \langle \lambda, x \rangle - \mathcal{B}_{\mathcal{L},\theta}(\lambda)$  and denote  $\lambda_x^*$  its maximal point. Now, we note that  $x + \nabla\mathcal{L}(\theta) = \nabla\mathcal{L}(\theta + \lambda_x^*)$  and thus

$$\begin{aligned} \langle \lambda, x \rangle - \mathcal{B}_{\mathcal{L},\theta}(\lambda) - \langle \lambda_x^*, x \rangle + \mathcal{B}_{\mathcal{L},\theta}(\lambda_x^*) &= \langle \lambda - \lambda_x^*, x + \nabla\mathcal{L}(\theta) \rangle + \mathcal{L}(\theta + \lambda_x^*) - \mathcal{L}(\theta + \lambda) \\ &= \langle \lambda - \lambda_x^*, \nabla\mathcal{L}(\theta + \lambda_x^*) \rangle + \mathcal{L}(\theta + \lambda_x^*) - \mathcal{L}(\theta + \lambda). \end{aligned}$$

Also, since it holds that  $x = \frac{n}{n+c}(\mu_n - \nabla\mathcal{L}(\theta))$ , then the quantity  $x + \nabla\mathcal{L}(\theta) = \nabla\mathcal{L}(\theta + \lambda_x^*)$  rewrites  $\frac{n}{n+c}\mu_n + \frac{c}{n+c}\nabla\mathcal{L}(\theta) = \nabla\mathcal{L}(\theta + \lambda_x^*)$ , which justifies to introduce the following regularized parameter estimate

$$\theta_{n,c}(\theta) = (\nabla\mathcal{L})^{-1} \left( \frac{n}{n+c}\mu_n + \frac{c}{n+c}\nabla\mathcal{L}(\theta) \right).$$

**Step 3. Martingale rewriting and conclusion.** Using this property, we note that

$$\begin{aligned} M_n &= \exp((n+c)\mathcal{B}_{\mathcal{L},\theta}^*(x)) \cdot G(\theta, c) \int_{\Lambda_\theta} \exp \left( (n+c) \left[ \langle \lambda - \lambda_x^*, \nabla\mathcal{L}(\theta + \lambda_x^*) \rangle - \mathcal{L}(\theta + \lambda) + \mathcal{L}(\theta + \lambda_x^*) \right] \right) d\lambda \\ &= \exp((n+c)\mathcal{B}_{\mathcal{L},\theta}^*(x)) \cdot G(\theta, c) \int_{\Lambda_\theta} \exp \left( \langle \lambda + \theta, (n+c)\nabla\mathcal{L}(\theta + \lambda_x^*) \rangle - (n+c)\mathcal{L}(\theta + \lambda) \right) d\lambda \times \\ &\quad \exp \left( -(n+c)\langle \lambda_x^* + \theta, \nabla\mathcal{L}(\theta + \lambda_x^*) + (n+c)\mathcal{L}(\theta + \lambda_x^*) \rangle \right) \\ &= \exp((n+c)\mathcal{B}_{\mathcal{L},\theta}^*(x)) \frac{G(\theta, c)}{G(\theta + \lambda_x^*, n+c)} = \exp((n+c)\mathcal{B}_{\mathcal{L},\theta}^*(x)) \frac{G(\theta, c)}{G(\theta_{n,c}(\theta), n+c)}. \end{aligned}$$

We then apply simple Markov inequality, which yields for all constant  $c$ ,

$$\forall \delta \in (0, 1], \quad \mathbb{P} \left( \mathcal{B}_{\mathcal{L},\theta}^* \left( \frac{n}{n+c}(\mu_n - \mu) \right) \geq \frac{1}{n+c} \log \left( \frac{G(\theta_{n,c}(\theta), n+c)}{G(\theta, c)} \frac{1}{\delta} \right) \right) \leq \delta.$$

To conclude, we use the duality property of the Bregman divergence (Lemma 1), considering that  $\nabla \mathcal{L}$  is invertible. Indeed, denoting  $\alpha = n/(n+c)$ , and  $\theta_n = (\nabla \mathcal{L})^{-1}(\mu_n)$ , it comes

$$\mathcal{B}_{\mathcal{L},\theta}^*(\alpha(\mu_n - \mu)) = \mathcal{B}_{\mathcal{L}}(\theta, \theta_\alpha) \quad \text{where } \theta_\alpha = \nabla \mathcal{L}^{-1}(\alpha \nabla \mathcal{L}(\theta_n) + (1-\alpha) \nabla \mathcal{L}(\theta)) = \theta_{n,c}(\theta).$$

Last, we note that this extends from a single  $n$  to being time-uniform thanks to Doob's maximal inequality for nonnegative supermartingale applied to  $M_n$  (also known as Ville (1939)).  $\square$

---

## A.2. Bregman concentration using Legendre function $\mathcal{L}_0$

In this section, we state a more general result that can handle the case of sequence  $(X_t)_{t \geq 1}$  of random variables that are not independent, each having possibly different distribution from others.

**Theorem 7 (Laplace method for non i.i.d. samples)** *Suppose  $\{\mathcal{H}_t\}_{t=0}^\infty$  is a filtration such that for each  $t \geq 1$ , (i)  $X_t$  is  $\mathcal{H}_t$ -measurable, (ii)  $F_t$  and  $h_t$  are  $\mathcal{H}_{t-1}$ -measurable and (iii) given  $\mathcal{H}_{t-1}$ ,  $X_t \sim p_{\theta_0,t}$  where  $p_{\theta_0,t}$  belongs to an exponential family with parameter  $\theta_0$ , feature function  $F_t$  and base function  $h_t$ . Let  $\mathcal{L}_t$  be the log-partition function corresponding to  $F_t$ . For any Legendre (i.e., strictly convex and continuously differentiable) function  $\mathcal{L}_0 : \Theta \rightarrow \mathbb{R}$  such that  $\exp(-\mathcal{L}_0)$  is integrable and any  $n \in \mathbb{N}$ , we introduce the parameter estimate and Bregman information gain*

$$\begin{aligned} \theta_{n,\mathcal{L}_0} &= \left( \sum_{t=0}^n \nabla \mathcal{L}_t \right)^{-1} \left( \sum_{t=1}^n F_t(X_t) \right), \\ \gamma_{n,\mathcal{L}_0} &= \log \left( \frac{\int_{\Theta} \exp(-\mathcal{L}_0(\theta')) d\theta'}{\int_{\Theta} \exp \left( -\sum_{t=0}^n \mathcal{B}_{\mathcal{L}_t}(\theta', \theta_{n,\mathcal{L}_0}) \right) d\theta'} \right), \end{aligned}$$

and then, for any  $\delta \in (0, 1]$ , the set

$$\Theta_{n,\mathcal{L}_0}(\delta) = \left\{ \theta : \sum_{t=0}^n \mathcal{B}_{\mathcal{L}_t}(\theta, \theta_{n,\mathcal{L}_0}) - \mathcal{L}_0(\theta) \leq \log \frac{1}{\delta} + \gamma_{n,\mathcal{L}_0} \right\}.$$

Then, the following time-uniform concentration holds:

$$\mathbb{P}[\exists n \in \mathbb{N} : \theta_0 \notin \Theta_{n,\mathcal{L}_0}(\delta)] \leq \delta.$$

In contrast to Theorem 3 that involves a local regularization using the true parameter  $\theta_0 \in \Theta$  of the family and a constant  $c > 0$ , here we make use of a global regularization, in the form of the Legendre function  $\mathcal{L}_0$ . Concretely, the regularized parameter estimate  $\theta_{n,\mathcal{L}_0}$  and the Bregman information gain  $\gamma_{n,\mathcal{L}_0}$  do not depend on the true parameter  $\theta_0$ , thus making for a more explicit confidence set  $\Theta_{n,\mathcal{L}_0}$ . The trade-off here is that the choice of regularizer is limited by the integrability assumption on  $\exp(-\mathcal{L}_0)$ , which is critical to build an appropriate prior for the method of mixtures.

The proof of this result follows a similar line of proof as before. However the use of  $\mathcal{L}_0$  instead of  $c$  induces a few changes that we detail below. In particular, we use a different prior to build the mixture of martingales.

---

**Proof of Theorem 7:**


---

**Step 1. Martingale and mixture martingale construction.** For any  $\lambda \in \mathbb{R}^d$ , we define

$$M_n^\lambda = \exp \left( \sum_{t=1}^n \left( \lambda^\top (F_t(X_t) - \mathbb{E}_\theta [F_t(X_t)]) - B_{\mathcal{L}_t, \theta}(\lambda) \right) \right),$$

where we introduced  $B_{\mathcal{L}_t, \theta}(\lambda) = \mathcal{L}_t(\theta + \lambda) - \mathcal{L}_t(\theta) - \langle \lambda, \nabla \mathcal{L}_t(\theta) \rangle$  for convenience.

Note that  $M_n^\lambda > 0$  and

$$\begin{aligned} \mathbb{E}_\theta \left[ \exp \left( \lambda^\top F_t(X_t) \right) \right] &= \exp (\mathcal{L}_t(\theta + \lambda) - \mathcal{L}_t(\theta)) , \\ \mathbb{E}_\theta \left[ \exp \left( \lambda^\top \mathbb{E}_\theta [F_t(X_t)] \right) \right] &= \exp \left( \lambda^\top \mathcal{L}_t(\theta) \right) . \end{aligned}$$

Note that  $M_n^\lambda$  is  $\mathcal{H}_n$ -measurable and in fact  $\mathbb{E}[M_n^\lambda | \mathcal{H}_{n-1}] = M_{n-1}^\lambda$ . Therefore  $\{M_n^\lambda\}_{n=0}^\infty$  is a nonnegative martingale adapted to the filtration  $\{\mathcal{H}_n\}_{n=0}^\infty$  and actually satisfies  $\mathbb{E}[M_n^\lambda] = 1$ . For any prior density  $q(\theta)$  for  $\theta$ , we now define a mixture of martingales

$$M_n = \int_{\Lambda_\theta} M_n^\lambda q(\theta + \lambda) d\lambda . \quad (1)$$

where  $\Lambda_\theta = \{\lambda : \theta + \lambda \in \Theta\}$ . Then  $M_n$  is also a martingale and actually satisfies  $\mathbb{E}[M_n] = 1$ . Now consider the prior density

$$q(\theta + \lambda) = c_0 \cdot \exp(-\mathcal{L}_0(\theta + \lambda)) , \quad (2)$$

where  $c_0 = \frac{1}{\int_\Theta \exp(-\mathcal{L}_0(\theta')) d\theta'}$  (which is well-defined since  $\exp(-\mathcal{L}_0)$  is integrable over  $\Theta$ ). We then have

$$M_n = c_0 \int_{\Lambda_\theta} \exp \left( \lambda^\top S_n - \sum_{t=1}^n B_{\mathcal{L}_t, \theta}(\lambda) - \mathcal{L}_0(\theta + \lambda) \right) d\lambda ,$$

where we denote  $S_n = \sum_{t=1}^n (F_t(X_t) - \mathbb{E}_\theta [F_t(X_t)])$ . Now, from the formula of parameter estimate, we have

$$\sum_{t=1}^n \nabla \mathcal{L}_t(\theta_{n, \mathcal{L}_0}) + \nabla \mathcal{L}_0(\theta_{n, \mathcal{L}_0}) = \sum_{t=1}^n F_t(X_t) . \quad (3)$$

This yields  $S_n = \sum_{t=1}^n (\nabla \mathcal{L}_t(\theta_{n, \mathcal{L}_0}) - \nabla \mathcal{L}_t(\theta)) + \nabla \mathcal{L}_0(\theta_{n, \mathcal{L}_0})$ .

**Step 2. Legendre function and Bregman properties.** We now introduce the function  $\mathcal{L}_{1:n}(\theta) = \sum_{t=1}^n \mathcal{L}_t(\theta)$ . Note that  $\mathcal{L}_{1:n}$  is a also Legendre function and its associated Bregman divergence satisfies

$$\mathcal{B}_{\mathcal{L}_{1:n}}(\theta', \theta) = \sum_{t=1}^n \left( \mathcal{L}_t(\theta') - \mathcal{L}_t(\theta) - (\theta' - \theta)^\top \nabla \mathcal{L}_t(\theta) \right) = \sum_{t=1}^n B_{\mathcal{L}_t}(\theta', \theta).$$

In this notation, we can rewrite  $\sum_{t=1}^n \mathcal{B}_{\mathcal{L}_t, \theta}(\lambda) = \mathcal{B}_{\mathcal{L}_{1:n}, \theta}(\lambda)$  and  $S_n = \nabla \mathcal{L}_n(\theta_{n, \mathcal{L}_0}) - \nabla \mathcal{L}_n(\theta) + \nabla \mathcal{L}_0(\theta_{n, \mathcal{L}_0})$ . We then obtain

$$M_n = c_0 \cdot \exp(-\mathcal{L}_0(\theta)) \int_{\Lambda_\theta} \exp\left(\lambda^\top x_n - \mathcal{B}_{\mathcal{L}_{1:n}, \theta}(\lambda) + \lambda^\top x_0 - \mathcal{B}_{\mathcal{L}_0, \theta}(\lambda)\right) d\lambda, \quad (4)$$

where we have introduced  $x_0 = \nabla \mathcal{L}_0(\theta_{n, \mathcal{L}_0}) - \nabla \mathcal{L}_0(\theta)$  and  $x_n = \nabla \mathcal{L}_{1:n}(\theta_{n, \mathcal{L}_0}) - \nabla \mathcal{L}_{1:n}(\theta)$ .

We now have from the Bregman-duality property recalled in Lemma 1 that

$$\begin{aligned} \sup_{\lambda \in \mathbb{R}^d} \left( \lambda^\top x_n - \mathcal{B}_{\mathcal{L}_{1:n}, \theta}(\lambda) \right) &= \mathcal{B}_{\mathcal{L}_{1:n}, \theta}^*(x_n) \\ &= \mathcal{B}_{\mathcal{L}_{1:n}, \theta}^*(\nabla \mathcal{L}_{1:n}(\theta_{n, \mathcal{L}_0}) - \nabla \mathcal{L}_{1:n}(\theta)) = \mathcal{B}_{\mathcal{L}_{1:n}}(\theta, \theta_{n, \mathcal{L}_0}). \end{aligned}$$

Further, any optimal  $\lambda$  must satisfy

$$\nabla \mathcal{L}_{1:n}(\theta + \lambda) - \nabla \mathcal{L}_{1:n}(\theta) = x_n \implies \nabla \mathcal{L}_{1:n}(\theta + \lambda) = \nabla \mathcal{L}_{1:n}(\theta_{n, \mathcal{L}_0}).$$

One possible solution is  $\lambda^* = \theta_{n, \mathcal{L}_0} - \theta$ . We then have

$$\begin{aligned} \lambda^\top x_n - \mathcal{B}_{\mathcal{L}_{1:n}, \theta}(\lambda) &= \lambda^\top x_n - \mathcal{B}_{\mathcal{L}_{1:n}, \theta}(\lambda) + \mathcal{B}_{\mathcal{L}_{1:n}}(\theta, \theta_{n, \mathcal{L}_0}) - (\lambda^{*\top} x_n - \mathcal{B}_{\mathcal{L}_{1:n}, \theta}(\lambda^*)) \\ &= \mathcal{B}_{\mathcal{L}_{1:n}}(\theta, \theta_{n, \mathcal{L}_0}) + (\lambda - \lambda^*)^\top \nabla \mathcal{L}_n(\theta + \lambda^*) + \mathcal{B}_{\mathcal{L}_{1:n}, \theta}(\lambda^*) - \mathcal{B}_{\mathcal{L}_{1:n}, \theta}(\lambda) - (\lambda - \lambda^*)^\top \nabla \mathcal{L}_{1:n}(\theta) \\ &= \mathcal{B}_{\mathcal{L}_{1:n}}(\theta, \theta_{n, \mathcal{L}_0}) + (\lambda - \lambda^*)^\top \nabla \mathcal{L}_{1:n}(\theta + \lambda^*) + \mathcal{L}_{1:n}(\theta + \lambda^*) - \mathcal{L}_{1:n}(\theta + \lambda). \end{aligned} \quad (5)$$

Similarly, we have

$$\lambda^\top x_0 - \mathcal{B}_{\mathcal{L}_0, \theta}(\lambda) = \mathcal{B}_{\mathcal{L}_0}(\theta, \theta_{n, \mathcal{L}_0}) + (\lambda - \lambda^*)^\top \nabla \mathcal{L}_0(\theta + \lambda^*) + \mathcal{L}_0(\theta + \lambda^*) - \mathcal{L}_0(\theta + \lambda). \quad (6)$$

**Step 3. Martingale rewriting and conclusion.** Plugin-in (6) and (5) in (4), we now obtain

$$\begin{aligned} M_n &= c_0 \exp\left(\sum_{j \in \{0,1:n\}} \mathcal{B}_{\mathcal{L}_j}(\theta, \theta_{n, \mathcal{L}_0}) - \mathcal{L}_0(\theta)\right) \\ &\quad \times \int_{\Lambda_\theta} \exp\left(\sum_{j \in \{0,1:n\}} \left((\lambda - \lambda^*)^\top \nabla \mathcal{L}_j(\theta + \lambda^*) + \mathcal{L}_j(\theta + \lambda^*) - \mathcal{L}_j(\theta + \lambda)\right)\right) d\lambda \\ &= c_0 \exp\left(\sum_{j \in \{0,1:n\}} \mathcal{B}_{\mathcal{L}_j}(\theta, \theta_{n, \mathcal{L}_0}) - \mathcal{L}_0(\theta)\right) \cdot \exp\left(-\sum_{j \in \{0,1:n\}} \left((\theta + \lambda^*)^\top \nabla \mathcal{L}_j(\theta + \lambda^*) - \mathcal{L}_j(\theta + \lambda^*)\right)\right) \\ &\quad \times \int_{\Lambda_\theta} \exp\left(\sum_{j \in \{0,1:n\}} \left((\theta + \lambda)^\top \nabla \mathcal{L}_j(\theta + \lambda^*) - \mathcal{L}_j(\theta + \lambda)\right)\right) d\lambda \\ &= \frac{c_0}{c_n} \exp\left(\sum_{j \in \{0,1:n\}} \mathcal{B}_{\mathcal{L}_j}(\theta, \theta_{n, \mathcal{L}_0}) - \mathcal{L}_0(\theta)\right) \\ &\quad \times \frac{\int_{\Lambda_\theta} \exp\left(\sum_{j \in \{0,1:n\}} \left((\theta + \lambda)^\top \nabla \mathcal{L}_j(\theta + \lambda^*) - \mathcal{L}_j(\theta + \lambda)\right)\right) d\lambda}{\int_{\Theta} \exp\left(\sum_{j \in \{0,1:n\}} \left((\theta')^\top \nabla \mathcal{L}_j(\theta + \lambda^*) - \mathcal{L}_j(\theta')\right)\right) d\theta'} \\ &= \frac{c_0}{c_n} \exp\left(\sum_{t=1}^n \mathcal{B}_{\mathcal{L}_t}(\theta, \theta_{n, \mathcal{L}_0}) + \mathcal{B}_{\mathcal{L}_0}(\theta, \theta_{n, \mathcal{L}_0}) - \mathcal{L}_0(\theta)\right), \end{aligned}$$

where we have introduced  $c_n = \frac{\exp\left(\sum_{j \in \{0,1:n\}} ((\theta + \lambda^*)^\top \nabla \mathcal{L}_j(\theta + \lambda^*) - \mathcal{L}_j(\theta + \lambda^*))\right)}{\int_\Theta \exp\left(\sum_{j \in \{0,1:n\}} ((\theta')^\top \nabla \mathcal{L}_j(\theta + \lambda^*) - \mathcal{L}_j(\theta'))\right) d\theta'}.$

We then obtain the following result by a simple application of Markov's inequality, which yields

$$\begin{aligned} & \mathbb{P}\left[\sum_{t=1}^n \mathcal{B}_{\mathcal{L}_t}(\theta, \theta_{n,\mathcal{L}_0}) + \mathcal{B}_{\mathcal{L}_0}(\theta, \theta_{n,\mathcal{L}_0}) - \mathcal{L}_0(\theta) \geq \log\left(\frac{c_n}{c_0} \frac{1}{\delta}\right)\right] \\ &= \mathbb{P}\left[M_n \geq \frac{1}{\delta}\right] \leq \delta \cdot \mathbb{E}[M_n] = \delta. \end{aligned} \quad (7)$$

Finally, since  $\lambda^* = \theta_{n,\mathcal{L}_0} - \theta$ , we have the more explicit form

$$\begin{aligned} \frac{c_n}{c_0} &= \frac{\int_\Theta \exp\left(-\mathcal{L}_0(\theta')\right) d\theta'}{\int_\Theta \exp\left(-\sum_{j \in \{0,1:n\}} \mathcal{B}_{\mathcal{L}_j}(\theta', \theta + \lambda^*)\right) d\theta'} \\ &= \frac{\int_\Theta \exp\left(-\mathcal{L}_0(\theta')\right) d\theta'}{\int_\Theta \exp\left(-\sum_{t=1}^n \mathcal{B}_{\mathcal{L}_t}(\theta', \theta_{n,\mathcal{L}_0}) - \mathcal{B}_{\mathcal{L}_0}(\theta', \theta_{n,\mathcal{L}_0})\right) d\theta'}. \end{aligned}$$

We obtain the time-uniform bound by applying Doob's maximal inequality for nonnegative supermartingale.  $\square$

The general result of Theorem 7 specifies straightforwardly to the case when all observations have same law, yielding the following corollary that we state now for completeness.

