

A NEW BOUND ON THE CUMULANT GENERATING FUNCTION OF DIRICHLET PROCESSES

BY PIERRE PERRAULT^{1,a}, DENIS BELOMESTNY^{2,b}, PIERRE MÉNARD^{3,c},
ÉRIC MOULINES^{4,e}, ALEXEY NAUMOV^{5,g}, DANIIL TIAPKIN^{4,f}, MICHAL VALKO^{3,d}

¹IDEMIA, ^apierre.perrault@outlook.com

²Duisburg-Essen University, Germany, ^bdenis.belomestny@uni-due.de

³Meta, ^cpmenard@meta.com; ^dmir@meta.com

⁴Centre de Mathématiques Appliquées, CNRS, École Polytechnique, Institut Polytechnique de Paris, France,
^eeric.moulines@polytechnique.edu; ^fdaniil.tiapkin@polytechnique.edu

⁵HSE University, Russia, ^ganaumov@hse.ru

In this paper, we introduce a novel approach for bounding the cumulant generating function (CGF) of a Dirichlet process (DP) $X \sim \text{DP}(\alpha\nu_0)$, using superadditivity. In particular, our key technical contribution is the demonstration of the superadditivity of $\alpha \mapsto \log \mathbb{E}_{X \sim \text{DP}(\alpha\nu_0)}[\exp(\mathbb{E}_X[\alpha f])]$, where $\mathbb{E}_X[f] = \int f dX$. This result, combined with Fekete's lemma and Varadhan's integral lemma, converts the known asymptotic large deviation principle into a practical upper bound on the CGF $\log \mathbb{E}_{X \sim \text{DP}(\alpha\nu_0)}[\exp(\mathbb{E}_X[f])]$ for any $\alpha > 0$. The bound is given by the convex conjugate of the scaled reversed Kullback-Leibler divergence $\alpha \text{KL}(\nu_0 \|\cdot)$. This new bound provides particularly effective confidence regions for sums of independent DPs, making it applicable across various fields.

1. Introduction. The Dirichlet Process (DP) is a fundamental stochastic process in which each realization is itself a probability distribution. Originally introduced by Ferguson in the early 1970s [27], the DP has become a fundamental tool in the field of nonparametric Bayesian statistics [15, 32, 43]. The use of DPs in statistical modeling provides a significant benefit: it enables the model's complexity to adjust according to the data rather than requiring a fixed structure a priori. Conceptually, the DP can be considered as a generalization of the Dirichlet distribution into infinite-dimensional spaces. Just as the Dirichlet distribution acts as the conjugate prior for categorical distributions, the DP serves as the conjugate prior for nonparametric, discrete distributions over infinite spaces. Before proceeding further, we introduce the necessary notation. We focus on a compact metric space Ω , equipped with its Borel σ -algebra $\mathcal{B}(\Omega)$. The set $\mathcal{M}(\Omega)$ (and respectively $\mathcal{M}_1(\Omega)$) denotes the space of finite non-negative (and probability) measures on Ω , and $\mathcal{B}(\mathcal{M}_1(\Omega))$ represents the Borel σ -algebra generated by the weak topology on $\mathcal{M}_1(\Omega)$. The set of continuous functions $f: \Omega \rightarrow \mathbb{R}$ is denoted by $\mathcal{C}(\Omega)$. For a given probability distribution p , \mathbb{E}_p and \mathbb{P}_p denote the expectation and probability, respectively, with respect to p . Alternatively, we use $\mathbb{E}_{\xi \sim p}$ and $\mathbb{P}_{\xi \sim p}$ to explicitly indicate that the random variable ξ follows the distribution p .

Consider a DP on Ω , characterised by a scale parameter $\alpha > 0$ and a base distribution $\nu_0 \in \mathcal{M}_1(\Omega)$ whose measure's law is denoted as $\text{DP}(\alpha\nu_0)$ and whose realization X is a random probability measure on the space Ω . The original definition of the DP, introduced by Ferguson [27], says that for any finite measurable partition $B_1 \sqcup \dots \sqcup B_k = \Omega$, $(X(B_1), \dots, X(B_k)) \sim$

MSC2020 subject classifications: Primary 60E15, 62E17, 62G15; secondary 60G57, 60F10, 39B62.

Keywords and phrases: Dirichlet Process, cumulant generating function, Varadhan's integral lemma, superadditivity, Fekete's lemma, confidence region.

$\text{Dir}(\alpha\nu_0(B_1), \dots, \alpha\nu_0(B_k))$, where $\text{Dir}(\alpha_1, \dots, \alpha_k)$ denotes the Dirichlet distribution with parameters $\alpha_1, \dots, \alpha_k$. An alternative definition, known as the "stick-breaking" construction, is $X = \sum_{k=1}^{\infty} \beta_k \prod_{\ell < k} (1 - \beta_\ell) \delta_{\omega_k}$, where $(\omega_k, \beta_k) \stackrel{iid}{\sim} \nu_0 \times \text{Beta}(1, \alpha)$ [47], where Beta denotes the beta distribution. Another important representation of the DP is analogous to the characterization of the Gamma distribution by [39] and expressed as $X = G/G(\Omega)$, where $G \sim \mathcal{G}(\alpha\nu_0)$ is the standard Gamma process¹ on Ω with shape parameter $\alpha\nu_0$, i.e., $G(A) \sim \text{Gamma}(\alpha\nu_0(A), 1)$ for any measurable set $A \subset \Omega$ [11].