**Corollary 8** *Let  $\mu_n = \frac{1}{n} \sum_{t=1}^n F(X_t)$ . For all Legendre function  $\mathcal{L}_0$  such that  $\exp(-\mathcal{L}_0)$  is integrable over  $\Theta$ , let*

$$\theta_{n,\mathcal{L}_0} = (\nabla \mathcal{L} + \frac{1}{n} \nabla \mathcal{L}_0)^{-1}(\mu_n) \text{ and } \gamma_{n,\mathcal{L}_0} = \log \left( \frac{\int_\Theta \exp\left(-\mathcal{L}_0(\theta')\right) d\theta'}{\int_\Theta \exp\left(-n\mathcal{B}_{\mathcal{L}}(\theta', \theta_{n,\mathcal{L}_0}) - \mathcal{B}_{\mathcal{L}_0}(\theta', \theta_{n,\mathcal{L}_0})\right) d\theta'} \right).$$

Then for all  $\delta \in (0, 1]$ ,

$$\mathbb{P}\left(\exists n \in \mathbb{N}, n\mathcal{B}_{\mathcal{L}}(\theta, \theta_{n,\mathcal{L}_0}) + \mathcal{B}_{\mathcal{L}_0}(\theta, \theta_{n,\mathcal{L}_0}) \geq \log \frac{1}{\delta} + \mathcal{L}_0(\theta) + \gamma_{n,\mathcal{L}_0}\right) \leq \delta.$$

**Choice of Legendre function  $\mathcal{L}_0$**  A natural choice for the Legendre regularizer  $\mathcal{L}_0$  is to use the log-partition function of the exponential family at hand. If  $\mathcal{L}_0 = c\mathcal{L}$ , where  $c > 0$  is a scaling coefficient, the formula above simplify to

$$\theta_{n,\mathcal{L}_0} = \frac{\theta_n}{1 + \frac{c}{n}} \text{ and } \gamma_{n,\mathcal{L}_0} = \log \left( \frac{\int_\Theta \exp\left(-c\mathcal{L}(\theta')\right) d\theta'}{\int_\Theta \exp\left(-(n+c)\mathcal{B}_{\mathcal{L}}(\theta', \theta_{n,\mathcal{L}_0})\right) d\theta'} \right),$$

where  $\theta_n = (\nabla \mathcal{L})^{-1}(\mu_n)$  is the standard maximum likelihood estimate of the parameter  $\theta_0$ .

In the special case of the one-dimensional Gaussian family  $\mathcal{N}(\theta_0, \sigma^2)$  where the variance  $\sigma^2$  is known, the log-partition function regularizer defined by  $\mathcal{L}_0(\theta) = \frac{\theta^2}{2\sigma^2}$  satisfies the integrability assumption  $\int_{\mathbb{R}} e^{-c\mathcal{L}_0(\theta)} d\theta < \infty$  with  $c > 0$ . Straightforward calculations show that the resulting confidence set is the same as the one derived from Theorem 3 with local regularization  $c$  (see Appendix C.1). For many other standard families however, this integrability assumption may fail, as in the case of the exponential distributions  $\mathcal{E}(-\theta)$ , for which  $\mathcal{L}(\theta) = -\log(-\theta)$  (see Appendix C.5).

For other choices of Legendre function, computing the regularized parameter estimate  $\theta_{n,\mathcal{L}_0}$  and the information gain  $\gamma_{n,\mathcal{L}_0}$  requires inverting the function  $\nabla\mathcal{L} + \frac{1}{n}\nabla\mathcal{L}_0$  and computing a tedious integral, both of which seldom result in closed-form expressions. For these reasons, we recommend as a first step the use of local regularization (Theorem 3) over Legendre regularizers. In Appendix F, we detail an application of Theorem 7 to Gaussian contextual bandits using a quadratic regularizer.

## Appendix B. Properties of Bregman Divergences

In this section, we detail the derivation of some useful properties of the Bregman divergence that are used in the proof of the Bregman deviation inequalities.

**Lemma 1 (Properties of Bregman divergences)** *For all  $\theta \in \Theta_{\mathcal{D}}$  and  $\lambda \in \mathbb{R}^d$  such that  $\theta + \lambda \in \Theta_{\mathcal{D}}$ ,*

$$\log \mathbb{E}_{\theta} [\exp (\langle \lambda, F(X) - \mathbb{E}_{\theta}[F(X)] \rangle)] = \mathcal{B}_{\mathcal{L},\theta}(\lambda).$$

*Furthermore, if  $\nabla\mathcal{L}$  is one-to-one, the following Bregman duality relations hold for any  $\theta, \theta' \in \Theta_{\mathcal{D}}$ :*

$$\mathcal{B}_{\mathcal{L}}(\theta', \theta) = \mathcal{B}_{\mathcal{L},\theta'}^*(\nabla\mathcal{L}(\theta) - \nabla\mathcal{L}(\theta')) = \mathcal{B}_{\mathcal{L}^*}(\nabla\mathcal{L}(\theta), \nabla\mathcal{L}(\theta')).$$

*More generally, the following holds for any  $\alpha \in [0, 1]$ :*

$$\mathcal{B}_{\mathcal{L},\theta'}^*(\alpha(\nabla\mathcal{L}(\theta) - \nabla\mathcal{L}(\theta'))) = \mathcal{B}_{\mathcal{L}}(\theta', \theta_\alpha), \text{ where } \theta_\alpha = \nabla\mathcal{L}^{-1}(\alpha\nabla\mathcal{L}(\theta) + (1-\alpha)\nabla\mathcal{L}(\theta')).$$

### Proof of Lemma 1:

The first equality is immediate, since  $\mathbb{E}_{\theta}[F(X)] = \nabla\mathcal{L}(\theta)$  and

$$\begin{aligned} \log \mathbb{E}_{\theta} [\exp(\langle \lambda, F(X) \rangle)] &= \log \int \exp(\langle \lambda, F(x) \rangle + \langle \theta, F(x) \rangle - \mathcal{L}(x)) dx \\ &= \mathcal{L}(\lambda + \theta) - \mathcal{L}(\theta). \end{aligned}$$

We now turn to the duality formula. Using the definition of each terms, we have

$$\begin{aligned} \mathcal{B}_{\mathcal{L},\theta'}^*(\nabla\mathcal{L}(\theta) - \nabla\mathcal{L}(\theta')) &= \sup_{\lambda} \langle \lambda, \nabla\mathcal{L}(\theta) - \nabla\mathcal{L}(\theta') \rangle - [\mathcal{L}(\theta' + \lambda) - \mathcal{L}(\theta') - \langle \lambda, \nabla\mathcal{L}(\theta') \rangle] \\ &= \sup_{\lambda} \langle \lambda, \nabla\mathcal{L}(\theta) \rangle - \mathcal{L}(\theta' + \lambda) + \mathcal{L}(\theta'). \end{aligned}$$

An optimal  $\lambda$  must satisfy  $\nabla\mathcal{L}(\theta) = \nabla\mathcal{L}(\theta' + \lambda)$ . Hence, provided that  $\nabla\mathcal{L}$  is invertible, this means  $\lambda = \theta - \theta'$ . Plugin-in this value, we obtain

$$\mathcal{B}_{\mathcal{L},\theta'}^*(\nabla\mathcal{L}(\theta) - \nabla\mathcal{L}(\theta')) = \langle \theta - \theta', \nabla\mathcal{L}(\theta) \rangle - \mathcal{L}(\theta) + \mathcal{L}(\theta') = \mathcal{B}_{\mathcal{L}}(\theta', \theta).$$

The remaining equality  $\mathcal{B}_{\mathcal{L}}(\theta', \theta) = \mathcal{B}_{\mathcal{L}^*}(\nabla \mathcal{L}(\theta), \nabla \mathcal{L}(\theta'))$  is a standard result. Finally, regarding the generalization, we note that

$$\mathcal{B}_{\mathcal{L}, \theta'}^*(\alpha(\nabla \mathcal{L}(\theta) - \nabla \mathcal{L}(\theta'))) = \sup_{\lambda} \langle \lambda, \alpha \nabla \mathcal{L}(\theta) + (1 - \alpha) \nabla \mathcal{L}(\theta') \rangle - \mathcal{L}(\theta' + \lambda) + \mathcal{L}(\theta').$$

Hence, an optimal  $\lambda$  must now satisfy  $\alpha \nabla \mathcal{L}(\theta) + (1 - \alpha) \nabla \mathcal{L}(\theta') = \nabla \mathcal{L}(\theta' + \lambda)$ , that is  $\lambda = \theta_\alpha - \theta'$ . This further yields

$$\mathcal{B}_{\mathcal{L}, \theta'}^*(\alpha(\nabla \mathcal{L}(\theta) - \nabla \mathcal{L}(\theta'))) = \langle \theta_\alpha - \theta', \nabla \mathcal{L}(\theta_\alpha) \rangle - \mathcal{L}(\theta_\alpha) + \mathcal{L}(\theta') = \mathcal{B}_{\mathcal{L}}(\theta', \theta_\alpha).$$

□

---

## Appendix C. Specification to Illustrative Exponential Families

In this section, we provide the technical derivations to specify our generic concentration result to a few classical distributions. The results are summarized in Table 2.

### C.1. Gaussian with unknown mean, known variance

Consider  $X \sim \mathcal{N}(\mu, \sigma^2)$ , where  $\mu$  is unknown and  $\sigma$  is known. This corresponds to an exponential family of distributions, with parameter  $\theta = \mu$ , feature function  $F(x) = \frac{1}{\sigma^2}x$  and log-partition function  $\mathcal{L}(\theta) = \frac{\theta^2}{2\sigma^2}$ . The Bregman divergence between two parameters  $\theta'$  and  $\theta$  associated with  $\mathcal{L}$  is given by  $\mathcal{B}_{\mathcal{L}}(\theta', \theta) = \text{KL}(P_\mu, P_{\mu'}) = \frac{(\mu - \mu')^2}{2\sigma^2}$ . We further have  $\mathcal{L}'(\theta) = \frac{\theta}{\sigma^2} = \frac{\mu}{\sigma^2}$  and  $\mathcal{L}''(\theta) = \frac{1}{\sigma^2}$ , implying that  $\mathcal{L}'$  is invertible. Denoting  $S_n = \sum_{t=1}^n X_t$ , we obtain that

$$\mu_{n,c}(\mu) = \theta_{n,c}(\theta) = \frac{S_n + c\mu}{n + c}.$$

The Bregman deviations simplify as follows

$$(n + c) \cdot \mathcal{B}_{\mathcal{L}}(\theta, \theta_{n,c}(\theta)) = (n + c) \frac{(\theta_{n,c}(\theta) - \theta)^2}{2\sigma^2} = \frac{1}{n + c} \frac{(S_n - n\mu)^2}{2\sigma^2}.$$

Now, on the other hand, let us see that the information gain is explicitly given by

$$\gamma_{n,c}(\mu) = \frac{1}{2} \log \frac{n + c}{c}.$$

We obtain from Theorem 3 w.p.  $\geq 1 - \delta$ ,  $\forall n \in \mathbb{N}$ ,

$$\frac{1}{n + c} \frac{(S_n - n\mu)^2}{2\sigma^2} \leq \log(1/\delta) + \frac{1}{2} \log \frac{n + c}{c}.$$

## C.2. Gaussian with known mean, unknown variance

We consider  $X \sim \mathcal{N}(\mu, \sigma^2)$ , where  $\mu$  is known and  $\sigma$  is unknown. This corresponds to a one-dimensional exponential family model with parameter  $\theta = -\frac{1}{2\sigma^2} \in \mathbb{R}_-^*$ , feature function  $F(x) = (x - \mu)^2$  and log-partition function  $\mathcal{L}(\theta) = -\frac{1}{2} \log(-2\theta)$ . The first two derivatives of  $\mathcal{L}$  are given by  $\mathcal{L}'(\theta) = -\frac{1}{2\theta}$  and  $\mathcal{L}''(\theta) = \frac{1}{2\theta^2}$ , which shows that  $\mathcal{L}'$  is invertible on the domain  $\Theta = \mathbb{R}_-^*$ . The Bregman divergence between two parameters  $\theta'$  and  $\theta$  is therefore  $\mathcal{B}_{\mathcal{L}}(\theta', \theta) = \frac{1}{2} \log(\theta/\theta') + \frac{1}{2}(\theta'/\theta - 1)$ .

Let  $Q_n = \sum_{t=1}^n (X_t - \mu)^2$ . A short calculation shows that the following expression holds:

$$\theta_{n,c}(\theta) = -\frac{n+c}{c/\theta - 2Q_n}.$$

To compute the Bregman information gain, note that the above expression and a change of variable implies

$$\int_{-\infty}^0 \exp(-c\mathcal{B}_{\mathcal{L}}(\theta', \theta)) d\theta' = -\theta (2/c)^{c/2+1} e^{c/2} \Gamma(c/2 + 2).$$

Combining the last two lines yields

$$\begin{aligned} \gamma_{n,c}(\theta) &= \log \frac{\theta}{\theta_{n,c}(\theta)} - (1+\log 2)\frac{n}{2} - (\frac{c}{2}+1) \log c + (\frac{n+c}{2}+1) \log(n+c) - \frac{\log \Gamma(\frac{n+c}{2}+2)}{\log \Gamma(\frac{c}{2}+2)} \\ &= \log \frac{c-2\theta Q_n}{n+c} - (1+\log 2)\frac{n}{2} - (\frac{c}{2}+1) \log c + (\frac{n+c}{2}+1) \log(n+c) - \frac{\log \Gamma(\frac{n+c}{2}+2)}{\log \Gamma(\frac{c}{2}+2)} \\ &= \log \frac{Q_n/\sigma^2 + c}{n+c} - (1+\log 2)\frac{n}{2} - (\frac{c}{2}+1) \log c + (\frac{n+c}{2}+1) \log(n+c) - \frac{\log \Gamma(\frac{n+c}{2}+2)}{\log \Gamma(\frac{c}{2}+2)}. \end{aligned}$$

Moreover, we deduce from the expression of  $\theta_{n,c}(\theta)$  and the Bregman divergence that:

$$\mathcal{B}_{\mathcal{L}}(\theta', \theta_{n,c}(\theta)) = -\frac{1}{2} \log \left( \frac{c\theta'/\theta - 2\theta'Q_n}{n+c} \right) + \frac{1}{2} \left( \frac{c\theta'/\theta - 2\theta'Q_n}{n+c} - 1 \right).$$

So, we have

$$(n+c)\mathcal{B}_{\mathcal{L}}(\theta, \theta_{n,c}(\theta)) = -\frac{n+c}{2} \log \left( \frac{Q_n/\sigma^2 + c}{n+c} \right) + \frac{Q_n/\sigma^2 + c}{2} - \frac{n}{2}.$$

After some simple algebra and using the natural parametrization  $\theta = -\frac{1}{2\sigma^2}$ , we derive from Theorem 3 that w.p.  $\geq 1 - \delta$ ,  $\forall n \in \mathbb{N}$ ,

$$\begin{aligned} &- (\frac{n+c}{2}+1) \log \frac{Q_n/\sigma^2 + c}{n+c} + \frac{Q_n}{2\sigma^2} \\ &\leq \log \frac{1}{\delta} - \frac{n}{2} \log 2 - (\frac{c}{2}+1) \log c + (\frac{n+c}{2}+1) \log(n+c) + \log \Gamma(\frac{c}{2}+2) - \log \Gamma(\frac{n+c}{2}+2). \end{aligned} \tag{8}$$

### C.3. Gaussian with unknown mean and variance

We consider  $X \sim \mathcal{N}(\mu, \sigma^2)$ , where both  $\mu \in \mathbb{R}$  and  $\sigma > 0$  are unknown. These distributions form a two-dimensional exponential family with parameter  $\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} \frac{\mu}{\sigma^2} \\ -\frac{1}{2\sigma^2} \end{pmatrix}$  belonging to the domain  $\Theta = \mathbb{R} \times \mathbb{R}_+$ . The corresponding feature and log-partition functions writes  $F(x) = \begin{pmatrix} x \\ x^2 \end{pmatrix}$  and  $\mathcal{L}(\theta) = -\frac{1}{2} \log(-\theta_2) - \frac{\theta_1^2}{4\theta_2}$  respectively. The gradient and Hessian of  $\mathcal{L}$  follows from straightforward calculations and write:

$$\nabla \mathcal{L}(\theta) = \begin{pmatrix} -\frac{\theta_1}{2\theta_2} \\ -\frac{1}{2\theta_2} + \frac{\theta_1^2}{4\theta_2^2} \end{pmatrix}, \quad \nabla^2 \mathcal{L}(\theta) = \begin{pmatrix} -\frac{1}{2\theta_2} & \frac{\theta_1}{2\theta_2^2} \\ \frac{\theta_1}{2\theta_2^2} & \frac{1}{2\theta_2^2} - \frac{\theta_1^2}{2\theta_2^3} \end{pmatrix}.$$

In particular,  $\text{Tr } \nabla^2 \mathcal{L}(\theta) = -\frac{1}{2\theta_2} + \frac{1}{2\theta_2^2} - \frac{\theta_1^2}{2\theta_2^3} > 0$  and  $\det \nabla^2 \mathcal{L}(\theta) = -\frac{1}{4\theta_2^3} > 0$ , therefore the symmetric matrix  $\nabla^2 \mathcal{L}(\theta)$  is positive definite (both its eigenvalues are positive). The inverse of the gradient mapping  $\nabla \mathcal{L}: \Theta \rightarrow \mathcal{P}$ , where the image set is  $\mathcal{P} = \{(x, y) \in \mathbb{R}^2, y > x^2\}$ , can be explicitly computed as:

$$(\nabla)^{-1} \mathcal{L} \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{y - x^2} \begin{pmatrix} x \\ -\frac{1}{2} \end{pmatrix}.$$

The expression of the Bregman divergence between two parameters  $\theta$  and  $\theta'$  can be calculated from the above results and reads:

$$\mathcal{B}_{\mathcal{L}}(\theta', \theta) = \frac{1}{2} \log \frac{\theta_2}{\theta'_2} + \frac{\theta'_2}{2\theta_2} - \theta'_2 \left( \frac{\theta'_1}{2\theta'_2} - \frac{\theta_1}{2\theta_2} \right)^2 - \frac{1}{2}.$$

Let  $S_n = \sum_{t=1}^n X_t$  and  $Q_n = \sum_{t=1}^n X_t^2$  and  $V_{n,c} = \frac{1}{n+c} Q_n - \left( \frac{1}{n+c} S_n \right)^2$ . We have that

$$\theta_{n,c}(\theta) = \frac{1}{V_{n,c} - \frac{c}{n+c} \frac{1}{2\theta_2} + \frac{nc}{(n+c)^2} \frac{\theta_1^2}{4\theta_2^2} + \frac{c}{(n+c)^2} \frac{\theta_1}{\theta_2} S_n} \begin{pmatrix} \frac{1}{n+c} S_n - \frac{c}{2(n+c)} \frac{\theta_1}{\theta_2} \\ -\frac{1}{2} \end{pmatrix}.$$

To compute the Bregman information gain, we will need to evaluate the integral  $\int_{\Theta} \exp(-c\mathcal{B}_{\mathcal{L}}(\theta', \theta)) d\theta'$ . As the integrand is nonnegative, we can integrate first along  $\theta'_1$  (Fubini's theorem), which writes:

$$\begin{aligned} \int_{\Theta} \exp(-c\mathcal{B}_{\mathcal{L}}(\theta', \theta)) d\theta' &= \int_{-\infty}^0 \int_{-\infty}^{\infty} \exp \left( -\frac{c}{2} \log \left( \frac{\theta_2}{\theta'_2} \right) - \frac{c}{2} \frac{\theta'_2}{\theta_2} + c\theta'_2 \left( \frac{\theta'_1}{2\theta'_2} - \frac{\theta_1}{2\theta_2} \right)^2 + \frac{c}{2} \right) d\theta'_1 d\theta'_2 \\ &= e^{\frac{c}{2}} \int_{-\infty}^0 \left( \frac{\theta'_2}{\theta_2} \right)^{\frac{c}{2}} e^{-\frac{c}{2} \frac{\theta'_2}{\theta_2}} \int_{-\infty}^{\infty} \exp \left( c\theta'_2 \left( \frac{\theta'_1}{2\theta'_2} - \frac{\theta_1}{2\theta_2} \right)^2 \right) d\theta'_1 d\theta'_2. \end{aligned}$$

Let  $m = \frac{\theta_1}{2\theta_2}$  and  $s = \frac{1}{\sqrt{-2c\theta'_2}}$ . The integral w.r.t  $\theta'_1$  can be rewritten using the change of variable  $y = \frac{\theta'_1}{2\theta'_2}$  as:

$$\int_{-\infty}^{\infty} \exp \left( c\theta'_2 \left( \frac{\theta'_1}{2\theta'_2} - \frac{\theta_1}{2\theta_2} \right)^2 \right) d\theta'_1 = -2\theta'_2 \sqrt{2\pi} s \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}s} e^{-\frac{(y-m)^2}{2s^2}} dy}_{=1}.$$

Therefore, the above calculation simplifies to:

$$\int_{\Theta} \exp(-c\mathcal{B}_{\mathcal{L}}(\theta', \theta)) d\theta' = 2\sqrt{\frac{\pi}{c}} e^{\frac{c}{2}} \int_{-\infty}^0 \left(\frac{\theta'_2}{\theta_2}\right)^{\frac{c}{2}} e^{-\frac{c}{2}\frac{\theta'_2}{\theta_2}} \sqrt{-\theta'_2} d\theta'_2,$$

which after a linear change of variable on  $\theta'_2$  can be related to the Gamma function as follows:

$$\int_{\Theta} \exp(-c\mathcal{B}_{\mathcal{L}}(\theta', \theta)) d\theta' = 2\sqrt{\frac{\pi}{c}} \left(\frac{2}{c}\right)^{\frac{c+1}{2}+1} \Gamma\left(\frac{c+1}{2} + 1\right) (-\theta_2)^{\frac{3}{2}}.$$

The Bregman information gain is thus:

$$\begin{aligned} \gamma_{n,c}(\theta) &= -\left(\frac{1+\log 2}{2}\right)n - \left(\frac{c}{2} + 2\right)\log c + \left(\frac{n+c}{2} + 2\right)\log(n+c) \\ &\quad + \log \Gamma\left(\frac{c+3}{2}\right) - \log \Gamma\left(\frac{n+c+3}{2}\right) + \frac{3}{2}\log \frac{\theta_2}{\theta_{n,c}(\theta)_2}, \end{aligned}$$

with  $\frac{\theta_2}{\theta_{n,c}(\theta)_2} = -2\theta_2 V_{n,c} + \frac{c}{n+c} - \frac{nc}{(n+c)^2} \frac{\theta_1^2}{2\theta_2} - \frac{2c\theta_1}{(n+c)^2} S_n$ . Now, applying the result of Theorem 3 shows that w.p.  $\geq 1 - \delta$ ,  $\forall n \in \mathbb{N}$ ,

$$\begin{aligned} &- \frac{n+c+3}{2} \log\left(\frac{\theta_2}{\theta_{n,c}(\theta)_2}\right) - \theta_2(n+c)V_{n,c} - \frac{n\theta_1^2}{4\theta_2} - \frac{\theta_2}{n+c} S_n^2 - \theta_1 S_n \\ &\leq \log \frac{1}{\delta} - \frac{n}{2} \log 2 - \left(\frac{c}{2} + 2\right) \log c + \left(\frac{n+c}{2} + 2\right) \log(n+c) + \log \Gamma\left(\frac{c+3}{2}\right) - \log \Gamma\left(\frac{n+c+3}{2}\right). \end{aligned}$$

After expanding the ratio  $\frac{\theta_2}{\theta_{n,c}(\theta)_2}$  and substituting the natural parametrization in terms of  $\mu$  and  $\sigma$ , we finally obtain that w.p.  $\geq 1 - \delta$ ,  $\forall n \in \mathbb{N}$ ,

$$\begin{aligned} &- \frac{n+c+3}{2} \log\left(\frac{1}{\sigma^2} V_{n,c} + \frac{c}{n+c} + \frac{nc}{(n+c)^2} \frac{\mu^2}{\sigma^2} - \frac{2c}{(n+c)^2} \frac{\mu}{\sigma^2} S_n\right) + \frac{n+c}{2\sigma^2} V_{n,c} + \frac{n\mu^2}{2\sigma^2} + \frac{1}{2(n+c)\sigma^2} S_n^2 - \frac{\mu}{\sigma^2} S_n \\ &\leq \log \frac{1}{\delta} - \frac{n}{2} \log 2 - \left(\frac{c}{2} + 2\right) \log c + \left(\frac{n+c}{2} + 2\right) \log(n+c) + \log \Gamma\left(\frac{c+3}{2}\right) - \log \Gamma\left(\frac{n+c+3}{2}\right). \end{aligned}$$

To simplify this formula, we introduce the standardized sum of squares  $Z_n(m, s) = \sum_{t=1}^n \left(\frac{X_t - m}{s}\right)^2$  for  $(m, s) \in \mathbb{R} \times \mathbb{R}_+^*$ . After rearranging terms and denoting  $\hat{\mu}_n = S_n/n$ , the above formula reads:

$$\begin{aligned} &\frac{1}{2} Z_n(\mu, \sigma) - \frac{n+c+3}{2} \log\left(\frac{n}{n+c} Z_n(\hat{\mu}_n, \sigma) + \frac{c}{n+c} Z_n(\mu, \sigma) + c\right) \\ &\leq \log \frac{1}{\delta} - \frac{n}{2} \log 2 - \left(\frac{c}{2} + 2\right) \log c + \frac{1}{2} \log(n+c) + \log \Gamma\left(\frac{c+3}{2}\right) - \log \Gamma\left(\frac{n+c+3}{2}\right). \end{aligned}$$

#### C.4. Bernoulli

We consider  $X \sim \text{Bernoulli}(\mu)$ . This corresponds to an exponential family model, with parameter  $\theta = \log \frac{\mu}{1-\mu}$ , feature function  $F(x) = x$  and log-partition function  $\mathcal{L}(\theta) = \log(1 + \exp(\theta))$ . The Bregman divergence between two parameters  $\theta'$  and  $\theta$  associated with  $\mathcal{L}$  is given

by  $\mathcal{B}_{\mathcal{L}}(\theta', \theta) = \text{KL}(\mu, \mu') = \mu \log \frac{\mu}{\mu'} + (1 - \mu) \log \frac{1 - \mu}{1 - \mu'}$ . We further have  $\mathcal{L}'(\theta) = \frac{\exp(\theta)}{1 + \exp(\theta)} = \mu$  and  $\mathcal{L}''(\theta) = \frac{\exp(\theta)}{(1 + \exp(\theta))^2} = \mu(1 - \mu)$ . Therefore  $\mathcal{L}'$  is invertible and we have the expression  $\theta_{n,c}(\theta) = (\mathcal{L}')^{-1} \left( \frac{1}{n+c} \sum_{t=1}^n X_t + \frac{c}{n+c} \frac{\exp(\theta)}{1 + \exp(\theta)} \right)$ . Then, denoting  $S_n = \sum_{t=1}^n X_t$ , we get

$$\mu_{n,c}(\mu) := \mathcal{L}'(\theta_{n,c}(\theta)) = \frac{\sum_{t=1}^n X_t + c\mu}{n+c} = \frac{S_n + c\mu}{n+c}.$$

Therefore, the Bregman deviation specifies to the following closed-form formula

$$\begin{aligned} (n+c) \cdot \mathcal{B}_{\mathcal{L}}(\theta, \theta_{n,c}(\theta)) &= (n+c) \cdot \text{KL}(\mu_{n,c}(\mu), \mu) \\ &= (S_n + c\mu) \log \frac{\mu_{n,c}(\mu)}{\mu} + (n - S_n + c(1 - \mu)) \log \frac{1 - \mu_{n,c}(\mu)}{1 - \mu}. \end{aligned}$$

Now, we observe that

$$\begin{aligned} \int_{\mathbb{R}^d} \exp(-c\mathcal{B}_{\mathcal{L}}(\theta', \theta)) d\theta' &= \frac{\mathbf{B}(c\mu, c(1-\mu))}{\mu^{c\mu}(1-\mu)^{c(1-\mu)}}, \\ \int_{\mathbb{R}^d} \exp(-(n+c)\mathcal{B}_{\mathcal{L}}(\theta', \theta_{n,c}(\theta))) d\theta' &= \frac{\mathbf{B}((n+c)\mu_{n,c}(\mu), (n+c)(1-\mu_{n,c}(\mu)))}{\mu_{n,c}(\mu)^{(n+c)\mu_{n,c}(\mu)}(1-\mu_{n,c}(\mu))^{(n+c)(1-\mu_{n,c}(\mu))}}. \end{aligned}$$

Therefore, we deduce that the Bregman information gain rewrites

$$\begin{aligned} \gamma_{n,c}(\mu) &= (S_n + c\mu) \log \mu_{n,c}(\mu) + (n - S_n + c(1 - \mu)) \log(1 - \mu_{n,c}(\mu)) - c\mu \log \mu - c(1 - \mu) \log(1 - \mu) \\ &\quad + \log \frac{\Gamma(c\mu)\Gamma(c(1-\mu))}{\Gamma(S_n + c\mu)\Gamma(n - S_n + c(1-\mu))} + \log \frac{\Gamma(n+c)}{\Gamma(c)}. \end{aligned}$$

Combining the above and using Theorem 3, we obtain that w.p.  $\geq 1 - \delta$ ,  $\forall n \in \mathbb{N}$ ,

$$S_n \log \frac{1}{\mu} + (n - S_n) \log \frac{1}{1 - \mu} + \log \frac{\Gamma(S_n + c\mu)\Gamma(n - S_n + c(1 - \mu))}{\Gamma(c\mu)\Gamma(c(1 - \mu))} \leq \log \frac{1}{\delta} + \log \frac{\Gamma(n+c)}{\Gamma(c)}.$$

### C.5. Exponential

We consider  $X \sim \text{Exp}(1/\mu)$  with unknown mean  $\mu$ . The distribution of  $X$  is supported on  $[0, +\infty)$  with density  $p_\mu(x) = \frac{1}{\mu} e^{-\frac{x}{\mu}}$ . This corresponds to an exponential family model with parameter  $\theta = -\frac{1}{\mu}$ , feature function  $F(x) = x$  and log-partition function  $\mathcal{L}(\theta) = \log(-\frac{1}{\theta})$ . The Bregman divergence between two parameters  $\theta'$  and  $\theta$  associated with  $\mathcal{L}$  is given by  $\mathcal{B}_{\mathcal{L}}(\theta', \theta) = \text{KL}(P_\mu, P_{\mu'}) = \frac{\mu}{\mu'} - 1 - \log \frac{\mu}{\mu'}$ .

We have  $\mathcal{L}'(\theta) = -\frac{1}{\theta} = \mu$  and  $\mathcal{L}''(\theta) = \frac{1}{\theta^2} = \mu^2$ . Therefore  $\mathcal{L}'$  is invertible and we have

$$\mu_{n,c}(\mu) := \mathcal{L}'(\theta_{n,c}(\theta)) = \frac{S_n + c\mu}{n+c}.$$

Therefore, we deduce that the Bregman divergence takes the following form

$$(n+c) \cdot \mathcal{B}_{\mathcal{L}}(\theta, \theta_{n,c}(\theta)) = (n+c) \cdot \text{KL}(P_{\mu_{n,c}(\mu)}, P_\mu) = \frac{S_n}{\mu} - (n+c) \log \frac{\mu_{n,c}(\mu)}{\mu} - n.$$

Now, we observe that

$$\int_{\mathbb{R}^d} \exp(-c\mathcal{B}_{\mathcal{L}}(\theta', \theta)) d\theta' = \frac{\Gamma(c+1)}{c\mu} \left(\frac{e}{c}\right)^c = \frac{\Gamma(c)}{\mu} \left(\frac{e}{c}\right)^c,$$

$$\int_{\mathbb{R}^d} \exp(-(n+c)\mathcal{B}_{\mathcal{L}}(\theta', \theta_{n,c}(\theta))) d\theta' = \frac{\Gamma(n+c+1)}{(n+c)\mu_{n,c}(\mu)} \left(\frac{e}{n+c}\right)^{n+c} = \frac{\Gamma(n+c)}{\mu_{n,c}(\mu)} \left(\frac{e}{n+c}\right)^{n+c}.$$

Therefore, the Bregman information gain writes explicitly as follows

$$\gamma_{n,c}(\theta) = \log \frac{\mu_{n,c}(\mu)}{\mu} + \log \left( \Gamma(c) \left(\frac{e}{c}\right)^c \right) - \log \left( \Gamma(n+c) \left(\frac{e}{n+c}\right)^{n+c} \right).$$

We can now specify the inequality  $(n+c) \cdot \mathcal{B}_{\mathcal{L}}(\theta, \theta_{n,c}(\theta)) \leq \log(1/\delta) + \gamma_{n,c}(\theta)$ . Combining the above, we obtain w.p.  $\geq 1 - \delta, \forall n \in \mathbb{N}$ ,

$$\frac{S_n}{\mu} - (n+c+1) \log \left( \frac{S_n + c\mu}{(n+c)\mu} \right) \leq (n+c) \log(n+c) + \log \frac{\Gamma(c)}{\Gamma(n+c)} + \log \frac{1}{\delta} - c \log c.$$

### C.6. Gamma with fixed shape

We consider  $X \sim \text{Gamma}(\lambda, k)$  with fixed shape  $k > 0$  and unknown scale  $\lambda > 0$ .<sup>2</sup> The distribution of  $X$  is supported on  $[0, \infty)$  with density  $p_\lambda(x) = \frac{1}{\Gamma(k)\lambda^k} x^{k-1} e^{-\frac{x}{\lambda}}$ . This corresponds to an exponential family model with parameter  $\theta = -\frac{1}{\lambda}$ , feature function  $F(x) = x$  and log-partition function  $\mathcal{L}(\theta) = \log(\lambda^k) = k \log(-\frac{1}{\theta})$ . The Bregman divergence between two parameters  $\theta'$  and  $\theta$  associated with  $\mathcal{L}$  is given by  $\mathcal{B}_{\mathcal{L}}(\theta', \theta) = \text{KL}(p_\lambda, p_{\lambda'}) = k \left( \frac{\lambda}{\lambda'} - 1 - \log \left( \frac{\lambda}{\lambda'} \right) \right)$ .

Note that  $\mathcal{L}'(\theta) = -\frac{k}{\theta} = k\lambda$  and  $\mathcal{L}''(\theta) = \frac{k}{\theta^2} = k\lambda$ . Therefore  $\mathcal{L}'$  is invertible, and we get

$$k\lambda_{n,c}(\lambda) := \mathcal{L}'(\theta_{n,c}(\theta)) = \frac{\sum_{t=1}^n X_t + ck\lambda}{n+c}$$

Therefore, the Bregman divergence takes the following form

$$(n+c) \cdot \mathcal{B}_{\mathcal{L}}(\theta, \theta_{n,c}(\theta)) = (n+c) \cdot \text{KL}(p_{\lambda_{n,c}(\lambda)}, p_\lambda)$$

$$= k(n+c) \left( \frac{\sum_{t=1}^n X_t + ck\lambda}{(n+c)k\lambda} \right) - k(n+c) \log \left( \frac{\sum_{t=1}^n X_t + ck\lambda}{(n+c)k\lambda} \right) - k(n+c).$$

Now, we observe that

$$\int_{\mathbb{R}^d} \exp(-c\mathcal{B}_{\mathcal{L}}(\theta', \theta)) d\theta' = \frac{\Gamma(ck)}{\lambda} \left(\frac{e}{ck}\right)^{ck},$$

$$\int_{\mathbb{R}^d} \exp(-(n+c)\mathcal{B}_{\mathcal{L}}(\theta', \theta_{n,c}(\theta))) d\theta' = \frac{\Gamma((n+c)k)}{\lambda_{n,c}(\lambda)} \left(\frac{e}{(n+c)k}\right)^{(n+c)k}.$$

Therefore, the Bregman information gain writes explicitly as follows

$$\gamma_{n,c}(\theta) = \log \left( \frac{\sum_{t=1}^n X_t + ck\lambda}{(n+c)k\lambda} \right) + \log \left( \Gamma(ck) \left(\frac{e}{ck}\right)^{ck} \right) - \log \left( \Gamma((n+c)k) \left(\frac{e}{(n+c)k}\right)^{(n+c)k} \right).$$

---

2. Note that  $\text{Gamma}(\lambda, k)$  with  $k = 1$  is  $\text{Exp}(1/\lambda)$ .

Using the above with Theorem 3, we obtain w.p.  $\geq 1 - \delta, \forall n \in \mathbb{N}$ ,

$$\begin{aligned} & k(n+c) \left( \frac{\sum_{t=1}^n X_t + ck\lambda}{(n+c)k\lambda} \right) - (k(n+c)+1) \log \left( \frac{\sum_{t=1}^n X_t + ck\lambda}{(n+c)k\lambda} \right) \\ & \leq \log \frac{1}{\delta} + \log \frac{\Gamma(ck)}{\Gamma((n+c)k)} + (n+c)k \log((n+c)k) + ck - ck \log ck. \end{aligned} \quad (9)$$

### C.7. Weibull with fixed shape

We consider  $X \sim \text{Weibull}(\lambda, k)$  with fixed shape  $k > 0$  and unknown scale  $\lambda > 0$ .<sup>3</sup> The distribution of  $X$  is supported on  $[0, \infty)$  with density  $p_\lambda(x) = \frac{k}{\lambda} \left( \frac{x}{\lambda} \right)^{k-1} e^{-\left( \frac{x}{\lambda} \right)^k}$ . This corresponds to an exponential family model with parameter  $\theta = -\frac{1}{\lambda^k}$ , feature function  $F(x) = x^k$  and log-partition function  $\mathcal{L}(\theta) = \log(\lambda^k) = \log(-\frac{1}{\theta})$ . The Bregman divergence between two parameters  $\theta'$  and  $\theta$  associated with  $\mathcal{L}$  is given by  $\mathcal{B}_{\mathcal{L}}(\theta', \theta) = \text{KL}(P_\lambda, P_{\lambda'}) = \left( \frac{\lambda}{\lambda'} \right)^k - 1 - \log \left( \frac{\lambda}{\lambda'} \right)^k$ .