Given the widespread use of Dirichlet Processes, understanding concentration phenomena [2, 33] and large deviation principles (LDPs, see Definition 2.2) [1, 16, 57] is crucial, particularly in the context of their applications to fields such as machine learning [4], reinforcement learning [45, 46], topic modeling [5, 55], among others. In this paper, we will focus on the study of concentration bounds for DPs. Indeed, the literature on LDPs for DPs is already well-established [21, 25, 26, 29, 40] and it is known that the probability that $X \sim \text{DP}(\alpha\nu_0)$ deviates from another distribution ν decreases exponentially (with respect to α) at a rate given by the Kullback–Leibler (KL) divergence (aka relative entropy) of ν_0 with respect to ν , defined as

$$\text{KL}(\nu_0 \parallel \nu) \triangleq \begin{cases} \mathbb{E}_{\nu_0} \left[\log \left(\frac{d\nu_0}{d\nu} \right) \right] & \text{if } \nu_0 \ll \nu \\ \infty & \text{otherwise.} \end{cases}$$

On the non-asymptotic side, research related to DPs remains sparse. The existing studies utilize the Gamma process-based representation defined above and afterward use a closed-form of the moment-generating function (MGF) formula for the Gamma process. This approach has been explored in various works; see [56, 59]. In particular, for any $f \in \mathcal{C}(\Omega)$, such that $f \leq 1$ and $\nu_0(f^{-1}(\{1\})) = 0$, the MGF of the Gamma process can be expressed as:

$$(1) \quad M_{\mathcal{G}(\alpha\nu_0)}(f) \triangleq \mathbb{E}_{G \sim \mathcal{G}(\alpha\nu_0)} \left[\exp \left(\int f dG \right) \right] = \exp \left(-\alpha \mathbb{E}_{\nu_0} [\log(1 - f)] \right).$$

While the Gamma process representation provides powerful tools for analyzing the concentration of an individual DP, extending this technique to multiple independent DPs is not feasible.

Our primary objective in this paper is to establish new concentration bounds for multiple independent DPs. Moving forward, we intend to focus specifically on the MGF of the DP, as opposed to that of the Gamma process. Indeed, the MGF and its logarithm, the cumulant-generating function (CGF), are essential in both asymptotic and non-asymptotic statistical analyses. Specifically, they play a critical role in establishing LDPs using the Gärtner-Ellis theorem (see [23, 31]). Within a non-asymptotic realm, the multiplicative nature of MGFs for sums of independent random variables can be utilized to derive concentration inequalities through Chernoff bounds [6, 8, 9, 50, 52]. In our context, bounding the CGF of a DP by a manageable expression would thus enable the derivation of new concentration inequalities for sums of independent DPs, which could have practical significance across various application domains.

The study of moment-generating function (MGF) bounds, particularly sub-Gaussian properties, has been extensively explored due to its broad applicability across various fields. This line of research focuses on identifying the optimal proxy variance for different types of random variables, including discrete distributions such as the Bernoulli distribution [7, 36], and continuous distributions like the Beta and Dirichlet distributions [42]. While sub-Gaussian bounds are widely used due to their general applicability, they often do not provide as tight

¹One of the formulations of the Gamma process is derived from a Poisson process over the space $\Omega \times [0, \infty)$ with mean measure $\mu(d\omega, dt) = \alpha\nu_0(d\omega)t^{-1}e^{-t}dt$. Drawing a sample from this Poisson process generates an infinite set of atoms $(\omega_k, t_k)_{k \geq 1}$. The Gamma process can then be constructed as $G \triangleq \sum_{k=1}^{\infty} t_k \delta_{\omega_k}$.

an estimate as those derived from the KL divergence (which are asymptotically optimal, as they match the LDP). Therefore, in this paper, we target a KL-based CGF bound on the DP.

The rest of the article is structured as follows. The following section reviews LDP results for DPs and their relevance to concentration bounds. Next is our main result (Theorem 3.3), which transforms a limit on a sequence of DP CGFs (focusing on the sequence considered in Varadhan’s integral lemma [57]) into a bound. Specifically, we utilize the superadditivity of the sequence (which is proved in the subsequent section), employing Fekete’s superadditive lemma [24]. The article concludes by applying this result to the stochastic semi-armed bandit problem [37].

2. Large deviation principle. We start by recalling the concept of a large deviation principle (LDP).

DEFINITION 2.1 (Rate function). A function I is a rate function if it is lower semi-continuous with values in $[0, \infty]$ (such that all level sets $\{x : I(x) \leq t\}$, for $t \in [0, \infty)$, are closed). A rate function is good if the level sets are compact. The effective domain of I is $D_I \triangleq \{x : I(x) < \infty\}$.