Note that  $\mathcal{L}'(\theta) = -\frac{1}{\theta} = \lambda^k$  and  $\mathcal{L}''(\theta) = \frac{1}{\theta^2} = \lambda^{2k}$ . Therefore  $\mathcal{L}'$  is invertible, and we get

$$(\lambda_{n,c}(\lambda))^k := \mathcal{L}'(\theta_{n,c}(\theta)) = \frac{\sum_{t=1}^n X_t^k + c\lambda^k}{n+c}.$$

Therefore, the Bregman divergence takes the following form

$$\begin{aligned} (n+c) \cdot \mathcal{B}_{\mathcal{L}}(\theta, \theta_{n,c}(\theta)) &= (n+c) \cdot \text{KL}(p_{\lambda_{n,c}(\lambda)}, p_\lambda) \\ &= (n+c) \left( \frac{\sum_{t=1}^n X_t^k + c\lambda^k}{(n+c)\lambda^k} \right) - (n+c) \log \left( \frac{\sum_{t=1}^n X_t^k + c\lambda^k}{(n+c)\lambda^k} \right) - (n+c). \end{aligned}$$

Now, we observe that

$$\begin{aligned} \int_{\mathbb{R}^d} \exp(-c\mathcal{B}_{\mathcal{L}}(\theta', \theta)) d\theta' &= \frac{\Gamma(c)}{\lambda^k} \left( \frac{e}{c} \right)^c, \\ \int_{\mathbb{R}^d} \exp(-(n+c)\mathcal{B}_{\mathcal{L}}(\theta', \theta_{n,c}(\theta))) d\theta' &= \frac{\Gamma(n+c)}{(\lambda_{n,c}(\lambda))^k} \left( \frac{e}{n+c} \right)^{n+c}. \end{aligned}$$

Therefore, the Bregman information gain writes explicitly as follows

$$\gamma_{n,c}(\theta) = \log \left( \frac{\sum_{t=1}^n X_t^k + c\lambda^k}{(n+c)\lambda^k} \right) + \log \left( \Gamma(c) \left( \frac{e}{c} \right)^c \right) - \log \left( \Gamma(n+c) \left( \frac{e}{n+c} \right)^{n+c} \right).$$

Using the above with Theorem 3, we obtain w.p.  $\geq 1 - \delta, \forall n \in \mathbb{N}$ ,

$$\begin{aligned} & (n+c) \left( \frac{\sum_{t=1}^n X_t^k + c\lambda^k}{(n+c)\lambda^k} \right) - (n+c+1) \log \left( \frac{\sum_{t=1}^n X_t^k + c\lambda^k}{(n+c)\lambda^k} \right) \\ & \leq \log \frac{1}{\delta} + \log \frac{\Gamma(c)}{\Gamma(n+c)} + (n+c) \log(n+c) + c - c \log c. \end{aligned} \quad (10)$$

---

3. Note that  $\text{Weibull}(\lambda, k)$  with  $k = 1$  is  $\text{Exp}(1/\lambda)$ .

### C.8. Pareto with fixed scale

We consider  $X \sim \text{Pareto}(\alpha)$ , where  $\alpha > 0$  is unknown.<sup>4</sup> The distribution of  $X$  is supported in  $[1, +\infty)$ , with density  $p_\alpha(x) = \frac{\alpha}{x^{\alpha+1}}$ , which corresponds to a one-dimensional exponential family model with parameter  $\theta = -\alpha - 1 \in (-\infty, -1)$ , feature function  $F(x) = \log x$  and log-partition function  $\mathcal{L}(\theta) = -\log(-1-\theta)$ . The first two derivatives of  $\mathcal{L}$  are given by  $\mathcal{L}'(\theta) = -\frac{1}{1+\theta}$  and  $\mathcal{L}''(\theta) = \frac{1}{(1+\theta)^2}$ , therefore  $\mathcal{L}'$  is invertible on the domain  $(-\infty, -1)$ . Using these expressions, the Bregman divergence between two parameters  $\theta'$  and  $\theta$  writes  $\mathcal{B}_{\mathcal{L}}(\theta', \theta) = -\log(-1-\theta') + \log(-1-\theta) + \frac{\theta' - \theta}{1+\theta}$ . Using the shorthand  $L_n = \sum_{t=1}^n \log X_t$ , it follows from the definition that:

$$\theta_{n,c}(\theta) = -1 + \frac{n+c}{\frac{c}{1+\theta} - L_n}.$$

To compute the Bregman information gain, we rewrite the following integral thanks to an affine change of variable in order to relate it to the Gamma function:

$$\begin{aligned} \int_{-\infty}^{-1} \exp(-c\mathcal{B}_{\mathcal{L}}(\theta', \theta)) d\theta' &= (-1-\theta)^{-c} e^{\frac{c\theta}{1+\theta}} \int_{-\infty}^{-1} (-1-\theta')^c e^{-\frac{c\theta'}{1+\theta'}} d\theta' \\ &= (-1-\theta) \Gamma(c) \left(\frac{e}{c}\right)^c. \end{aligned}$$

The expression of the Bregman information then follows immediately:

$$\gamma_{n,c}(\theta) = -\log\left(\frac{n+c}{(-1-\theta)L_n + c}\right) - n - c \log c + (n+c) \log(n+c) + \log \Gamma(c) - \log \Gamma(n+c).$$

Moreover, we deduce from the expression of  $\theta_{n,c}(\theta)$  and the Bregman divergence that:

$$\mathcal{B}_{\mathcal{L}}(\theta', \theta_{n,c}(\theta)) = -\log(-1-\theta') + \log\left(\frac{n+c}{L_n - \frac{c}{1+\theta}}\right) - \frac{(L_n - \frac{c}{1+\theta})(\theta' + 1)}{n+c} - 1.$$

Therefore, Theorem 3 combined with the natural parameter  $\theta = -1 - \alpha$  yields that w.p.  $\geq 1 - \delta$ ,  $\forall n \in \mathbb{N}$ ,

$$\alpha L_n - (n+c+1) \log(\alpha L_n + c) \leq \log \frac{1}{\delta} - c \log c - \log(n+c) + \log \Gamma(c) - \log \Gamma(n+c). \quad (11)$$

### C.9. Chi-square

We finally consider  $X \sim \chi^2(k)$  or, equivalently,  $X \sim \text{Gamma}\left(\frac{k}{2}, \frac{1}{2}\right)$ , i.e.,  $p_k(x) = \frac{(\frac{1}{2})^{\frac{k}{2}}}{\Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}$ ,  $x \geq 0$ ,  $k \in \mathbb{N}$  (or  $k \in \mathbb{R}_+ \setminus \{0\}$  if one considers Gamma distributions). This corresponds to an exponential family model with parameter  $\theta = \frac{k}{2} - 1$ , feature function  $F(x) = \log x$  and log-partition function  $\mathcal{L}(\theta) = (\theta + 1) \log 2 + \log \Gamma(\theta + 1)$ . The Bregman divergence between two parameters  $\theta'$  and  $\theta$  associated with  $\mathcal{L}$  is given by

$$\mathcal{B}_{\mathcal{L}}(\theta', \theta) = \text{KL}(P_k, P_{k'}) = \frac{1}{2}(k - k')\psi_0(k/2) - \log \frac{\Gamma(k/2)}{\Gamma(k'/2)},$$

---

4. We assume scale is fixed to the value 1.

where  $\psi_0(x) = \frac{d}{dx} \log \Gamma(x)$  denotes the digamma function.

We further have  $\mathcal{L}'(\theta) = \log 2 + \psi_0(\theta+1)$  and  $\mathcal{L}''(\theta) = \psi_1(\theta+1)$ , where  $\psi_1(x) = \frac{d^2}{dx^2} \log \Gamma(x)$  denotes the trigamma function. Therefore  $\mathcal{L}'$  is invertible, and the parameter estimate is given by

$$\log 2 + \psi_0(\theta_{n,c}(\theta) + 1) = \frac{1}{n+c} \sum_{t=1}^n \log X_t + \frac{c}{n+c} (\log 2 + \psi_0(\theta + 1)),$$

yielding

$$\begin{aligned} k_{n,c}(k) &:= 2(1 + \theta_{n,c}(\theta)) = 2\psi_0^{-1} \left( \frac{1}{n+c} \sum_{t=1}^n \log X_t + \frac{c}{n+c} \psi_0(k/2) - \frac{n}{n+c} \log 2 \right) \\ &= 2\psi_0^{-1} \left( \frac{1}{n+c} K_n + \frac{c}{n+c} \psi_0(k/2) \right), \end{aligned}$$

where  $K_n = \sum_{t=1}^n \log \frac{X_t}{2}$ . Therefore, the Bregman divergence rewrites as follows

$$\begin{aligned} (n+c) \cdot \mathcal{B}_{\mathcal{L}}(\theta, \theta_{n,c}(\theta)) &= (n+c) \cdot \text{KL}(P_{k_{n,c}(k)}, P_k) \\ &= \frac{1}{2} (k_{n,c}(k) - k) \left( K_n + c \psi_0 \left( \frac{k}{2} \right) \right) - (n+c) \log \frac{\Gamma \left( \frac{k_{n,c}(k)}{2} \right)}{\Gamma \left( \frac{k}{2} \right)}. \end{aligned}$$

Now, we see that

$$\begin{aligned} \int_{\mathbb{R}^d} \exp(-c\mathcal{B}_{\mathcal{L}}(\theta', \theta)) d\theta' &= \int \frac{1}{2} \exp \left( -\frac{c}{2}(k - k') \psi_0(k/2) \right) \left( \frac{\Gamma(k/2)}{\Gamma(k'/2)} \right)^c dk', \\ \int_{\mathbb{R}^d} \exp(-(n+c)\mathcal{B}_{\mathcal{L}}(\theta', \theta_{n,c}(\theta))) d\theta' &= \int \frac{1}{2} \exp \left( -\frac{n+c}{2}(k_{n,c}(k) - k') \psi_0(k_{n,c}(k)/2) \right) \left( \frac{\Gamma(k_{n,c}(k)/2)}{\Gamma(k'/2)} \right)^{n+c} dk'. \end{aligned}$$

Therefore, the Bregman information gain writes

$$\begin{aligned} \gamma_{n,c}(\theta) &= c \log \Gamma \left( \frac{k}{2} \right) - (n+c) \log \Gamma \left( \frac{k_{n,c}(k)}{2} \right) - c \frac{k}{2} \psi_0 \left( \frac{k}{2} \right) + (n+c) \frac{k_{n,c}(k)}{2} \psi_0 \left( \frac{k_{n,c}(k)}{2} \right) \\ &\quad + \log \frac{\int (\Gamma(k'/2))^{-c} \exp \left( c \frac{k'}{2} \psi_0 \left( \frac{k}{2} \right) \right) dk'}{\int (\Gamma(k'/2))^{-(n+c)} \exp \left( (n+c) \frac{k'}{2} \psi_0 \left( \frac{k_{n,c}(k)}{2} \right) \right) dk'} \\ &= c \log \Gamma \left( \frac{k}{2} \right) - (n+c) \log \Gamma \left( \frac{k_{n,c}(k)}{2} \right) - c \frac{k}{2} \psi_0 \left( \frac{k}{2} \right) + \frac{k_{n,c}(k)}{2} \left( K_n + c \psi_0 \left( \frac{k}{2} \right) \right) \\ &\quad + \log \frac{\int (\Gamma(k'/2))^{-c} \exp \left( c \frac{k'}{2} \psi_0 \left( \frac{k}{2} \right) \right) dk'}{\int (\Gamma(k'/2))^{-(n+c)} \exp \left( \frac{k'}{2} (K_n + c \psi_0 \left( \frac{k}{2} \right)) \right) dk'} \\ &= c \log \Gamma \left( \frac{k}{2} \right) - (n+c) \log \Gamma \left( \frac{k_{n,c}(k)}{2} \right) - c \frac{k}{2} \psi_0 \left( \frac{k}{2} \right) + \frac{k_{n,c}(k)}{2} \left( K_n + c \psi_0 \left( \frac{k}{2} \right) \right) \\ &\quad + \log J \left( c, c \psi_0 \left( \frac{k}{2} \right) \right) - \log J \left( n+c, K_n + c \psi_0 \left( \frac{k}{2} \right) \right), \end{aligned}$$

where we define the auxiliary function  $J(a, b) = \int \exp\left(-a \log \Gamma\left(\frac{k'}{2}\right) + b \frac{k'}{2}\right) dk'$ . Combining the above with Theorem 3, and after some simple algebra, we obtain that w.p.  $\geq 1 - \delta$ ,  $\forall n \in \mathbb{N}$ ,

$$n \log \Gamma\left(\frac{k}{2}\right) - \frac{k}{2} K_n - \log J\left(c, c\psi_0\left(\frac{k}{2}\right)\right) + \log J\left(n+c, K_n + c\psi_0\left(\frac{k}{2}\right)\right) \leq \log \frac{1}{\delta}. \quad (12)$$

**Remark 9** The integral terms  $\int dk'$  in the above derive from the martingale construction in A.1 and the mixture distribution  $q(\theta|\alpha, \beta)$  over the parameter  $\theta \in \Theta$  of the exponential family. In the case of  $\text{Gamma}\left(\frac{k}{2}, \frac{1}{2}\right)$  with unknown shape  $\frac{k}{2} > 0$ , we have  $\theta = \frac{k}{2}$  and  $\Theta = (0, +\infty)$ , therefore  $dk'$  corresponds to the Lebesgue measure over  $(0, +\infty)$ . When restricted to the Chi-square family,  $\Theta = \mathbb{N}$ , and  $dk'$  is instead the counting measure, effectively turning integrals into discrete sums. In Appendix D, we report figures using both versions, see Figure 6 and Figure 11.

**Remark 10** The ratio of integrals (or infinite sums) in (12) can be efficiently implemented using a simple integration scheme (or by truncation). Indeed, for a given  $k_{\max} \in \mathbb{N}$ ,  $b \ll 1$  and  $B \gg 1$ , we define  $x_{k'} = b + \frac{k'}{k_{\max}}(B - b)$  for  $k' = 0, \dots, k_{\max}$ , so that

$$\begin{aligned} & \log \frac{\int_0^{+\infty} \left(\Gamma\left(\frac{k'}{2}\right)\right)^{-c} \exp\left(\frac{k'}{2} c\psi_0\left(\frac{k}{2}\right)\right) dk'}{\int_0^{+\infty} \left(\Gamma\left(\frac{k'}{2}\right)\right)^{-(n+c)} \exp\left(\frac{k'}{2} (c\psi_0\left(\frac{k}{2}\right) + \sum_{t=1}^n \log \frac{X_t}{2})\right) dk'} \\ & \approx \underset{k'=1, \dots, k_{\max}}{\text{logsumexp}} \left( -c \log \Gamma\left(\frac{x_{k'}}{2}\right) + \frac{cx_{k'}}{2} \psi_0\left(\frac{k}{2}\right) + \log(x_{k'} - x_{k'-1}) \right) \\ & \quad - \underset{k'=1, \dots, k_{\max}}{\text{logsumexp}} \left( -(n+c) \Gamma\left(\frac{x_{k'}}{2}\right) + \frac{x_{k'}}{2} \left( c\psi_0\left(\frac{k}{2} + \sum_{t=1}^n \log \frac{X_t}{2}\right) \right) + \log(x_{k'} - x_{k'-1}) \right). \end{aligned}$$

Similarly, we have

$$\begin{aligned} & \log \frac{\sum_{k'=1}^{+\infty} \left(\Gamma\left(\frac{k'}{2}\right)\right)^{-c} \exp\left(\frac{k'}{2} c\psi_0\left(\frac{k}{2}\right)\right)}{\sum_{k'=1}^{+\infty} \left(\Gamma\left(\frac{k'}{2}\right)\right)^{-(n+c)} \exp\left(\frac{k'}{2} (c\psi_0\left(\frac{k}{2}\right) + \sum_{t=1}^n \log \frac{X_t}{2})\right)} \\ & \approx \underset{k'=1, \dots, k_{\max}}{\text{logsumexp}} \left( -c \log \Gamma\left(\frac{k'}{2}\right) + \frac{k'}{2} c\psi_0\left(\frac{k}{2}\right) \right) \\ & \quad - \underset{k'=1, \dots, k_{\max}}{\text{logsumexp}} \left( -(n+c) \Gamma\left(\frac{k'}{2}\right) + \frac{k'}{2} \left( c\psi_0\left(\frac{k}{2} + \sum_{t=1}^n \log \frac{X_t}{2}\right) \right) \right), \end{aligned}$$

The final steps correspond to the right-rectangular scheme over  $(b, B)$  with  $k_{\max}$  steps and the truncation to the first  $k_{\max}$  terms respectively. Note the use of  $\text{logsumexp}(z) = \log \sum_{k'=1}^{k_{\max}} \exp(z_{k'})$  for  $z \in \mathbb{R}^{k_{\max}}$ , which is efficiently implemented in many libraries for scientific computing and better handles summation of large numbers. Empirically, we found that  $\log b = -10$ ,  $\log B = 10$  and  $k_{\max} = 2000$  provided sufficient accuracy and that using finer approximation schemes did not significantly impact the numerical results.

### C.10. Poisson

We consider  $X \sim \text{Poisson}(\lambda)$ , where  $\lambda > 0$  is unknown. We recall that the distribution of  $X$  is supported on  $\mathbb{N}$  with probability mass function  $\mathbb{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$ . This corresponds to a one-dimensional exponential family model with parameter  $\theta = \log \lambda \in \mathbb{R}$ , feature function  $F(x) = x$  and log-partition function  $\mathcal{L}(\theta) = e^\theta$  (which is invertible on  $\mathbb{R}$ ). The Bregman divergence between two parameters  $\theta'$  and  $\theta$  is therefore  $\mathcal{B}_{\mathcal{L}}(\theta', \theta) = e^{\theta'} - e^\theta - (\theta' - \theta)e^\theta$ .

Using the shorthand  $S_n = \sum_{t=1}^n X_t$ , it follows from the definition that:

$$\theta_{n,c}(\theta) = \log \left( \frac{S_n + ce^\theta}{n+c} \right).$$

The Bregman information gain is expressed using the auxiliary function  $I(a, b) = \int_{-\infty}^{+\infty} e^{-ae^\theta + b\theta} d\theta$  as:

$$\begin{aligned} \gamma_{n,c}(\theta) &= c(1 - \theta)e^\theta - (n+c)(1 - \theta_{n,c}(\theta))e^{\theta_{n,c}(\theta)} + \log I(c, ce^\theta) - \log I(n+c, (n+c)e^{\theta_{n,c}(\theta)}) \\ &= c(1 - \theta)e^\theta - \left( 1 - \log \left( \frac{S_n + ce^\theta}{n+c} \right) \right) (S_n + ce^\theta) + \log I(c, ce^\theta) - \log I(n+c, S_n + ce^\theta). \end{aligned}$$

Moreover, we deduce from the expression of  $\theta_{n,c}(\theta)$  and the Bregman divergence that:

$$\mathcal{B}_{\mathcal{L}}(\theta', \theta_{n,c}(\theta)) = e^{\theta'} - \left( \frac{S_n + ce^\theta}{n+c} \right) - \left( \theta' - \log \left( \frac{S_n + ce^\theta}{n+c} \right) \right) \left( \frac{S_n + ce^\theta}{n+c} \right).$$

Therefore, Theorem 3 combined with the natural parametrization  $\theta = \log \lambda$  yields that w.p.  $\geq 1 - \delta$ ,  $\forall n \in \mathbb{N}$ ,

$$n\lambda - S_n \log \lambda \leq \log \frac{1}{\delta} + \log I(c, c\lambda) - \log I(n+c, S_n + c\lambda).$$

**Remark 11** *Although, to the best of our knowledge, the integral  $I(a, b)$  does not have a closed-form expression, it can be numerically estimated up to arbitrary precision. We recommend the same implementation as discussed in Remark 10, using the logsumexp operator for stability, and refer to the code for further details.*

### C.11. Summary table

We summarize the high probability confidence sets of Theorem 3 derived in this Table 2.

Table 2: Summary of Bregman confidence sets given by Theorem 3 for representative families. Throughout, the following notations are used:

$$\begin{aligned}
 S_n &= \sum_{t=1}^n X_t, \quad \hat{\mu}_n = \frac{S_n}{n}, \quad Q_n(\mu) = \sum_{t=1}^n (X_t - \mu)^2, \quad Z_n(\mu, \sigma) = \frac{Q_n(\mu)}{\sigma^2}, \\
 S_n^{(k)} &= \sum_{t=1}^n X_t^k, \quad L_n = \sum_{t=1}^n \log X_t, \quad K_n = \sum_{t=1}^n \log \frac{X_t}{2}, \\
 I(a, b) &= \int_{-\infty}^{+\infty} e^{-ae^\theta + b\theta} d\theta, \\
 J(a, b) &= \int_0^\infty \exp(-a \log \Gamma(\frac{k}{2}) + \frac{bk}{2}) dk \text{ (}dk \text{ is the Lebesgue measure if } k \in \mathbb{R}_+ \text{ and the counting measure if } k \in \mathbb{N}\text{).}
 \end{aligned}$$

Name	Parameters	Formula
Gaussian	$\mu \in \mathbb{R}$	$\frac{1}{n+c} \frac{(S_n - n\mu)^2}{2\sigma^2} \leq \log \frac{1}{\delta} + \frac{1}{2} \log \frac{n+c}{c}$
Gaussian	$\sigma \in \mathbb{R}_+$	$  \begin{aligned}  &\frac{Q_n(\mu)}{2\sigma^2} - \left(\frac{n+c}{2} + 1\right) \log \left(\frac{Q_n(\mu) + c}{2\sigma^2}\right) \\  &\leq \log \frac{1}{\delta} + \log \frac{\Gamma(\frac{c+2}{2})}{\Gamma(\frac{n+c+2}{2})} - \frac{n}{2} \log 2 - \left(\frac{c}{2} + 1\right) \log c + \left(\frac{n+c}{2} + 1\right) \log(n+c)  \end{aligned}  $
	$\mu \in \mathbb{R}$ $\sigma \in \mathbb{R}_+$	$  \begin{aligned}  &\frac{1}{2} Z_n(\mu, \sigma) - \frac{n+c+3}{2} \log \left(\frac{n}{n+c} Z_n(\hat{\mu}_n, \sigma) + \frac{c}{n+c} Z_n(\mu, \sigma) + c\right) \\  &\leq \log \frac{1}{\delta} - \frac{n}{2} \log 2 - \left(\frac{c}{2} + 2\right) \log c + \frac{1}{2} \log(n+c) \\  &\quad + \log \Gamma\left(\frac{c+3}{2}\right) - \log \Gamma\left(\frac{n+c+3}{2}\right)  \end{aligned}  $
Bernoulli	$\mu \in [0, 1]$	$  \begin{aligned}  &S_n \log \frac{1}{\mu} + (n - S_n) \log \frac{1}{1-\mu} + \log \frac{\Gamma(S_n + c\mu) \Gamma(n - S_n + c(1-\mu))}{\Gamma(c\mu) \Gamma(c(1-\mu))} \\  &\leq \log \frac{1}{\delta} + \log \frac{\Gamma(n+c)}{\Gamma(c)}  \end{aligned}  $
Exponential	$\mu \in \mathbb{R}_+$	$  \begin{aligned}  &\frac{S_n}{\mu} - (n+c+1) \log \left(\frac{S_n}{\mu} + c\right) \\  &\leq \log \frac{1}{\delta} + \log \frac{\Gamma(c)}{\Gamma(n+c)} - \log(n+c) - c \log c  \end{aligned}  $
Gamma	$\lambda \in \mathbb{R}_+$	$  \begin{aligned}  &\frac{S_n}{\lambda} - ((n+c)k + 1) \log \left(\frac{S_n}{\lambda} + ck\right) \\  &\leq \log \frac{1}{\delta} + \log \frac{\Gamma(ck)}{\Gamma((n+c)k)} - \log((n+c)k) - ck \log ck  \end{aligned}  $
Weibull	$\lambda \in \mathbb{R}_+$	$  \begin{aligned}  &\frac{S_n^{(k)}}{\lambda^k} - (n+c+1) \log \left(\frac{S_n^{(k)}}{\lambda^k} + c\right) \\  &\leq \log \frac{1}{\delta} + \log \frac{\Gamma(c)}{\Gamma(n+c)} - \log(n+c) - c \log c  \end{aligned}  $
Pareto	$\alpha \in \mathbb{R}$	$  \begin{aligned}  &\alpha L_n - (n+c+1) \log(\alpha L_n + c) \\  &\leq \log \frac{1}{\delta} + \log \frac{\Gamma(c)}{\Gamma(n+c)} - \log(n+c) - c \log c  \end{aligned}  $
Poisson	$\lambda \in \mathbb{R}_+$	$  \begin{aligned}  &n\lambda - S_n \log \lambda \\  &\leq \log \frac{1}{\delta} + \log I(c, c\lambda) - \log I(n+c, S_n + c\lambda)  \end{aligned}  $
Chi-square	$k \in \mathbb{N}$ or $k \in \mathbb{R}_+$	$  \begin{aligned}  &n \log \Gamma\left(\frac{k}{2}\right) - \frac{k}{2} K_n - \log J\left(c, c\psi_0\left(\frac{k}{2}\right)\right) \\  &\quad + \log J\left(n+c, K_n + c\psi_0\left(\frac{k}{2}\right)\right) \leq \log \frac{1}{\delta}  \end{aligned}  $

## Appendix D. Empirical Comparison with Existing Time-uniform Confidence Sequences

We illustrate the time-uniform confidence sequences derived from Bregman concentration on several instances of classical exponential families detailed in 3.2. In each setting, when available, we also report confidence sequences based on existing methods in the literature<sup>5</sup>.

In what follows, we fix  $\delta \in (0, 1)$  the uniform confidence level. For each confidence sequence  $(\Theta_n)_{n \in \mathbb{N}}(\delta)$ , we report in the figures the intersection sequence  $(\cap_{k \leq n} \Theta_k(\delta))_{n \in \mathbb{N}}$ , which also holds with confidence  $1 - \delta$ , for  $n$  up to 200. Typical realizations of Bregman confidence sequences are reported in Figure 3 (Gaussian), Figure 4 (Bernoulli, Poisson), Figure 5 (Exponential, Gamma, Weibull, Pareto), and Figure 6 (Chi-square).

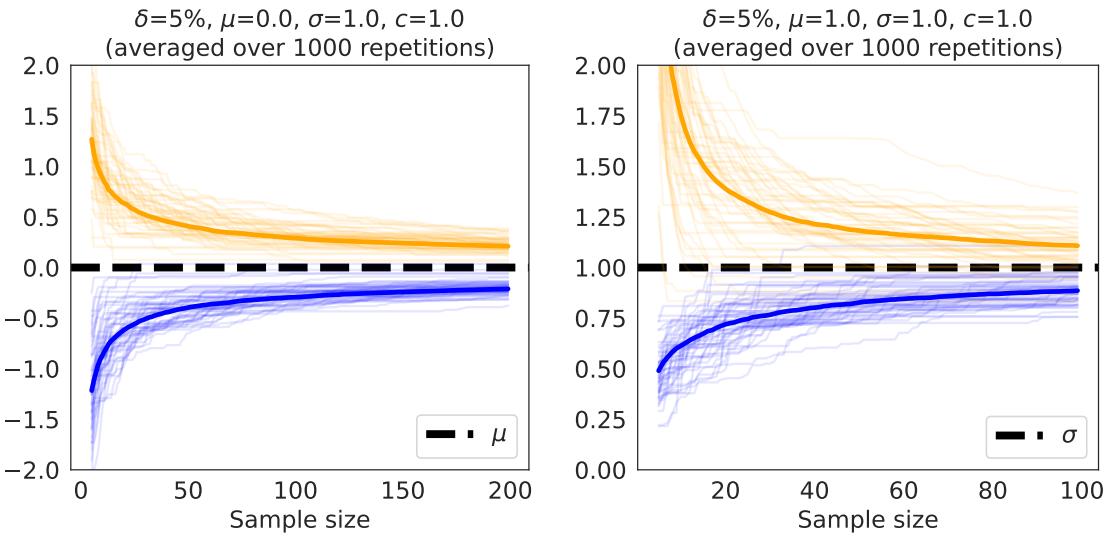


Figure 3: Examples of confidence upper and lower envelopes around unknown expectation  $\mu = 0$  for  $\mathcal{N}(\mu, 1)$  (left, cf. Table 2) and unknown standard deviation  $\sigma = 1$  for  $\mathcal{N}(1, \sigma)$  (right, cf. Table 2), as a function of the number of observations  $n$ . The thick lines indicate the median curve over 1000 replicates.

---

5. The code is provided in the supplementary material for reproducibility.

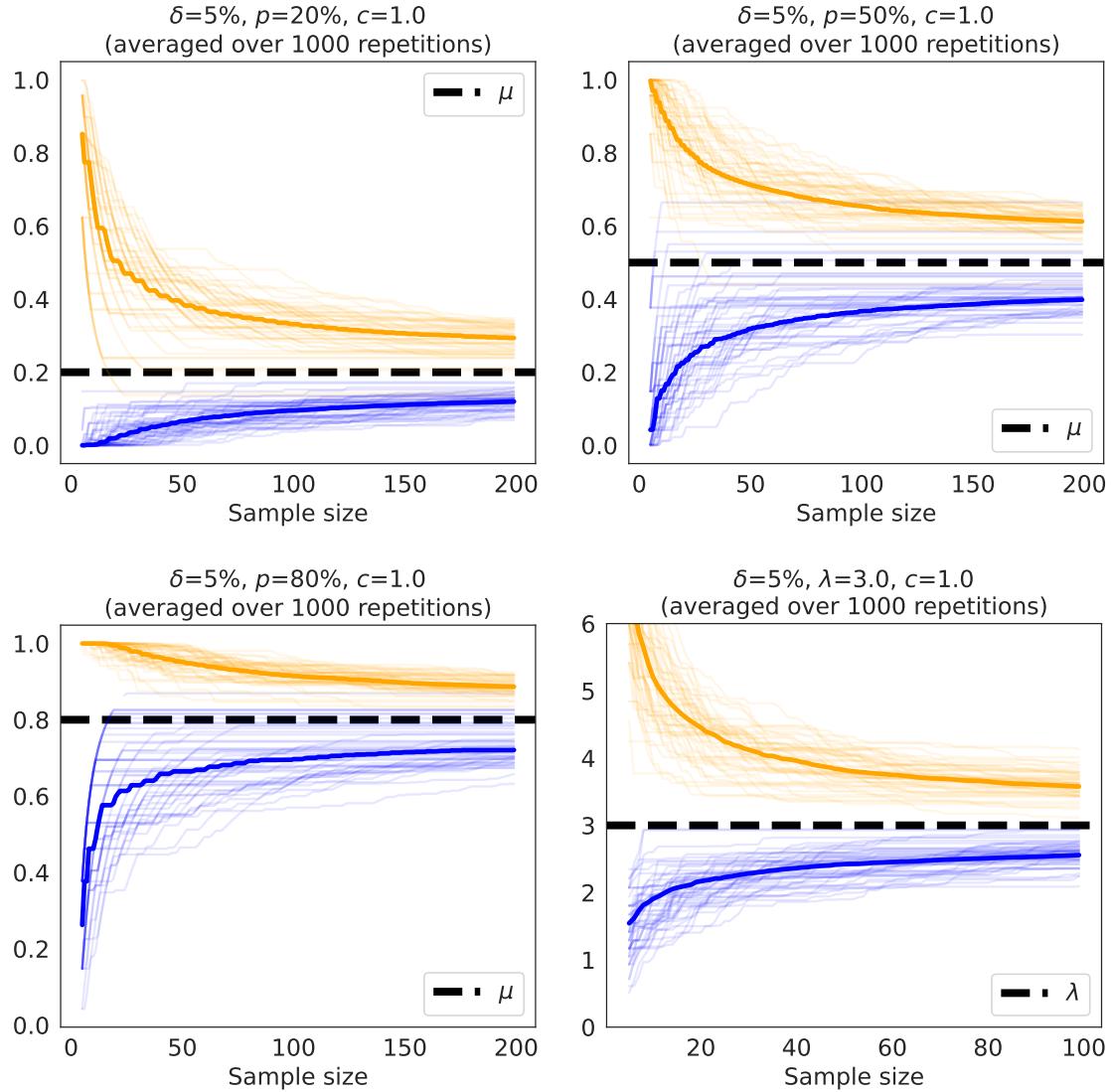


Figure 4: Examples of confidence upper and lower envelopes around expectation  $\mu$  for discrete distributions. From top to bottom: Bernoulli( $p$ ),  $p \in \{0.2, 0.5, 0.8\}$  (cf. Table (2)) and Poisson(3) (cf. Table 2) on several realizations (each dashed lines) as a function of the number of observations  $n$ . The thick lines indicate the median curve over 1000 replicates.

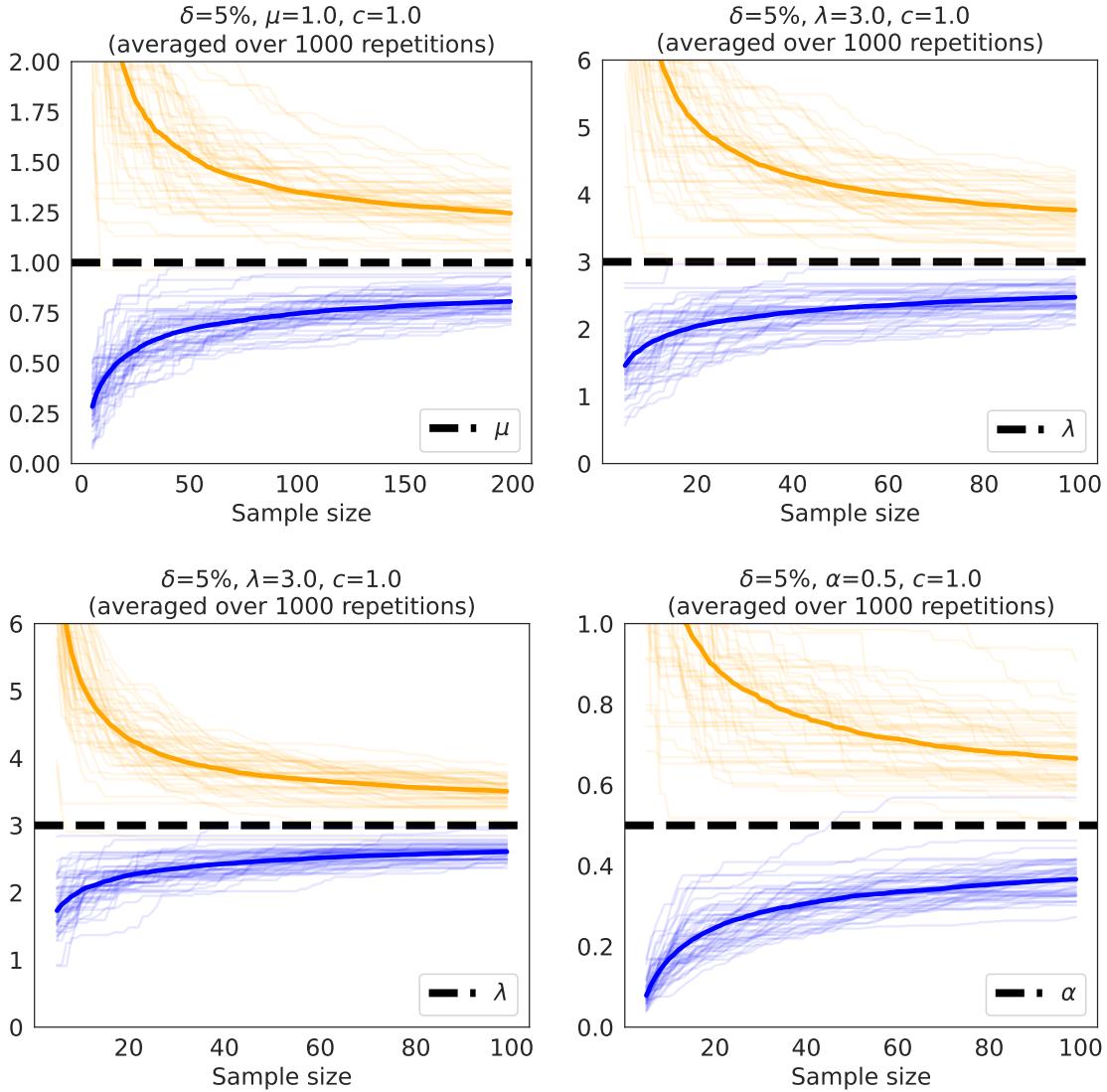


Figure 5: Examples of confidence upper and lower envelopes (from top to bottom) around scale parameter  $\lambda$  for  $\text{Exp}(1)$  (cf. Table (2)), fixed shape  $\text{Gamma}(3, 2)$  (cf. Table 2), fixed shape  $\text{Weibull}(3, 2)$  (cf. Table 2), and exponent  $\alpha$  for  $\text{Pareto}(\frac{1}{2})$  (cf. Table 2), on several realizations (each dashed lines) as a function of the number of observations  $n$ . The thick lines indicate the median curve over 1000 replicates.

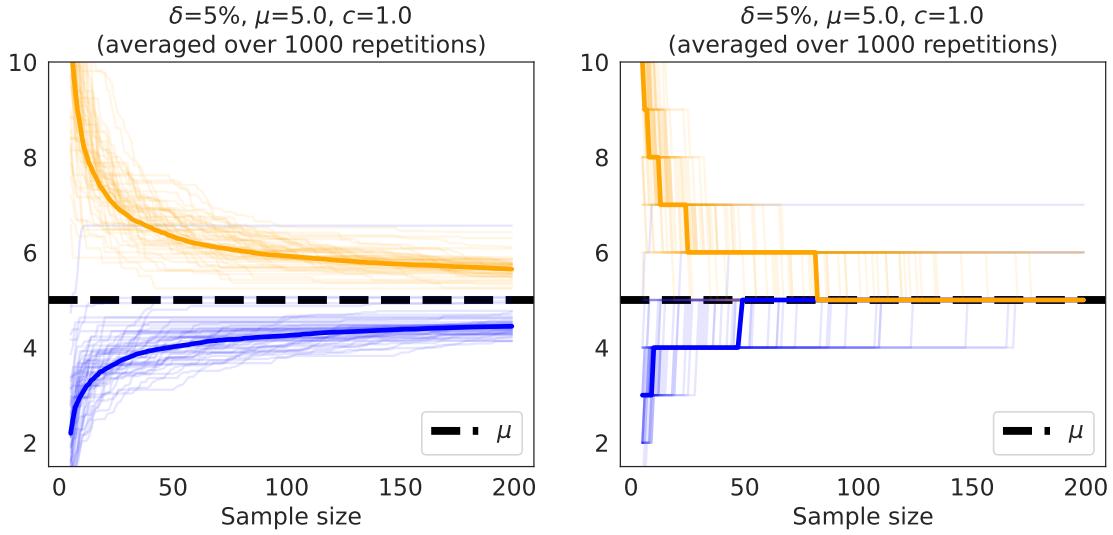


Figure 6: Examples of confidence upper and lower envelopes around expectation  $\mu$  for (continuous prior) Gamma  $(\frac{k}{2})$  with unknown shape  $\frac{k}{2} > 0$  and (discrete prior) Chi-square  $\chi^2(k)$ ,  $k \in \mathbb{N}^*$  (cf. Table (2) and Remark 9) on several realizations (each dashed lines), as a function of the number of observations  $n$ . In particular for the Chi-square, confidence lower and upper bounds are ceiled and floored to integers. The thick lines indicate the median curve over 1000 replicates.

#### D.1. Gaussian with unknown mean and variance

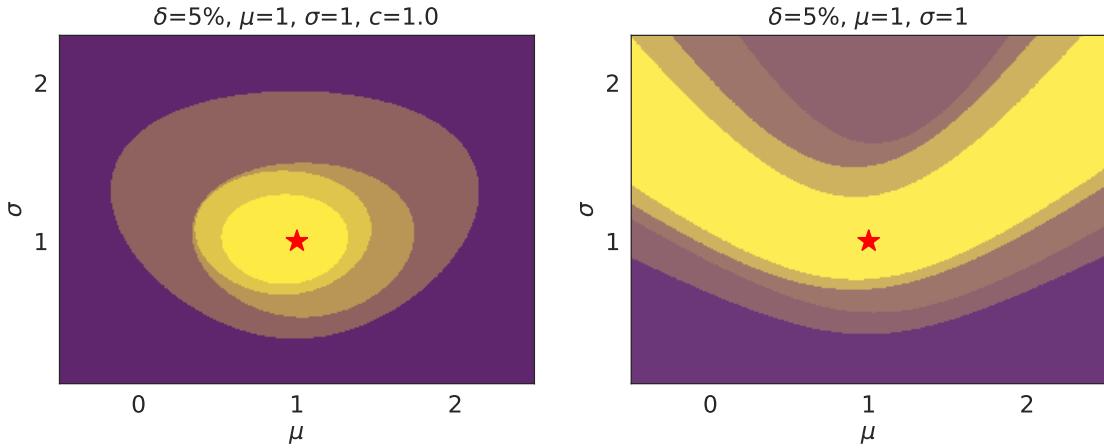


Figure 7: Example of time-uniform, joint confidence sets around  $(\mu, \sigma)$  for  $\mathcal{N}(1, 1)$  (cf. Table 2 for  $n \in \{10, 25, 50, 100\}$  observations (smaller confidence sets correspond to larger sample sizes). The red star indicates the true parameters  $(\mu, \sigma) = (1, 1)$ .

We consider the two-dimensional family  $\{\mathcal{N}(\mu, \sigma^2), (\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_+^*\}$ . To the best of our knowledge, there does not exist time-uniform joint confidence sets for  $(\mu, \sigma)$  prior to this work. In order to compare ourselves against some baseline, we derive a crude one based on Chi-square quantiles and a union bound.

Let  $n \in \mathbb{N}, \delta \in (0, 1)$ ,  $X_1, \dots, X_n$  i.i.d samples from  $\mathcal{N}(\mu_0, \sigma_0^2)$ . The standardized sum of squares

$$Z_t(\mu_0, \sigma_0) = \sum_{s=1}^t \left( \frac{X_s - \mu_0}{\sigma_0} \right)^2$$

follows a  $\chi^2(t)$  distribution, therefore, denoting by  $q_{\chi^2(t)}$  the corresponding quantile function, we have:

$$\mathbb{P} \left( q_{\chi^2(t)} \left( \frac{\delta}{2} \right) \leq Z_t(\mu_0, \sigma_0) \leq q_{\chi^2(t)} \left( 1 - \frac{\delta}{2} \right) \right) \geq 1 - \delta.$$

By a union bound argument, the intersection of  $n$  such events with confidence  $\frac{\delta}{n}$ , i.e

$$\Theta_{t,n}^Z(\delta) = \bigcap_{s=1}^t \left\{ (\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_+^*: q_{\chi^2(s)} \left( \frac{\delta}{2n} \right) \leq Z_s(\mu, \sigma) \leq q_{\chi^2(s)} \left( 1 - \frac{\delta}{2n} \right) \right\}$$

describes a sequence of confidence sets at level  $\delta$  that hold uniformly over  $t \in \{1, \dots, n\}$ .

We report our confidence sets (cf. Equation 2) and the above on Figure 7. The most striking drawback of  $\Theta_{t,n}^Z(\delta)$  is that it is not convex nor even bounded; in particular, projecting onto the axes of  $\mu$  and  $\sigma$  only provides trivial confidence sets,  $\mathbb{R}$  and  $\mathbb{R}_+^*$  respectively, rendering this result vacuous. To better grasp this phenomenon, let us informally consider  $\mu = \alpha + \mu_0$  and  $\sigma = \alpha$  for some  $\alpha > 0$ . We have:

$$\begin{aligned} Z_n(\mu, \sigma) &= \sum_{t=1}^n \left( \frac{X_t - \mu_0}{\alpha} - 1 \right)^2 \\ &= \underbrace{\frac{1}{\alpha^2} \sum_{t=1}^n (X_t - \mu_0)^2}_{\approx \frac{\sigma_0^2}{\alpha^2} n} - \underbrace{\frac{2}{\alpha} \sum_{t=1}^n (X_t - \mu_0)}_{\approx \frac{2}{\alpha} \sqrt{n}} + n. \end{aligned}$$

Hence, for  $\alpha \rightarrow +\infty$ ,  $Z_n(\mu_0 + \alpha, \alpha) \approx n$ . On the other hand, Lemma 1 in [Laurent and Massart \(2000\)](#) shows that  $q_{\chi^2(n)}(1 - \delta/2n) = \mathcal{O}(n + \sqrt{n \log n})$  and  $q_{\chi^2(s)}(\delta/2n) = \mathcal{O}(n)$ . Therefore, even with increasing the sample size  $n$ , there exists arbitrary large  $\alpha$  such that  $(\mu_0 + \alpha, \alpha)$  may belong to  $\Theta_{t,n}^Z(\delta)$ .