Notice, in the previous definition, we did not specify on which domain the rate function I is defined. In our context, I will be defined on $\mathcal{M}_1(\Omega)$.

DEFINITION 2.2 (Large deviation principle: LDP). A sequence of probability measures (μ_n) satisfies an LDP with speed n (we can avoid explicitly stating the speed if it is clear from the context) and rate function I if:

- (i) For all closed sets F , $\limsup_n \frac{1}{n} \log \mu_n(F) \leq -\inf_{x \in F} I(x)$,
- (ii) For all open sets G , $\liminf_n \frac{1}{n} \log \mu_n(G) \geq -\inf_{x \in G} I(x)$.

The standard, and likely the most well-known, LDP result for DP is as follows.

THEOREM 2.3 (see [29]). $(\text{DP}(\alpha\nu_0))_\alpha$ satisfies a LDP with speed α and rate function $I(\nu) = \text{KL}(\nu_0\|\nu)$, i.e., for all $B \in \mathcal{B}(\mathcal{M}_1(\Omega))$, if B° (resp. \bar{B}) denotes the interior of B (resp. the closure),

$$-\inf_{\nu \in B^\circ} I(\nu) \leq \liminf_{\alpha} \frac{1}{\alpha} \log \mathbb{P}_{\text{DP}(\alpha\nu_0)}[B] \leq \limsup_{\alpha} \frac{1}{\alpha} \log \mathbb{P}_{\text{DP}(\alpha\nu_0)}[B] \leq -\inf_{\nu \in \bar{B}} I(\nu).$$

The rate function in Theorem 2.3 is given by the reverse KL divergence $I(\cdot) = \text{KL}(\nu_0\|\cdot)$, which is a dual to the rate function in the Sanov theorem [53]. In [28], the reverse relation of the rate functions is explained as follows: “in Sanov’s theorem we ask how likely the empirical distribution is to be close to ν_0 , given that the true distribution is ν ; whereas in the Bayesian context we ask how likely it is that the true distribution is close to ν_0 , given that the empirical distribution is close to ν .” There are several ways to get Theorem 2.3. For instance, [40] utilize an LDP on the Gamma process, linking to DPs through the characterization mentioned above, while [29] rely on Varadhan’s integral lemma (see Fact 3).

For some probability distribution $\nu \in \mathcal{M}_1(\Omega)$, one of the central information-theoretic measures we are considering is defined as an infimum of Kullback-Leibler divergences: for some real-valued continuous function $f \in \mathcal{C}(\Omega)$ and some $u \in \mathbb{R}$, we define

$$\mathcal{K}_{\text{inf}}(\nu, u, f) \triangleq \inf_{\mu \in \mathcal{M}_1(\Omega), \mathbb{E}_\mu[f] \geq u} \text{KL}(\nu\|\mu),$$

where by convention, the infimum of the empty set equals to ∞ . Notice, \mathcal{K}_{inf} is used to express the bounds of Theorem 2.3 when B is some deviation event. This quantity can be interpreted as a distance from the measure ν to the set of all measures $\mu \in \mathcal{M}_1(\Omega)$, $\mathbb{E}_\mu[f] \geq u$, where the distance is measured by the KL-divergence. The measure μ solving this optimization problem is called moment projection (M -projection) or reversed information projection (rI -projection), see [3, 17, 44]. This is different from the more common information projection (I -projection), $\inf_{\mu \in \mathcal{M}_1(\Omega), \mathbb{E}_\mu[f] \geq u} \text{KL}(\mu \| \nu)$, appearing, for example, in Sanov-type deviation bounds [54]. The I -projections have a geometric interpretation because the KL can be viewed as a Bregman divergence. The M -projections are not Bregman divergences and lack geometric interpretation. However, they are deeply connected to the maximum likelihood estimation when the measure ν is the empirical measure of a sample [18, Lemma 3.1]. Additionally, as we saw in Theorem 2.3, M -projections naturally appear as a rate function for a LDP in a Bayesian framework [28]. They also naturally appear in lower (and sometimes upper) bounds for multi-armed bandits² [10, 38]. Like the KL divergence, $\mathcal{K}_{\text{inf}}(\nu, u, f)$ admits the following variational formula.

LEMMA 2.4 (Variational formula for \mathcal{K}_{inf} [30, 34]). *For all $\nu \in \mathcal{M}_1(\Omega)$, $f \in \mathcal{C}(\Omega)$, $u \in [f_{\min}, f_{\max})$, where $f_{\max} \triangleq \max_{x \in \Omega} f(x)$, $f_{\min} \triangleq \min_{x \in \Omega} f(x)$, we have*

$$\mathcal{K}_{\text{inf}}(\nu, u, f) = \max_{\lambda \in [0, 1/(f_{\max} - u)]} \mathbb{E}_\nu[\log(1 - \lambda(f - u))].$$

Moreover, if λ^* is the value at which the above maximum is reached, then

$$\mathbb{E}_\nu[1/(1 - \lambda^*(f - u))] \leq 1.$$

In particular, $\nu(f^{-1}(\{f_{\max}\})) = 0$ in the case $\lambda^* = 1/(f_{\max} - u)$.

This formula is essential for deriving the deviation and concentration results involving \mathcal{K}_{inf} : as a simple example, with the same assumptions and notations as in Lemma 2.4, one can consecutively use Chernoff bound, Equation 1 and Lemma 2.4 to get

$$\begin{aligned} \mathbb{P}_{X \sim \text{DP}(\alpha\nu_0)}[\mathbb{E}_X[f] \geq u] &\leq M_{\mathcal{G}(\alpha\nu_0)}(\lambda^*(f - u)) \\ (2) \quad &= e^{-\alpha \mathbb{E}_{\nu_0}[\log(1 - \lambda^*(f - u))]} = e^{-\alpha \mathcal{K}_{\text{inf}}(\nu_0, u, f)}. \end{aligned}$$

In fact, LDP rate functions, in general, are often used in non-asymptotic concentration inequalities. For example, suppose we have a process consisting of real-valued i.i.d. random variables (Y_i) . Then, it is known that the sequence $\log \mathbb{P}(\sum_{i=1}^m Y_i/m \geq x)$ is superadditive w.r.t. $m \in \mathbb{N}^*$ (we recall that a function $f : \mathbb{R} \rightarrow \mathbb{R}$ is called superadditive on $A \subset \mathbb{R}$ if $f(x + y) \geq f(x) + f(y)$ for all $x, y \in A$), so from the superadditive lemma due to Fekete [24], for all $n \geq 1$,

$$\begin{aligned} \frac{1}{n} \log \mathbb{P}\left(\frac{1}{n} \sum_{i=1}^n Y_i \geq x\right) &\leq \sup_{m \geq 1} \frac{1}{m} \log \mathbb{P}\left(\frac{1}{m} \sum_{i=1}^m Y_i \geq x\right) \\ &= \lim_{m \rightarrow \infty} \frac{1}{m} \log \mathbb{P}\left(\frac{1}{m} \sum_{i=1}^m Y_i \geq x\right), \end{aligned}$$

where the last quantity is the corresponding LDP rate function. Our aim in this paper is to use a similar superadditivity approach, but to bound the DP CGF instead.

²We apply our results to this domain in Section 5.

3. A bound through superadditivity of the CGF. In this section, we explore how superadditivity for the DP CGF can translate a limit on the CGF into a bound. Specifically, our focus lies on the limit given by Varadhan's integral lemma [57]. A frequently encountered formulation of this lemma is the following.

FACT (Varadhan's integral lemma [58]). Let \mathcal{X} be a complete separable metric space. Let $(P_n) \in \mathcal{M}_1(\mathcal{X})^{\mathbb{N}}$ satisfying an LDP with a rate function $I(\cdot)$ and let $\varphi \in \mathcal{C}(\mathcal{X})$. Then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}_{P_n}[\exp(n\varphi)] = \sup_{x \in \mathcal{X}} (\varphi(x) - I(x)).$$

We can use this fact with the continuous mapping $\varphi: \nu \in \mathcal{X} \triangleq \mathcal{M}_1(\Omega) \mapsto \mathbb{E}_\nu[f]$, where $f \in \mathcal{C}(\Omega)$. Indeed, since Ω is compact, $\mathcal{X} = \mathcal{M}_1(\Omega)$ is compact in the weak topology; additionally, it is separable and metrizable, for example, by the Lévy-Prokhorov metric [51]. We thus get the following result for DPs, using Theorem 2.3 with $I(\cdot) = \text{KL}(\nu_0 \|\cdot)$ to get the needed LDP result.

COROLLARY 3.1. Let $f \in \mathcal{C}(\Omega)$. Then,

$$\lim_{\alpha \rightarrow \infty} \frac{1}{\alpha} \log \mathbb{E}_{X \sim \text{DP}(\alpha\nu_0)}[\exp(\mathbb{E}_X[\alpha f])] = \sup_{\nu \in \mathcal{M}_1(\Omega)} (\mathbb{E}_\nu[f] - \text{KL}(\nu_0 \|\nu)).$$

Our next step involves demonstrating the following Lemma 3.2, with the aim of converting the preceding limit into an upper bound.

LEMMA 3.2 (Superadditivity for the DP cumulant-generating function). *Let $f \in \mathcal{C}(\Omega)$. Then, the function $\alpha \mapsto \log \mathbb{E}_{X \sim \text{DP}(\alpha\nu_0)}[\exp(\mathbb{E}_X[\alpha f])]$ is superadditive on $(0, \infty)$.*

The proof of Lemma 3.2 is postponed to section 4. This in turn leads to the following Theorem 3.3.