By contrast, our Bregman confidence sets are convex and bounded, which we interpret as the result of exploiting the true geometry of the two-dimensional Gaussian family. In addition to the unboundedness,  $\Theta_{t,n}^Z(\delta)$  is built using a crude union bound, which is not anytime (depends on the terminal time  $n$ ) and rather loose.

## D.2. Bernoulli

We consider the Bernoulli distribution  $\text{Bernoulli}(\mu)$  for some unknown  $\mu \in [0, 1]$ . Confidence bounds are displayed in Figure 8 for  $\mu \in \{0.2, 0.5, 0.8\}$ . We detail below a few alternative state-of-the-art bounds.

**Chernoff-Laplace method for sub-Gaussian distributions** We recall that  $B(\mu)$  being supported in  $[0, 1]$ , it is  $1/2$ -sub-Gaussian (Hoeffding lemma, [Hoeffding \(1963\)](#)). The method of mixtures for exponential martingales ([Peña et al., 2008](#)) shows that  $\mathbb{P}(\forall n \in \mathbb{N}, \mu \in \Theta_n^{\text{Laplace}}(\delta)) \geq 1 - \delta$ , with

$$\Theta_n^{\text{Laplace}}(\delta) = \left[ \widehat{\mu}_n - \sqrt{\left(1 + \frac{1}{n}\right) \frac{\log(2\sqrt{1+n}/\delta)}{2n}}, \widehat{\mu}_n + \sqrt{\left(1 + \frac{1}{n}\right) \frac{\log(2\sqrt{1+n}/\delta)}{2n}} \right].$$

Note that a sharper sub-Gaussian control holds for  $B(\mu)$ , see [Kearns and Saul \(1998\)](#), [Berend and Kontorovich \(2013\)](#). However, the optimal sub-Gaussian parameter they suggest ( $\frac{1/2-\mu}{\log(1/\mu-1)}$  instead of  $1/2$ ) is a function of the unknown  $\mu$ .

**Bentkus concentration with geometric time-peeling** [Kuchibhotla and Zheng \(2021, Theorem 4\)](#) show that  $\mathbb{P}(\forall n \in \mathbb{N}, \mu \in \Theta_n^{\text{Bentkus}}(\delta)) \geq 1 - \delta$  with

$$\Theta_n^{\text{Bentkus}}(\delta) = \left[ \widehat{\mu}_n - \frac{1}{n} q\left(\frac{\delta}{2h(k_n)}; c_n, 1/2, 1\right), \widehat{\mu}_n + \frac{1}{n} q\left(\frac{\delta}{2h(k_n)}; c_n, 1/2, 1\right) \right].$$

Here  $q(\delta'; N, A, B)$  is the Bentkus bound, where  $\delta'$  is the confidence level,  $N$  is the sample size,  $B$  is the *almost sure* upper bound of the random variables and  $A^2$  is the variance upper bound. The function  $h(k)$  is defined as  $h(k) = \zeta(1.1)(k+1)^{1.1}$ , where  $\zeta(\cdot)$  is the Riemann-zeta function. The integer  $k_n$  and the real  $c_n$  are defined as  $k_n = \min\{k \in \mathbb{N} : \lceil \eta^k \rceil \leq n \leq \lfloor \eta^{k+1} \rfloor\}$  and  $c_n = \lfloor \eta^{k_n+1} \rfloor$ , with  $\eta = 1.1$ . The referenced theorem is stated with an empirical overestimate of the standard deviation instead of the fixed bound  $1/2$ , which would lead to replace the  $\delta/2$  present in the confidence set with  $\delta/3$  by union bound; we found this step to be of negligible, and even slightly detrimental impact, in the case of Bernoulli distributions.

**Hedged Capital martingale method** [Waudby-Smith and Ramdas \(2023, Theorem 3\)](#) present a nonnegative martingale construction from observations  $X_1, \dots, X_n$  bounded in  $[0, 1]$ . Following the recommendations of the authors, we define the following quantities:

$$\begin{aligned} \mathcal{K}_n(m) &= \max(\mathcal{K}_n^+, \mathcal{K}_n^-(m)), \\ \mathcal{K}_n^\pm(m) &= \prod_{k=1}^n (1 \pm \lambda_k^\pm(m) (X_k - m)), \\ \lambda_k^+(m) &= \min\left(|\lambda_k|, \frac{1/2}{m}\right), \lambda_k^-(m) = \min\left(|\lambda_k|, \frac{1/2}{1-m}\right), \\ \lambda_k &= \sqrt{\frac{2 \log(2/\delta)}{\widehat{\sigma}_{k-1}^2 k \log(k+1)}}, \\ \widehat{\sigma}_k^2 &= \frac{1/4 + \sum_{i=1}^k (X_i - \widehat{\mu}_i)^2}{k+1}, \\ \widehat{\mu}_k &= \frac{1/2 + \sum_{i=1}^k X_i}{k+1}. \end{aligned}$$

Then, it holds that  $\mathbb{P}(\forall n \in \mathbb{N}, \mu \in \Theta_n^{\text{HC}}(\delta)) \geq 1 - \delta$ , where  $\Theta_n^{\text{HC}}(\delta) = \{m \in [0, 1] : \mathcal{K}_n(m) < 1/\delta\}$ .

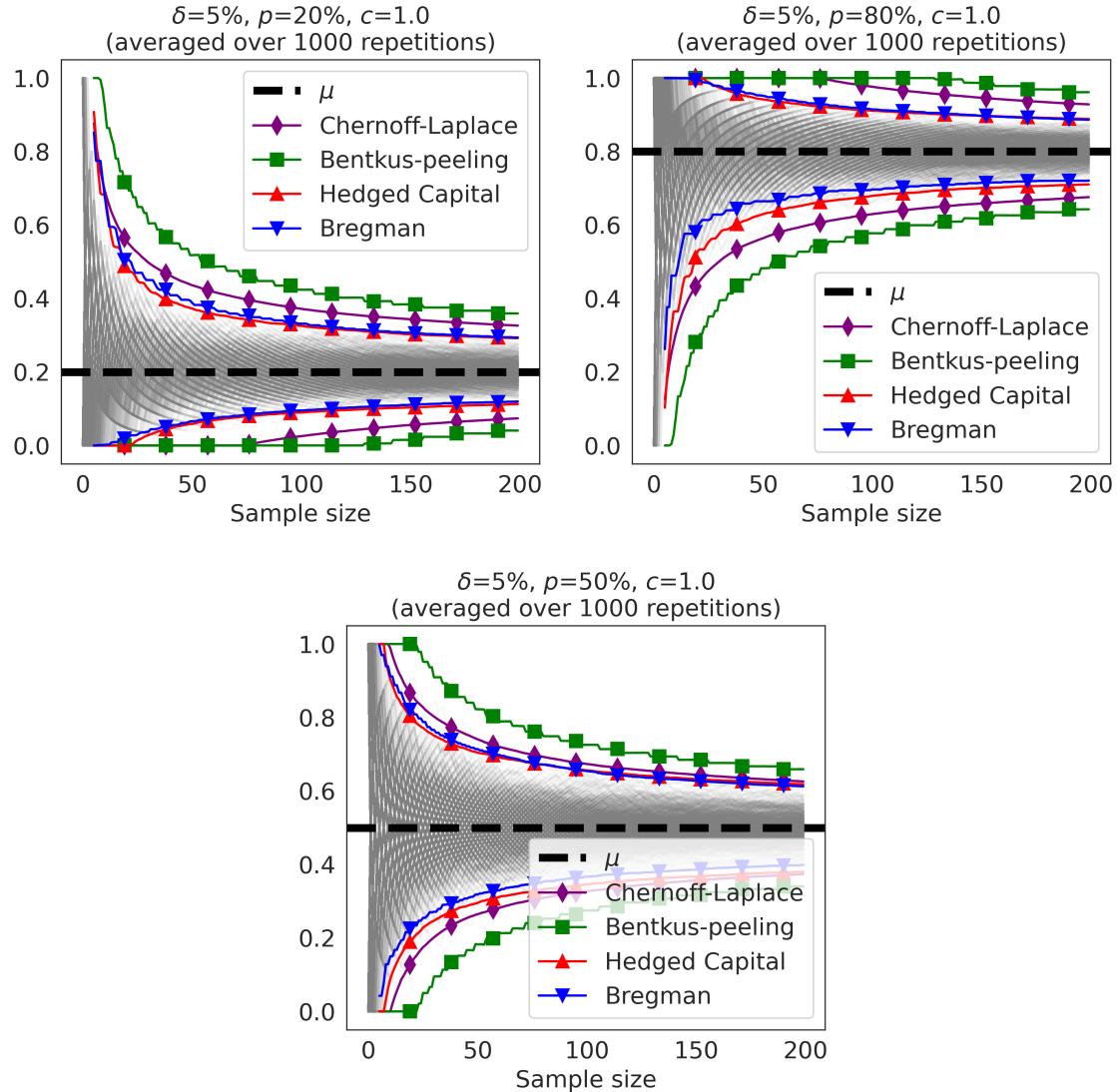


Figure 8: Confidence upper and lower envelopes built for  $\text{Bernoulli}(\mu)$ ,  $\mu \in \{0.2, 0.5, 0.8\}$ , as a function of the number of observations  $n$ , averaged over 1000 independent simulations. Grey lines are trajectories of empirical means  $\hat{\mu}_n$ . Confidence bounds are clipped between 0 and 1.

### D.3. Gaussian

We consider Gaussian distributions  $\mathcal{N}(\mu, \sigma^2)$  for some unknown  $\mu \in \mathbb{R}$  and known variance  $\sigma^2$ . The confidence bounds are displayed in Figure 9 for  $\mu = 0$ .

**Chernoff-Laplace** Similarly to the Bernoulli case,  $\mathcal{N}(\mu, \sigma^2)$  is  $\sigma$ -sub-Gaussian, therefore  $\mathbb{P}(\forall n \in \mathbb{N}, \mu \in \Theta_n^{\text{Laplace}}(\delta)) \geq 1 - \delta$ , with

$$\Theta_n^{\text{Laplace}}(\delta) = \left[ \hat{\mu}_n - \sigma \sqrt{\left(1 + \frac{1}{n}\right) \frac{\log(2\sqrt{1+n}/\delta)}{n}}, \hat{\mu}_n + \sigma \sqrt{\left(1 + \frac{1}{n}\right) \frac{\log(2\sqrt{1+n}/\delta)}{n}} \right].$$

**Kaufmann-Koolen** [Kaufmann and Koolen \(2021\)](#) introduces a martingale construction for exponential families to derive time-uniform deviation inequalities under bandit sampling. However, application of their result is limited to Gaussian distribution with known variance and Gamma distribution with known shape, which is just a scaled version of exponential distribution. Restricting Corollary 10 of [Kaufmann and Koolen \(2021\)](#) to the case of a single arm yields  $\mathbb{P}(\forall n \in \mathbb{N}, \mu \in \Theta_n^{\text{KK}}(\delta)) \geq 1 - \delta$ , with

$$\begin{aligned} \Theta_n^{\text{KK}}(\delta) &= \left\{ \mathcal{B}_{\mathcal{L}}(\hat{\mu}_n, \mu) \leq \frac{2}{n} \log(4 + \log(n)) + \frac{1}{n} C^g(\log 1/\delta) \right\}, \\ g: \lambda \in (1/2, 1] &\mapsto 2\lambda(1 - \log(4\lambda)) + \log \zeta(2\lambda) - \frac{1}{2} \log(1 - \lambda), \\ C^g: x \in (0, +\infty) &\mapsto \min_{\lambda \in (1/2, 1]} \frac{g(\lambda) + x}{\lambda}, \\ \mathcal{L}(\theta) &= \frac{\theta^2}{2\sigma^2} \quad (\text{log-partition function of } \mathcal{N}(\cdot, \sigma^2)). \end{aligned}$$

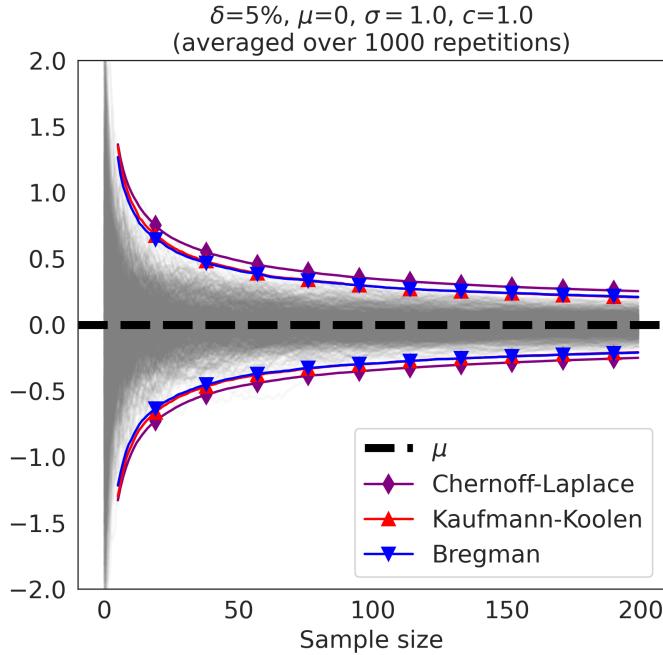


Figure 9: Confidence upper and lower envelopes built for Gaussian distributions  $\mathcal{N}(0, 1)$  as a function of the number of observations  $n$ , averaged over 1000 independent simulations. Grey lines are trajectories of empirical means  $\hat{\mu}_n$ .

#### D.4. Exponential

We now consider exponential distributions  $\text{Exp}(1/\mu)$  for some unknown mean  $\mu \in \mathbb{R}$ . We report in Figure 10 the confidence bounds for the case when  $\mu = 1$ .

**Kaufmann-Koolen** [Kaufmann and Koolen \(2021, Corollary 12\)](#) show that  $\mathbb{P}(\forall n \in \mathbb{N}, \mu \in \Theta_n^{\text{KK}}(\delta)) \geq 1 - \delta$ , with the same definition as for the Gaussian case except

$$g: \lambda \in (1/2, 1] \mapsto 2\lambda(1 - \log(4\lambda)) + \log \zeta(2\lambda) - \log(1 - \lambda), \quad (13)$$

$$\mathcal{L}(\theta) = \log(-1/\mu) \quad (\text{log-partition function of } \mathcal{E}(1/\mu)). \quad (14)$$

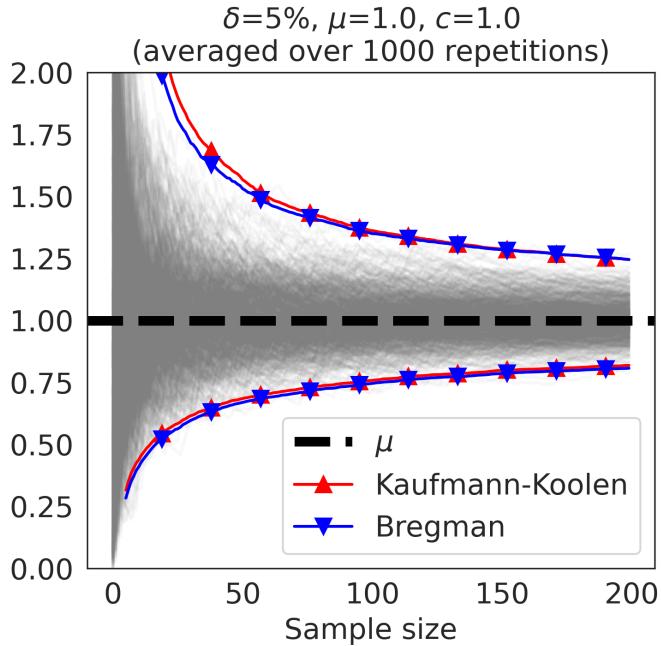


Figure 10: Confidence upper and lower envelopes built for Exponential distributions  $\text{Exp}(1)$  as a function of the number of observations  $n$ , averaged over 1000 independent simulations. Grey lines are trajectories of empirical means  $\hat{\mu}_n$ .

#### D.5. Chi-square

We consider Gamma  $(\frac{k}{2}, \frac{1}{2})$  and  $\chi^2(k)$  distributions for some unknown  $k > 0$  and  $k \in \mathbb{N}^*$  respectively. They are in fact the same distributions, however we distinguish both as the restriction on the domain for  $k$  (real or integer) bears two consequences:

- the mixture distribution in the martingale construction is either continuous or discrete, resulting in integrals or sums in the expression of the confidence sequence (Equation 2);
- confidence bounds are constrained to be integers in the discrete case, thus ceiling and flooring the lower and upper bounds respectively.

We find that the former is negligible as both the continuous and discrete mixtures yield the same bound within numerical precision. The latter however allows to drastically shrink the size of the confidence sequences, resulting in perfect identification the mean within 95% confidence with less than 100 observations in half the simulations (see Figure 11).

### D.6. Key Observations

On the studied examples, confidence intervals based on time-uniform Bregman concentration are either comparable with state-of-the art methods for the corresponding setting or result in sharper bounds, especially for small sample sizes (lower bounds for Bernoulli compared to Hedged Capital, upper bound for Exponential compared to Kaufmann-Koolen). In particular, due to the formulation in terms of Bregman divergence, our intervals are naturally asymmetric when the underlying distribution is, and respect the support constraints (for instance Bernoulli Bregman bounds are in  $[0, 1]$  without the need for clipping). Moreover, we also provide bounds in novel settings for which, to the best of our knowledge, time-uniform confidence sets are lacking (Chi-square, Poisson, Weibull, Pareto, mean-variance for Gaussian).

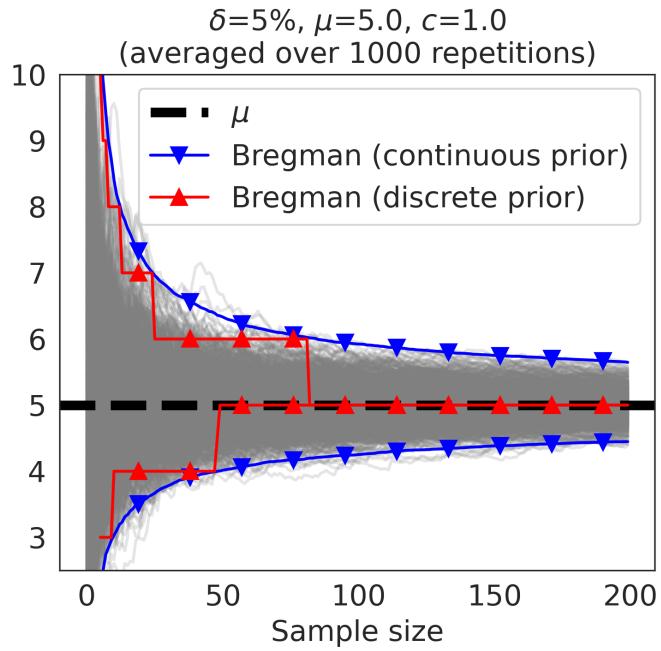


Figure 11: Confidence upper and lower envelopes built for Gamma  $(\frac{5}{2}, \frac{1}{2})$  (blue) and  $\chi^2(5)$  (red) distributions as a function of the number of observations  $n$ , over 1000 independent simulations (colored lines are median of simulated bounds). Grey lines are trajectories of empirical means  $\hat{\mu}_n$ .

### D.7. Tuning of The Regularization Parameter $c$

In this section, we provide additional experiments regarding the local regularization parameter  $c$ . We study the sensitivity of the bounds to the values of  $c$ , reported in Figure 12. Then, we report in Figure 13 plots showing a tuning of parameter  $c$  for each fixed  $n = n_0$  illustrated for Bernoulli,

Gaussian, exponential and Chi-square distributions. Specifically,  $c$  is chosen to minimize the width of the confidence set at  $n = n_0$ , i.e.,  $c_{n_0} = \min_{c>0} |\Theta_{n_0,c}(\delta)|$ . Let us recall that such a tuning is only valid for a single value. Hence, while using  $c = c_{n_0}$  for a fixed  $n_0$  is allowed by the theory to produce time-uniform bounds for all  $n$ , using  $c = c_n$ , that is a different value for different  $n$ , is not supported by the theory as it would break the martingale property necessary for the mixture construction to hold, which is manifested by a contradiction with the law of iterated logarithm (see Section 3.1).