THEOREM 3.3 (Bound on the cumulant-generating function). *Let $f \in \mathcal{C}(\Omega)$. Then,*

$$\log M_{\text{DP}(\alpha\nu_0)}(f) = \log \mathbb{E}_{X \sim \text{DP}(\alpha\nu_0)}[\exp(\mathbb{E}_X[f])] \leq \sup_{\nu \in \mathcal{M}_1(\Omega)} (\mathbb{E}_\nu[f] - \alpha \text{KL}(\nu_0 \|\nu)).$$

PROOF. The proof is a simple use of the Fekete's lemma [24] coupled with Corollary 3.1. More precisely, Lemma 3.2 and the Fekete's lemma give

$$\lim_{\alpha \rightarrow \infty} \frac{1}{\alpha} \log \mathbb{E}_{X \sim \text{DP}(\alpha\nu_0)}[\exp(\mathbb{E}_X[\alpha f])] = \sup_{\alpha > 0} \frac{1}{\alpha} \log \mathbb{E}_{X \sim \text{DP}(\alpha\nu_0)}[\exp(\mathbb{E}_X[\alpha f])],$$

so from Corollary 3.1, for any $\alpha > 0$,

$$\sup_{\nu \in \mathcal{M}_1(\Omega)} (\mathbb{E}_\nu[f] - \text{KL}(\nu_0 \|\nu)) \geq \frac{1}{\alpha} \log \mathbb{E}_{X \sim \text{DP}(\alpha\nu_0)}[\exp(\mathbb{E}_X[\alpha f])].$$

Multiplying this inequality by α and dividing f by α , we get the desired result. \square

REMARK 3.4. To the best of our knowledge, this bound is new even for the special case of $X \sim \text{Beta}(a, b)$. In this case, we have for $\lambda \geq 0$,

$$(3) \quad \psi(\lambda) \triangleq \log \mathbb{E}[e^{\lambda(X - \mathbb{E}[X])}] \leq \max_{s \in [a/(a+b), 1]} \left(\lambda \left(s - \frac{a}{a+b} \right) - (a+b) \text{kl} \left(\frac{a}{a+b}, s \right) \right),$$

where $\text{kl}(p, q) \triangleq p \log(\frac{p}{q}) + (1 - p) \log(\frac{1-p}{1-q})$. The maximizer in (3) is given by $s = \frac{\lambda - (a+b) + \sqrt{(\lambda - (a+b))^2 + 4\lambda a}}{2\lambda}$. Order-reversing property of the convex conjugate ψ^* of ψ implies that

$$\psi^*(\varepsilon) \geq (a+b) \text{kl}\left(\frac{a}{a+b}, \frac{a}{a+b} + \varepsilon\right),$$

and the Cramer method allows us to reproduce the proof for the tail probability bounds for the Beta distribution in terms of KL-divergence from [22] (see Remark 3.6 for the generalization to DPs). However, let us stress that our result is more general since it controls CGF, thus yielding similar inequalities for sums of independent Beta random variables (see Corollary 3.7 for the generalization to DPs), where the Gamma distribution-based techniques of [22] becomes inapplicable.

REMARK 3.5. In Theorem 3.3, the bound is the convex conjugate of some reversed KL-divergence (as in Theorem 2.3). Since the KL is not symmetric, this is different from the KL used in the Donsker and Varadhan's formula [20] or in the original Sanov theorem [53].

REMARK 3.6. Theorem 3.3 can be used to recover (2), using Sion's theorem ($\mathcal{M}_1(\Omega)$ is a compact convex subset of the topological vector space of finite signed measures on Ω):

$$\begin{aligned} \mathbb{P}_{X \sim \text{DP}(\alpha\nu_0)}[\mathbb{E}_X[f] \geq u] &\leq \inf_{\lambda \geq 0} M_{\text{DP}(\alpha\nu_0)}(\lambda(f - u)) \\ &\leq \exp\left(\inf_{\lambda \geq 0} \sup_{\nu \in \mathcal{M}_1(\Omega)} (\lambda \mathbb{E}_\nu[f - u] - \alpha \text{KL}(\nu_0 \| \nu))\right) \\ &= \exp\left(\sup_{\nu \in \mathcal{M}_1(\Omega)} \inf_{\lambda \geq 0} (\lambda \mathbb{E}_\nu[f - u] - \alpha \text{KL}(\nu_0 \| \nu))\right) \\ &= e^{-\alpha \mathcal{K}_{\text{inf}}(\nu_0, u, f)}, \end{aligned}$$

letting λ go to 0 (resp. ∞) when $\mathbb{E}_\nu[f - u] \geq 0$ (resp. $\mathbb{E}_\nu[f - u] < 0$) in the last equality.

We can go beyond Remark 3.6 with the following result for the sum of independent DPs. Notably, the direct use of the multiplicative behavior of MGF allows us to bypass the need for the representation of DPs via Gamma processes.