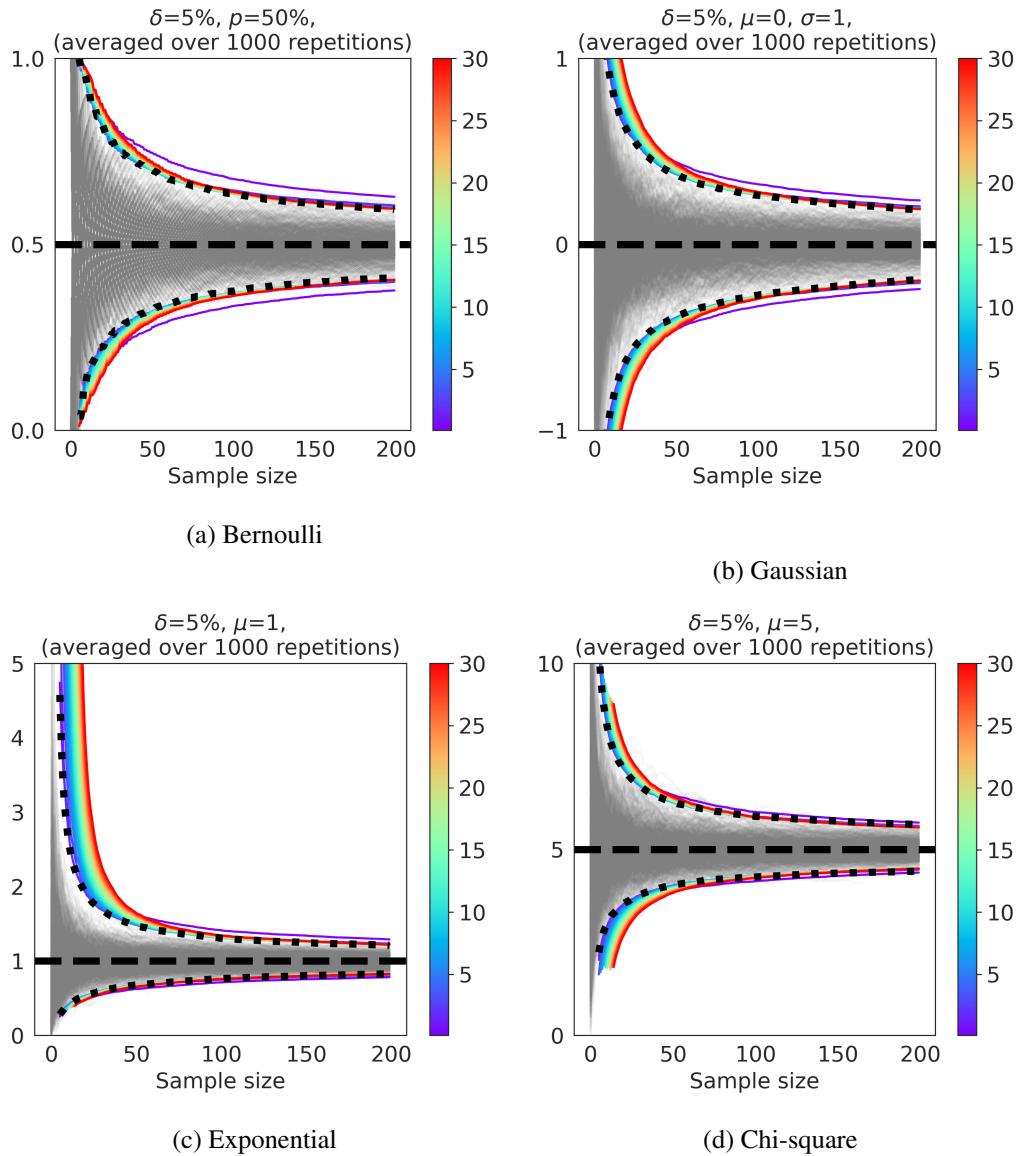


Figure 12: Confidence envelopes for varying  $c \in [0.1, 30]$  for  $\mathcal{B}(0.5)$  (top left),  $\mathcal{N}(0, 1)$  (top right),  $\mathcal{E}(1)$  (bottom left) and  $\chi^2(5)$  (bottom right). The dotted black line corresponds to the heuristics  $c_n \approx 0.12n$ .

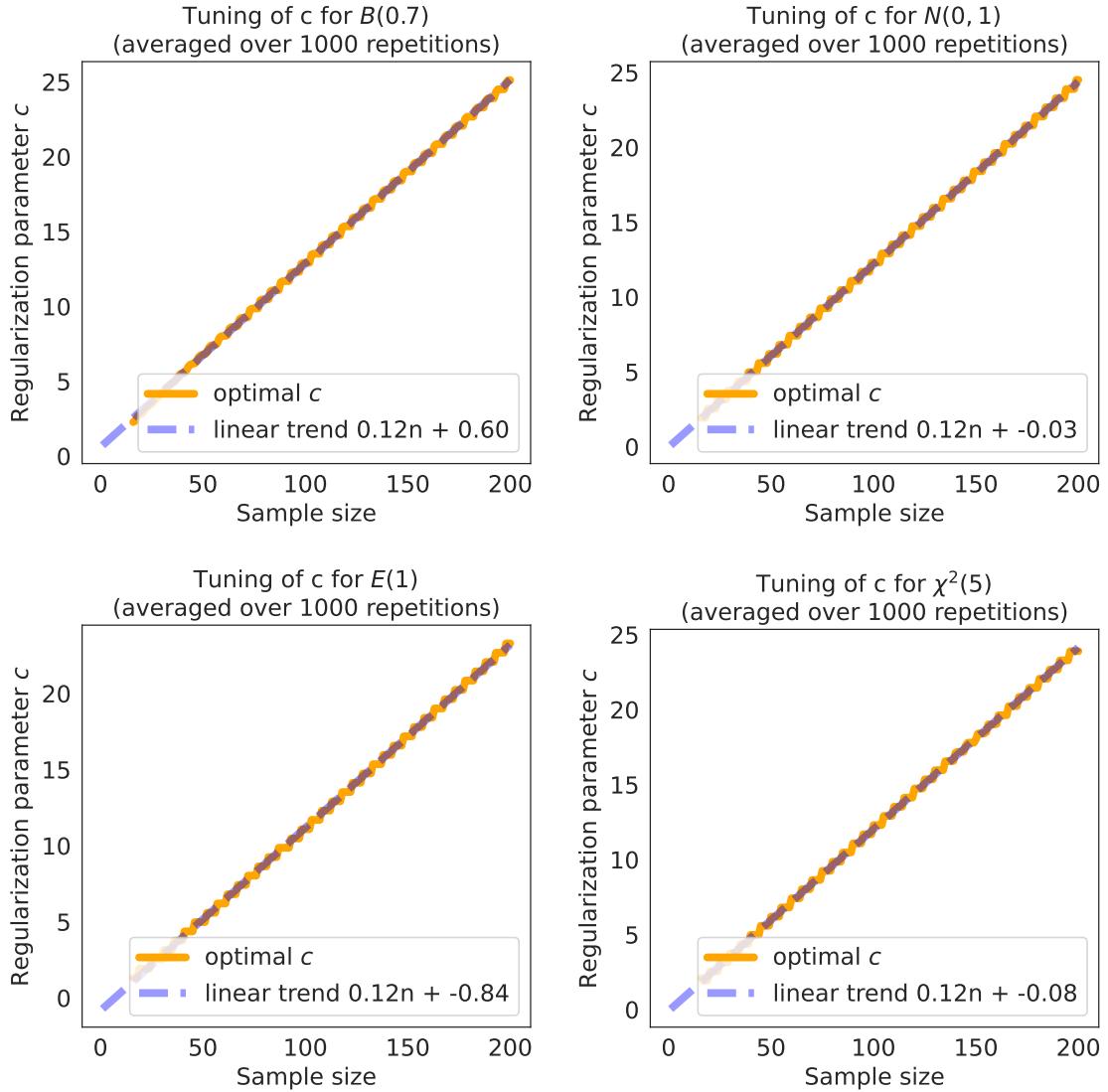


Figure 13: Optimal (smallest confidence interval diameter) local regularization parameter  $c_n^* = \operatorname{argmin}_c |\Theta_{n,c}(\delta)|$  for  $\delta = 5\%$  on Bernoulli (top left), Gaussian (top right), exponential (bottom left) and Chi-square (bottom right). Note the apparent common linear trend  $c_n^* \approx 0.12n$ . Results are averaged over 1000 independent simulations.

## Appendix E. Application: Generalized Likelihood Ratio Test in Exponential Families

In this section, we apply our result to revisit GLR tests in exponential families. The Generalized Likelihood Ratio (GLR) in the exponential family model  $\mathcal{E}$  writes as follows

$$\begin{aligned} G_{1:s:t}^{\mathcal{E}} &= \log \left( \frac{\sup_{\theta_1, \theta_2} \prod_{t'=1}^s p_{\theta_1}(X_{t'}) \prod_{t'=s+1}^t p_{\theta_2}(X_{t'})}{\sup_{\theta} \prod_{t'=1}^t p_{\theta}(X_{t'})} \right) \\ &= \sup_{\theta_1, \theta_2} \sum_{t'=1}^s \log p_{\theta_1}(X_{t'}) + \sum_{t'=s+1}^t \log p_{\theta_2}(X_{t'}) - \sup_{\theta} \sum_{t'=1}^t \log p_{\theta}(X_{t'}) \\ &= s\mathcal{L}^*(F_{1:s}) + (t-s)\mathcal{L}^*(F_{s+1:t}) - t\mathcal{L}^*(F_{1:t}) \\ &= s(\mathcal{L}^*(F_{1:s}) - \mathcal{L}^*(F_{1:t})) + (t-s)(\mathcal{L}^*(F_{s+1:t}) - \mathcal{L}^*(F_{1:t})), \end{aligned}$$

where  $F_{s+1:t} = \frac{1}{t-s} \sum_{t'=s+1}^t F(X_{t'})$ , and it makes sense to introduce  $\theta_{s+1:t}$  such that  $\nabla \mathcal{L}(\theta_{s+1:t}) = F_{s+1:t}$ . The GLR test is then defined, for a threshold  $\alpha$  as

$$\tau(\alpha; \mathcal{E}) = \min\{t \in \mathbb{N} : \max_{s \in [0,t]} G_{1:s:t}^{\mathcal{E}} \geq \alpha\}$$

An alternative formulation of the GLR shows that

$$\begin{aligned} G_{1:s:t}^{\mathcal{E}} &= \inf_{\theta} \sup_{\theta_1, \theta_2} \log \left( \frac{\prod_{t'=1}^s p_{\theta_1}(X_{t'}) \prod_{t'=s+1}^t p_{\theta_2}(X_{t'})}{\prod_{t'=1}^t p_{\theta}(X_{t'})} \right) \\ &= \inf_{\theta} \sup_{\theta_1, \theta_2} s \left( \langle \theta_1, F_{1:s} \rangle - \mathcal{L}(\theta_1) - \langle \theta, F_{1:t} \rangle + \mathcal{L}(\theta) \right) \\ &\quad + (t-s) \left( \langle \theta_2, F_{s+1:t} \rangle - \mathcal{L}(\theta_2) - \langle \theta, F_{1:t} \rangle + \mathcal{L}(\theta) \right). \end{aligned}$$

The first supremum is obtained for  $\theta_1 = [\nabla \mathcal{L}]^{-1}(F_{1:s})$ , with value equal to  $\mathcal{B}_{\mathcal{L}}(\theta, \theta_1) + \langle \theta - \theta_1, \nabla \mathcal{L}(\theta_1) \rangle + \langle \theta_1, F_{1:s} \rangle + \langle \theta_1, F_{1:t} \rangle$ , that is  $\mathcal{B}_{\mathcal{L}}(\theta, \theta_{1:s}) + \langle \theta, F_{1:s} - F_{1:t} \rangle$ . Likewise, the second supremum is obtained for  $\theta_2 = [\nabla \mathcal{L}]^{-1}(F_{s+1:t})$ , with value equal to  $\mathcal{B}_{\mathcal{L}}(\theta, \theta_{s+1:t}) + \langle \theta, F_{s+1:t} - F_{1:t} \rangle$ . Combining these two results, and remarking that  $s(F_{1:s} - F_{1:t}) + (t-s)(F_{s+1:t} - F_{1:t}) = 0$ , we obtain that

$$G_{1:s:t}^{\mathcal{E}} = \inf_{\theta} s\mathcal{B}_{\mathcal{L}}(\theta, \theta_{1:s}) + (t-s)\mathcal{B}_{\mathcal{L}}(\theta, \theta_{s+1:t}) \quad (15)$$

We further note that a solution to this optimization problem is obtained for  $\theta$  such that  $s(\nabla L(\theta) - \nabla L(\theta_{1:s})) + (t-s)(\nabla L(\theta) - \nabla L(\theta_{s+1:t})) = 0$ . Reorganizing the terms, this entails that  $\nabla L(\theta) = \frac{1}{t}(sF_{1:s} + (t-s)F_{s+1:t}) = F_{1:t}$ , and thus (without surprise)  $\theta = \theta_{1:t}$ .

Written in the form (15), the GLR satisfies for any  $\theta'$ ,

$$G_{1:s:t}^{\mathcal{E}} \leq s\mathcal{B}_{\mathcal{L}}(\theta', \theta_{1:s}) + (t-s)\mathcal{B}_{\mathcal{L}}(\theta', \theta_{s+1:t}).$$

In particular, the false alarm rate of the GLR test, when all observations are generated from parameter  $\theta$ , can be bounded using

$$\begin{aligned} \mathbb{P}_{\theta}(\tau(\alpha; \mathcal{E}) < \infty) &= \mathbb{P}_{\theta}(\exists(t, s) \in \mathbb{N}^2, s < t : G_{1:s:t}^{\mathcal{E}} \geq \alpha) \\ &\leq \mathbb{P}_{\theta}\left(\exists(t, s) \in \mathbb{N}^2, s < t : s\mathcal{B}_{\mathcal{L}}(\theta, \theta_{1:s}) + (t-s)\mathcal{B}_{\mathcal{L}}(\theta, \theta_{s+1:t}) \geq \alpha\right). \end{aligned}$$

At this point, provided that for appropriate terms  $\alpha_1, \alpha_2$ ,

$$\begin{aligned}\mathbb{P}_\theta\left(\exists s \in \mathbb{N} : s\mathcal{B}_{\mathcal{L}}(\theta, \theta_{1:s}) \geq \alpha_2\right) &\leq \frac{\delta}{2}, \\ \mathbb{P}_\theta\left(\exists(t, s) \in \mathbb{N}^2, s < t : (t-s)\mathcal{B}_{\mathcal{L}}(\theta, \theta_{s+1:t}) \geq \alpha_1\right) &\leq \frac{\delta}{2}.\end{aligned}$$

we deduce by a simple union bound argument, and for  $\alpha = \alpha_1 + \alpha_2$  that  $\mathbb{P}_\theta(\tau(\alpha : \mathcal{E}) < \infty) \leq \delta$ .

**A regularized GLR test** Our concentration result shows that such a controlled can be obtained for a penalized version of the maximum likelihood parameter estimates  $\theta_{1:s}, \theta_{s+1:t}, \theta_{1:t}$ . Namely, let us consider the set

$$\begin{aligned}\Theta_{s,c}(\delta) &= \left\{ \theta \in \Theta : \exists s \in \mathbb{N} : s\mathcal{B}_{\mathcal{L}}(\theta, \theta_{1:s,c}(\theta)) \geq \alpha_1 \right\}, \\ \Theta_{s+1:t,c}(\delta) &= \left\{ \theta \in \Theta : \exists(t, s) \in \mathbb{N}^2, s < t : (t-s)\mathcal{B}_{\mathcal{L}}(\theta, \theta_{s+1:t,c}(\theta)) \geq \alpha_2 \right\}.\end{aligned}$$

Provided that  $\alpha_1$  and  $\alpha_2$  are defined to ensure that these are high-probability confidence sets for the parameter  $\theta_0$ , then they should have non-empty intersections under  $\mathbb{P}_\theta$ . This motivates the following definition of the regularized GLR-test:

$$\tau_c^\delta(\mathcal{E}) = \min \{t \in \mathbb{N} : \exists s < t, \Theta_{s,c}(\delta/2) \cap \Theta_{s+1:t,c}(\delta/2) = \emptyset\}.$$

The appropriate quantity for  $\alpha_1$  is immediately obtained by Theorem 3 as  $\frac{s}{s+c}(\log(1/\delta) + \gamma_{1:s,c}(\theta))$ . Now for  $\alpha_2$ , we need to study doubly-time uniform concentration inequalities over both  $s$  and  $t$ .

### E.1. Doubly-time uniform concentration inequalities for exponential families

**Theorem 5 (Doubly time-uniform concentration)** Let  $\delta \in (0, 1]$  and  $g: \mathbb{N} \rightarrow \mathbb{R}_+^*$  such that  $\sum_{t=1}^\infty 1/g(t) \leq 1$  (e.g.,  $g(t) = \kappa(1+t) \log^2(1+t)$  where  $\kappa = 2.10975$ ). Then, under the assumptions of Definition 2, it holds that

$$\mathbb{P}_\theta\left[\exists t \in \mathbb{N}, \exists s < t : (t-s+c)\mathcal{B}_{\mathcal{L}}(\theta, \theta_{s+1:t,c}(\theta)) \geq \log \frac{g(t)}{\delta} + \gamma_{s+1:t,c}(\theta)\right] \leq \delta.$$

#### Proof of Theorem 5:

We consider that all the  $(X_{t'})_{t' \in \mathbb{N}}$  come from a distribution with same parameter  $\theta$ . Let  $t \in \mathbb{N}$  and  $s \in \{0, \dots, t-1\}$ . We introduce the scan-mean  $\mu_{s+1:t} = \frac{1}{t-s} \sum_{t'=s+1}^t F(X_{t'})$  and its mean  $\mu = \nabla \mathcal{L}(\theta)$ . We define as in the proof of Theorem 3, for each  $\lambda$  the martingale

$$M_{s+1:t}^\lambda = \exp\left(\langle \lambda, (t-s)(\mu_{s+1:t} - \mu) \rangle - (t-s)\mathcal{B}_{\mathcal{L},\theta}(\lambda)\right).$$

Then, applying the method of mixture and replacing each term with its scan version from  $s+1$  to  $t$ , we introduce the quantity

$$M_{s+1,t} = \exp\left((t-s+c)\mathcal{B}_{\mathcal{L},\theta}^\star(x)\right) \frac{G(\theta, c)}{G(\theta_{s+1:t,c}(\theta_0), t-s+c)}$$

where  $x = \frac{t-s}{t-s+c}(\mu_{s+1:t} - \nabla \mathcal{L}(\theta))$  and  $\theta_{s+1:t,c}(\theta)$  is the regularized estimate from the scan samples

$$\theta_{s+1:t,c}(\theta) = (\nabla \mathcal{L})^{-1} \left( \frac{\sum_{t'=s+1}^t F(X_{t'}) + c \nabla \mathcal{L}(\theta)}{t-s+c} \right).$$

Remarking that  $\mathbb{E}[M_{t,t}] \leq \mathbb{E}[M_{t+1,t}] = 1$  and that  $(M_{s+1,t})_s$  is a nonnegative supermartingale, we can now control its doubly-time uniform deviations following the proof of [Maillard \(2019b\)](#)[Th.3.2, p.58]. More precisely, for any non-decreasing function  $g$ , it can be shown that

$$\mathbb{P}_\theta \left[ \exists t, \exists s < t : M_{s+1,t} \geq \frac{g(t)}{\delta} \right] \leq \delta \mathbb{E}_\theta \left[ \max_t \max_{s < t} \frac{M_{s+1:t}}{g(t)} \right] \leq \delta \sum_{t=1}^{\infty} \frac{1}{g(t)}.$$

Rewriting the terms thanks to the duality formulas, yields

$$\mathbb{P}_\theta \left[ \exists t, \exists s < t : (t-s-c)\mathcal{B}_\mathcal{L}(\theta, \theta_{s+1:t,c}(\theta)) \geq \log \left( \frac{g(t)}{\delta} \right) + \gamma_{s+1:t,c}(\theta) \right] \leq \delta \sum_{t=1}^{\infty} \frac{1}{g(t)}.$$

This suggests to set  $\alpha_2 = \frac{t-s}{t-s-c} \left( \log \left( \frac{g(t)}{\delta} \right) + \gamma_{s+1:t,c}(\theta) \right)$  in the definition of the regularized GLR test. Putting all terms together, this leads to the definition of the following regularized GLR test, for a given false-detection error probability  $\delta \in [0, 1]$  and regularization parameter  $c$ ,

$$\begin{aligned} \Theta_{s,c}(\delta) &= \{\theta \in \Theta : (s+c)\mathcal{B}_\mathcal{L}(\theta, \theta_{1:s,c}(\theta)) \leq \log(1/\delta) + \gamma_{1:s,c}(\theta)\}, \\ \Theta_{s+1:t,c}(\delta) &= \{\theta \in \Theta : (t-s+c)\mathcal{B}_\mathcal{L}(\theta, \theta_{s+1:t,c}(\theta)) \leq \log(g(t)/\delta) + \gamma_{s+1:t,c}(\theta)\}, \\ \tau_c^\delta(\mathcal{E}) &= \min \{t \in \mathbb{N} : \exists s < t, \Theta_{s,c}(\delta/2) \cap \Theta_{s+1:t,c}(\delta/2) = \emptyset\}. \end{aligned}$$

The factor  $g(t)$  that inflates the width of the confidence set can be tuned to satisfy  $\sum_{t=1}^{+\infty} 1/g(t) \leq 1$ , thus controlling the above deviation probability by at most  $\delta$  (ideally,  $\sum_{t=1}^{+\infty} 1/g(t)$  should be close to 1 to avoid overinflating the confidence width). A natural choice for this is  $g(t) = \kappa(1+t) \log^2(1+t)$  where  $\kappa = 2.10974 \geq \sum_{t=1}^{\infty} \frac{1}{(1+t) \log^2(1+t)}$  (for completeness, we derive in the next section an elementary way to computing suitable values for  $\kappa$ ). By construction, this test is guaranteed to have a false-detection probability controlled by  $\delta$  (that is,  $\mathbb{P}_\theta(\tau_c^\delta(\mathcal{E}) < \infty) \leq \delta$ ).