COROLLARY 3.7 (Confidence region for independent DPs). Consider $r \in \mathbb{N}^*$, $f_1, \dots, f_r \in \mathcal{C}(\Omega)$ and $\alpha_1 \nu_1, \dots, \alpha_r \nu_r \in \mathcal{M}(\Omega)$. For $\delta \in (0, 1)$, let

$$M_\delta \triangleq \left\{ (\mu_j)_{j \in [r]} \in (\mathcal{M}_1(\Omega))^r, \sum_{j=1}^r \alpha_j \text{KL}(\nu_j \| \mu_j) \leq \log(1/\delta) \right\}.$$

Then,

$$\mathbb{P}_{(X_j) \sim \otimes_{j \in [r]} \text{DP}(\alpha_j \nu_j)} \left[\sum_{j=1}^r \mathbb{E}_{X_j}[f_j] > \sup_{(\mu_j) \in M_\delta} \sum_{j=1}^r \mathbb{E}_{\mu_j}[f_j] \right] \leq \delta.$$

In addition, if $u \in \mathbb{R}$, then,

$$\mathbb{P}_{(X_j) \sim \otimes_{j \in [r]} \text{DP}(\alpha_j \nu_j)} \left[\sum_{j=1}^r \mathbb{E}_{X_j}[f_j] \geq u \right] \leq \exp \left(- \inf_{\substack{(u_j) \in \mathbb{R}^r, \\ \sum_{j \in [r]} u_j = u}} \sum_{j=1}^r \alpha_j \mathcal{K}_{\text{inf}}(\nu_j, u_j, f_j) \right).$$

REMARK 3.8. The bound obtained in the second inequality of Corollary 3.7 matches the LDP rate function obtained using the joint LDP [12] and the contraction principle [19]. However, our bound is non-asymptotic.

PROOF OF COROLLARY 3.7. Let's start by proving the first inequality. From strong duality via Slater condition, we have

$$\sup_{(\mu_j) \in M_\delta} \sum_{j=1}^r \mathbb{E}_{\mu_j} [f_j] = \min_{\lambda \geq 0} g(\lambda) = g(\lambda^*),$$

where

$$g(\lambda) \triangleq \sup_{(\mu_j) \in \mathcal{M}_1(\Omega)^r} \lambda \left(\log(1/\delta) - \sum_{j=1}^r \alpha_j \text{KL}(\nu_j \| \mu_j) \right) + \sum_{j=1}^r \mathbb{E}_{\mu_j} [f_j].$$

If $\lambda^* = 0$, then the result is trivial as

$$\mathbb{P}_{(X_j) \sim \otimes_{j \in [r]} \text{DP}(\alpha_j \nu_j)} \left[\sum_{j=1}^r \mathbb{E}_{X_j} [f_j] > \sup_{(\mu_j) \in M_\delta} \sum_{j=1}^r \mathbb{E}_{\mu_j} [f_j] \right] = 0.$$

Thus, We consider the case $\lambda^* > 0$. From Chernoff inequality and Theorem 3.3,

$$\begin{aligned} & \mathbb{P}_{(X_j) \sim \otimes_{j \in [r]} \text{DP}(\alpha_j \nu_j)} \left[\sum_{j=1}^r \mathbb{E}_{X_j} [f_j] > \sup_{(\mu_j) \in M_\delta} \sum_{j=1}^r \mathbb{E}_{\mu_j} [f_j] \right] \\ & \leq \mathbb{E}_{(X_j) \sim \otimes_{j \in [r]} \text{DP}(\alpha_j \nu_j)} \left[\exp \left(\frac{1}{\lambda^*} \sum_{j=1}^r \mathbb{E}_{X_j} [f_j] \right) \right] \exp \left(-\frac{1}{\lambda^*} \sup_{(\mu_j) \in M_\delta} \sum_{j=1}^r \mathbb{E}_{\mu_j} [f_j] \right) \\ & \leq \exp \left(\sup_{(\mu_j) \in \mathcal{M}_1(\Omega)^r} \sum_{j=1}^r \left(\frac{1}{\lambda^*} \mathbb{E}_{\mu_j} [f_j] - \alpha_j \text{KL}(\nu_j \| \mu_j) \right) - \frac{1}{\lambda^*} \sup_{(\mu_j) \in M_\delta} \sum_{j=1}^r \mathbb{E}_{\mu_j} [f_j] \right) \\ & = \delta \cdot \exp \left(\frac{1}{\lambda^*} g(\lambda^*) - \frac{1}{\lambda^*} \sup_{(\mu_j) \in M_\delta} \sum_{j=1}^r \mathbb{E}_{\mu_j} [f_j] \right) = \delta. \end{aligned}$$

For the second inequality, we have using Sion's minimax theorem

$$\begin{aligned} & \mathbb{P}_{(X_j) \sim \otimes_{j \in [r]} \text{DP}(\alpha_j \nu_j)} \left[\sum_{j=1}^r \mathbb{E}_{X_j} [f_j] \geq u \right] \\ & \leq \exp \left(\inf_{\lambda \geq 0} \sup_{(\mu_j) \in \mathcal{M}_1(\Omega)^r} \lambda \left(-u + \sum_{j=1}^r \mathbb{E}_{\mu_j} [f_j] \right) - \sum_{j=1}^r \alpha_j \text{KL}(\nu_j \| \mu_j) \right) \\ & = \exp \left(\sup_{(\mu_j) \in \mathcal{M}_1(\Omega)^r} \inf_{\lambda \geq 0} \lambda \left(-u + \sum_{j=1}^r \mathbb{E}_{\mu_j} [f_j] \right) - \sum_{j=1}^r \alpha_j \text{KL}(\nu_j \| \mu_j) \right) \\ & = \exp \left(- \inf_{\substack{(u_j) \in \mathbb{R}^r, \\ \sum_{j \in [r]} u_j = u}} \sum_{j=1}^r \alpha_j \mathcal{K}_{\text{inf}}(\nu_j, u_j, f_j) \right). \end{aligned}$$

□

4. Superadditivity for CGF of DPs.

PROOF OF LEMMA 3.2. Without loss of generality, we can assume $0 \leq f \leq 1$ since $f \in \mathcal{C}(\Omega)$ is bounded as Ω is compact.