**Doubly time-uniform control of supermartingale sequences** For completeness, we reproduce below the derivation from [Maillard \(2019b\)](#)[Th.3.2, p.58], applied to our setup. First, let us note that

$$\begin{aligned} \mathbb{P} \left( \exists t \in \mathbb{N}_*, \exists s \in [0, t-1], M_{s+1:t} \geq \frac{g(t)}{\delta} \right) &= \mathbb{P} \left( \max_{t \in \mathbb{N}_*} \max_{s \in [0, t-1]} \frac{M_{s+1:t}}{g(t)} \geq \frac{1}{\delta} \right) \\ &\leq \delta \mathbb{E} \left[ \max_{t \in \mathbb{N}_*} \max_{s \in [0, t-1]} \frac{M_{s+1:t}}{g(t)} \right]. \end{aligned}$$

Let us also denote  $\tau$  the random stopping time corresponding to the first occurrence  $t$  of the event  $\max_{s \in [0, t-1]} \frac{M_{s+1:t}}{g(t)} \geq \frac{1}{\delta}$ . It is convenient to introduce the quantity  $\bar{M}_t = \frac{\sum_{s \in \{0, \dots, t-1\}} M_{s+1,t}}{g(t)}$  for each  $t \in \mathbb{N}_*$ . Since each  $M_{s+1,t}$  and  $g(t)$  is nonnegative, we first get that for every random stopping time  $\tau \in \mathbb{N}_*$ , the following inequality holds:

$$\mathbb{E} \left[ \frac{\max_{s < \tau} M_{s+1,\tau}}{g(\tau)} \right] \leq \mathbb{E} \left[ \bar{M}_\tau \right] = \mathbb{E} \left[ \bar{M}_1 + \sum_{t=1}^{\infty} (\bar{M}_{t+1} - \bar{M}_t) \mathbb{I}\{\tau > t\} \right].$$

Furthermore, we note that, conveniently

$$\overline{M}_{t+1} - \overline{M}_t = \frac{M_{t+1,t+1}}{g(t+1)} + \sum_{s=0}^{t-1} \left( \frac{M_{s+1,t+1}}{g(t+1)} - \frac{M_{s+1,t}}{g(t)} \right).$$

Next, by assumption, we note that  $\mathbb{E}[M_{s+1,t+1} | \mathcal{F}_t] \leq M_{s+1,t}$ . Thus, since  $\mathbb{I}\{\tau > t\} \in \mathcal{F}_t$ , we deduce that

$$\begin{aligned} & \mathbb{E} \left[ \frac{\max_{s<\tau} M_{s+1,\tau}}{g(\tau)} \right] \\ & \leq \mathbb{E}[\overline{M}_1] + \sum_{t=1}^{\infty} \frac{\mathbb{E}[M_{t+1,t+1}]}{g(t+1)} + \sum_{t=1}^{\infty} \sum_{s< t} \mathbb{E} \left[ \left( \frac{1}{g(t+1)} - \frac{1}{g(t)} \right) M_{s+1,t} \mathbb{I}\{\tau > t\} \right] \\ & = \mathbb{E}[\overline{M}_1] + \sum_{t=1}^{\infty} \frac{\mathbb{E}[M_{t+1,t+1}]}{g(t+1)} + \sum_{t=1}^{\infty} \sum_{s< t} \left( \frac{1}{g(t+1)} - \frac{1}{g(t)} \right) \underbrace{\mathbb{E}[M_{s+1,t} \mathbb{I}\{\tau > t\}]}_{\geq 0}. \end{aligned}$$

Hence, the assumption that  $g$  is non-decreasing ensures that the last sum is upper bounded by 0. Since on the other hand  $\mathbb{E}[M_{t+1,t+1}] \leq 1$  holds for all  $t$  (and thus  $\mathbb{E}[\overline{M}_1] \leq 1/g(1)$ ), we deduce that the following inequality holds:

$$\mathbb{E} \left[ \frac{\max_{s<\tau} M_{s+1,\tau}}{g(\tau)} \right] \leq \frac{1}{g(1)} + \sum_{t=1}^{\infty} \frac{1}{g(t+1)} = \sum_{t=1}^{\infty} \frac{1}{g(t)} \leq 1.$$

□

## E.2. Computing the inflating factor $g(t)$

The right-hand side of the doubly time-uniform confidence set of Theorem 5 involves a  $\log g(t)/\delta$  term instead of  $\log 1/\delta$  as in Theorem 3, which is a by-product of the union-like argument used in the proof. Ideally,  $g(t)$  should grow as slow as possible with  $t$  to avoid unnecessary looseness in the confidence bound. That is however limited by the constraint  $\sum_{t=1}^{\infty} 1/g(t) \leq 1$ , which prohibits the use of a linearly growing  $g(t)$  (since  $\sum_{t=1}^{\infty} 1/t$  diverges). The choice  $g(t) = \kappa(1+t) \log^{1+\eta}(1+t)$  for some  $\eta > 0$  and  $\kappa > 0$  guarantees that the series of inverses converges, although its limit is not available in closed-form. However, it is sufficient to set  $\kappa$  to an upper bound on  $\sum_{t=1}^{\infty} \frac{1}{(1+t) \log^{1+\eta}(1+t)}$ . The following elementary lemma explains how to compute a tight value for  $\kappa$ .

**Lemma 12** *Let  $f: t \in \mathbb{N} \mapsto (1+t) \log^{1+\eta}(1+t)$  and define  $S_p = \sum_{t=1}^p 1/f(t)$  for  $p \in \mathbb{N} \cup \{+\infty\}$ . Then  $S_{\infty} \leq S_p + \frac{1}{\eta \log(1+p)^{\eta}}$ .*

### Proof of Lemma 12:

The function  $f$  is nondecreasing and positive, therefore for  $t \in \mathbb{N}$  and  $x \in [t, t+1]$ , it holds that  $1/f(t+1) \leq 1/f(x)$ . Integrating both terms as functions of  $x$  over the interval  $[t, t+1]$  yields

$1/f(t+1) \leq \int_t^{t+1} 1/f(x)dx$ . Summing starting at  $t = p$ , we further get the following sum-integral comparison:

$$S_\infty - S_p \leq \int_p^\infty \frac{dx}{f(x)}.$$

Finally, note that  $\frac{d}{dx} \frac{1}{\eta \log^\eta(1+x)} = -\frac{1}{f(x)}$ , and thus  $\int_p^\infty \frac{dx}{f(x)} = \frac{1}{\eta \log(1+p)^\eta}$ .  $\square$

---

This lemma shows that  $\kappa = S_p + 1/\log(1+p)$  for some  $p \geq 1$  ( $\eta = 1$ ) is a valid choice for the definition of the inflating factor  $g(t)$ , and is straightforward to compute numerically. For  $p = 100$ , we obtain  $\kappa \lesssim 2.10974$  and increasing  $p$  only changes further digits.

### E.3. Experiments on change point detection for Gaussian with unknown variance

We illustrate the change point GLR test above in a numerical experiment as follows. We consider the one-dimensional exponential family of centered Gaussian distributions with unknown variance  $\mathcal{E} = \{\mathcal{N}(0, \sigma^2), \sigma > 0\}$ . A sequence of independent random variables  $(X_t)_{t \in \mathbb{N}}$  is drawn from  $X_t \sim \mathcal{N}(0, \sigma_0^2)$  if  $t \leq t^*$  and from  $X_t \sim \mathcal{N}(0, \sigma_1^2)$  if  $t > t^*$ , where  $t^* = 50$  and  $\sigma_0 = 1$  and  $\sigma_1 \in [1, 2, 3, 4]$ . Of note, this setting corresponds to an open question in [Maillard \(2019a\)](#), which studies GLR tests in sub-Gaussian families, which are adapted to the detection of changing means but not variances. Detection times are reported as  $\min(\tau_c^\delta(\mathcal{E}), T)$  with  $T = 100$ , and  $\tau_c^\delta(\mathcal{E}) > T$  is interpreted as no change being detected. For the doubly time-uniform confidence set in the definition of the regularized GLR test, we use the factor  $g(t) = \kappa(1+t) \log^2(1+t)$  with  $\kappa = 2.10974$ .

A practical motivation for this setting is, for instance, the design of maintenance models for equipment: with time, a component of a physical system (mechanical component, measuring instrument...) starts to wear off and while it is still functioning (same output on average), it is more imprecise or less stable (higher output variance); the goal of the maintenance agent is to detect as early as possible such changes of regime in order to replace the failing component before it breaks completely.

In Figure 14, we report the histograms of detection times across 1000 independent simulations for increasingly abrupt changes of variance ( $\sigma_1 \in \{1, 2, 3, 4\}$ ). As expected, the distribution of detection times shifts closer to the actual change point  $t^*$  when  $\sigma_1$  increases, i.e., more obvious changes are detected earlier. This is confirmed in Figure 15, where we report the median and interquartile range of detection times for  $\sigma_1 \in [1, 4]$ .

We empirically validate in Figure 14a the control of the false positive probability by at most  $\delta = 5\%$ . However, the actual false positive rate appears much lower (0.2%), which is a consequence of the looseness of the union bound involving the factor  $g(t)$  in the doubly time-uniform confidence set. Following [Maillard \(2019a\)](#), a sharper approach would involve a direct concentration result on the pair  $(\theta_{s,c}, \theta_{s+1:t,c})$  (i.e., in the terminology of [Maillard \(2019a\)](#), a joint bound, as opposed to the current disjoint one). This is a nontrivial result which we leave for future work. Finally, note that the high false negative rate of Figure 14b (27.6%) is an artifact of thresholding the detection time at  $T = 100$ ; increasing  $T$  would enable later detections, thus reducing the number of observed false negatives, at the cost of increasing the average detection delay (as well as the computational burden of the experiment).

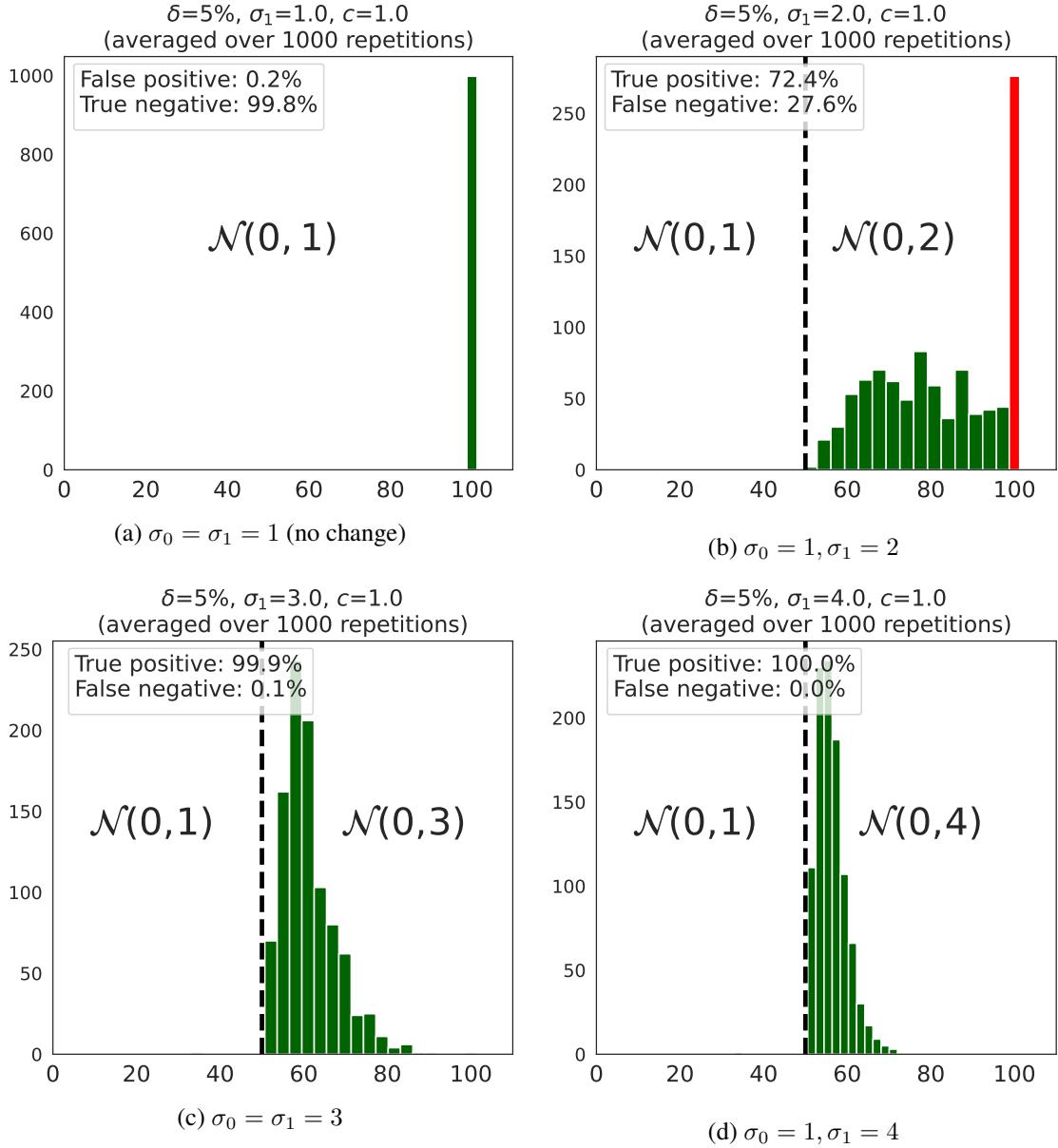


Figure 14: Histograms of detection times  $\tau_c^\delta(\mathcal{E})$  in the Gaussian setting described in Appendix E.3. Green: correct detection (negative in (a), positive in (b), (c) and (d)). Red: incorrect detection (positive in (a), negative in (b), (c) and (d)). Black dashed line: change point  $t^*$ .

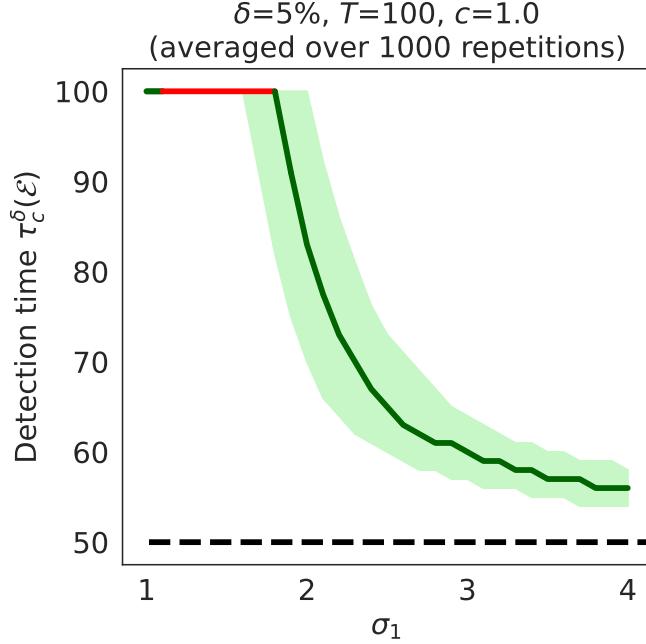


Figure 15: Detection time  $\tau_c^\delta(\mathcal{E})$  in the Gaussian setting described in Appendix E.3. Solid line: median detection time. Shaded area: interquartile range of detection times (25th and 75th percentiles). Black dashed line: change point  $t^*$ .

## Appendix F. Application: Linear Contextual Bandits

In this section, we apply Theorem 7 to build confidence sets in the well-known linear bandit setting. We consider the setting of linear contextual bandits (Auer et al., 2002; Abbasi-Yadkori et al., 2011), but with possibly arm-dependent noise variance (heteroscedastic noise). An algorithm for this problem chooses, at each round  $t \in \mathbb{N}$ , an action or arm  $A_t \in \mathcal{A}$ , and subsequently observes a reward  $X_t = \theta^\top \varphi(A_t) + \varepsilon_t$ , where  $\varphi: \mathcal{A} \rightarrow \mathbb{R}^d$  is a (fixed) map from actions to their context vectors,  $\theta \in \mathbb{R}^d$  is a vector of weights unknown to the algorithm and given the action  $A_t$ ,  $\varepsilon_t \sim \mathcal{N}(0, \sigma_{A_t}^2 I)$  is a Gaussian noise with known variance  $\sigma_{A_t}^2$ . The goal is to select actions to maximize the expected cumulative reward  $\sum_t \theta^\top \varphi(A_t)$ . Any rational agent would choose the action  $A_t$  causally depending upon the history of arms and reward sequences available before the start of round  $t$ . Naturally, for each  $n \in \mathbb{N}$ , it boils down to controlling the deviation between the unknown parameter  $\theta$  and a suitable estimate  $\theta_n$  built from  $n$  observations  $(A_t, X_t)_{t \leq n}$ .

Under this observation model, the feature function and the log-partition function are given by  $F_t(x) = \frac{x}{\sigma_{A_t}^2} \varphi(A_t)$  and  $\mathcal{L}_t(\theta) = \frac{1}{2\sigma_{A_t}^2} \|\theta\|_{\varphi(A_t)\varphi(A_t)^\top}^2$ . Let us define the matrix

$$V_n = \sum_{t=1}^n \frac{1}{\sigma_{A_t}^2} \varphi(A_t) \varphi(A_t)^\top,$$

and introduce the Legendre function  $\mathcal{L}_0(\theta) = \frac{1}{2}\|\theta\|_{V_0}^2$ , where  $V_0$  is some fixed positive definite matrix. Then, the parameter estimate and the Bregman information gain take the form

$$\theta_{n,\mathcal{L}_0} = (V_n + V_0)^{-1} \sum_{t=1}^n \frac{X_t}{\sigma_{A_t}^2} \varphi(A_t), \quad \gamma_{n,\mathcal{L}_0} = \log \frac{\det(V_0 + V_n)^{1/2}}{\det(V_0)^{1/2}}.$$

Now, defining  $\mathcal{H}_{t-1} := \{A_1, X_1, \dots, A_{t-1}, X_{t-1}, A_t\}$  to be the set of all information available before observing  $X_t$ , one can see that the confidence set  $\Theta_{n,\mathcal{L}_0}(\delta)$  given by Theorem 7 is the set of all  $\theta \in \Theta$  satisfying

$$\|\theta - \theta_{n,\mathcal{L}_0}\|_{V_n + V_0}^2 \leq \|\theta\|_{V_0}^2 + 2 \log \frac{\det(V_0 + V_n)^{1/2}}{\delta \det(V_0)^{1/2}}.$$

**Comparison with Existing Results** Under the assumption that the noise  $\varepsilon_t$  is  $\sigma_{A_t}$ -sub-Gaussian conditioned on  $\mathcal{H}_{t-1}$ , Abbasi-Yadkori et al. (2011); Kirschner and Krause (2018) obtain a high-probability confidence set  $\tilde{\Theta}_{n,\mathcal{L}_0}(\delta)$ , which is the set of all  $\theta \in \Theta$  satisfying

$$\|\theta - \theta_{n,\mathcal{L}_0}\|_{V_n + V_0} \leq \|\theta\|_{V_0} + \sqrt{2 \log \frac{\det(V_0 + V_n)^{1/2}}{\delta \det(V_0)^{1/2}}}.$$

Interestingly, since  $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$  for positive  $a, b$ , it is then clear that  $\Theta_{n,\mathcal{L}_0}(\delta) \subset \tilde{\Theta}_{n,\mathcal{L}_0}(\delta)$ . In particular, as a side result of our approach, we obtain a tighter confidence set compared to them.