By the series decomposition of the exponential function, we observe that the desired inequality for $\alpha, \beta > 0$ can be obtained from the following moment inequality: for any integer $k \geq 0$,

$$\mathbb{E}_{(X, X') \sim \text{DP}(\alpha\nu_0) \otimes \text{DP}(\beta\nu_0)} \left[(\mathbb{E}_X[\alpha f] + \mathbb{E}_{X'}[\beta f])^k \right] \leq (\alpha + \beta)^k \mathbb{E}_{X \sim \text{DP}((\alpha+\beta)\nu_0)} \left[\mathbb{E}_X[f]^k \right].$$

This moment inequality can be reformulated as: for all $k \in \mathbb{N}$, $Q_k(f, \dots, f) \leq R_k(f, \dots, f)$, where we define the following symmetric polynomials for f_1, \dots, f_k measurable on Ω :

$$Q_k(f_1, \dots, f_k) \triangleq \mathbb{E}_{(X, X') \sim \text{DP}(\alpha\nu_0) \otimes \text{DP}(\beta\nu_0)} \left[\prod_{\ell \in [k]} (\alpha \mathbb{E}_X[f_\ell] + \beta \mathbb{E}_{X'}[f_\ell]) \right],$$

$$R_k(f_1, \dots, f_k) \triangleq (\alpha + \beta)^k \mathbb{E}_{X \sim \text{DP}((\alpha+\beta)\nu_0)} \left[\prod_{\ell \in [k]} \mathbb{E}_X[f_\ell] \right].$$

Now, let $(U_\ell)_{\ell \in [k]} \stackrel{iid}{\sim} \mathcal{U}([0, 1])$ and define the indicator functions $F_\ell(\omega) \triangleq \mathbb{I}\{U_\ell < f(\omega)\}$. By Fubini's theorem, the multi-linearity of these symmetric polynomials, and using the independence of the uniform random variables, for all $k \in \mathbb{N}$,

$$\mathbb{E}[Q_k(F_1, \dots, F_k)] = Q_k(f, \dots, f), \quad \mathbb{E}[R_k(F_1, \dots, F_k)] = R_k(f, \dots, f).$$

It is thus sufficient to prove $Q_k(F_1, \dots, F_k) \leq R_k(F_1, \dots, F_k)$ with probability 1. Let us consider the ordering $U_{(1)} \geq \dots \geq U_{(k)}$ and let us define the corresponding sets $A_\ell \triangleq \{\omega \in \Omega : U_{(\ell)} < f(\omega)\}$. By Lemma 4.1, for any $\alpha > 0$ and $S \triangleq \{\ell_1, \dots, \ell_{|S|}\} \subset [k] \triangleq \{1, \dots, k\}$, $\ell_1 < \dots < \ell_{|S|}$, we have

$$T(\alpha, S) \triangleq \mathbb{E}_{X \sim \text{DP}(\alpha\nu_0)} \left[\prod_{\ell \in S} X(A_\ell) \right] = \prod_{j=1}^{|S|} \frac{\alpha\nu_0(A_{\ell_j}) + j - 1}{\alpha + j - 1}.$$

In particular, if $k \in S$, then $k = \ell_{|S|}$ and

$$(4) \quad T(\alpha, S) = T(\alpha, S \setminus \{k\}) \cdot \frac{\alpha\nu_0(A_k) + |S| - 1}{\alpha + |S| - 1}.$$

Expanding the products inside the expectation in the definition of Q_k and using the independence between X and X' , we have

$$Q_k(F_1, \dots, F_k) = \sum_{S \subset [k]} \underbrace{\alpha^{|S|} \beta^{k-|S|} T(\alpha, S) \cdot T(\beta, [k] \setminus S)}_{A(k, S)}.$$

By (4), we get

$$\begin{aligned} \sum_{S \subset [k]} A(k, S) &= \sum_{S \subset [k]: k \in S} A(k, S) + \sum_{S \subset [k]: k \notin S} A(k, S) = \sum_{S \subset [k-1]} (A(k, S \cup \{k\}) + A(k, S)) \\ &= \sum_{S \subset [k-1]} \alpha^{|S|} \beta^{k-|S|} \left(\frac{\alpha}{\beta} T(\alpha, S \cup \{k\}) T(\beta, [k-1] \setminus S) + T(\alpha, S) T(\beta, [k] \setminus S) \right) \\ &= \sum_{S \subset [k-1]} A(k-1, S) \left(\alpha \cdot \frac{\alpha\nu_0(A_k) + |S|}{\alpha + |S|} + \beta \cdot \frac{\beta\nu_0(A_k) + k - 1 - |S|}{\beta + k - 1 - |S|} \right). \end{aligned}$$

Finally, we apply Proposition 4.2 for $z = |S| \in [0, k-1]$ and get

$$\alpha \cdot \frac{\alpha \nu_0(A_k) + |S|}{\alpha + |S|} + \beta \cdot \frac{\beta \nu_0(A_k) + k - |S| - 1}{\beta + k - |S| - 1} \leq (\alpha + \beta) \cdot \frac{(\alpha + \beta) \nu_0(A_k) + k - 1}{\alpha + \beta + k - 1},$$

so that $\sum_{S \subset [k]} A(k, S) \leq (\alpha + \beta) \frac{(\alpha + \beta) \nu_0(A_k) + k - 1}{\alpha + \beta + k - 1} \sum_{S \subset [k-1]} A(k-1, S)$. By a simple induction on $k \in \mathbb{N}$, we thus get $\sum_{S \subset [k]} A(k, S) \leq (\alpha + \beta)^k \prod_{\ell \in [k]} \frac{(\alpha + \beta) \nu_0(A_\ell) + \ell - 1}{\alpha + \beta + \ell - 1} = (\alpha + \beta)^k T(\alpha + \beta, [k]) = R_k(F_1, \dots, F_k)$. \square

LEMMA 4.1. *For any increasing sequence $A_1 \subset \dots \subset A_m$, $A_i \in \mathcal{B}(\Omega)$, $m \in \mathbb{N}^*$,*

$$\mathbb{E}_{X \sim \text{DP}(\alpha \nu_0)} \left[\prod_{\ell \in [m]} X(A_\ell) \right] = \prod_{\ell \in [m]} \frac{\alpha \nu_0(A_\ell) + \ell - 1}{\alpha + \ell - 1}.$$

PROOF. We proceed by induction. For $m = 1$, this is simply the expectation formula for the DP. Assume this is true for $m - 1$, for any DP and for any measurable increasing sequence of length $m - 1$. Now, let's fix some measurable increasing sequence $A_1 \subset \dots \subset A_m \subset \Omega$. Consider any finite measurable partition $\sqcup_{i \in [k]} B_i = A_m$. Then, by definition of the DP, $(X(B_1), \dots, X(B_k), 1 - X(A_m)) \sim \text{Dir}(\alpha \nu_0(B_1), \dots, \alpha \nu_0(B_k), \alpha(1 - \nu_0(A_m)))$. By the neutrality property of the Dirichlet distribution (see e.g. [35]), it holds

$$X(A_m) \perp\!\!\!\perp \left(\frac{X(B_1)}{X(A_m)}, \dots, \frac{X(B_k)}{X(A_m)} \right) \sim \text{Dir}(\alpha \nu_0(B_1), \dots, \alpha \nu_0(B_k)).$$

This is true for all finite measurable partition $\sqcup_{i \in [k]} B_i = A_m$, so $X(A_m) \perp\!\!\!\perp \frac{X(\cdot \cap A_m)}{X(A_m)} \sim \text{DP}(\alpha \nu_0(\cdot \cap A_m))$. From the induction hypothesis on $\frac{X(\cdot \cap A_m)}{X(A_m)}$ and the moment formula for the beta distribution, we get

$$\begin{aligned} \mathbb{E}_{X \sim \text{DP}(\alpha \nu_0)} \left[\prod_{\ell \in [m]} X(A_\ell) \right] &= \mathbb{E}_{X \sim \text{DP}(\alpha \nu_0)} \left[\prod_{\ell \in [m-1]} \frac{X(A_\ell)}{X(A_m)} \right] \mathbb{E}_{X \sim \text{DP}(\alpha \nu_0)} [X(A_m)^m] \\ &= \prod_{\ell \in [m-1]} \frac{\alpha \nu_0(A_\ell) + \ell - 1}{\alpha \nu_0(A_m) + \ell - 1} \cdot \prod_{\ell \in [m]} \frac{\alpha \nu_0(A_m) + \ell - 1}{\alpha + \ell - 1} \\ &= \prod_{\ell \in [m]} \frac{\alpha \nu_0(A_\ell) + \ell - 1}{\alpha + \ell - 1}. \end{aligned}$$

\square

PROPOSITION 4.2. Let $s, t, j > 0$ and let $x \in [0, 1]$. For $z \in [0, j]$, we define

$$h(z) \triangleq \frac{(sx + z)s}{s + z} + \frac{(tx + j - z)t}{t + j - z}.$$

Then,

$$\max_{z \in [0, j]} h(z) = h\left(\frac{js}{s+t}\right) = \frac{(s+t)((s+t)x + j)}{s+t+j}.$$

PROOF. We have that h is concave on $[0, j]$, as

$$h'(z) = \frac{s^2(1-x)}{(s+z)^2} - \frac{t^2(1-x)}{(t+j-z)^2}$$

and

$$h''(z) = -\frac{2s^2(1-x)(s+z)}{(s+z)^4} - \frac{2t^2(1-x)(t+j-z)}{(t+j-z)^4} \leq 0.$$

Solving for $h'(z) = 0$, we get that the maximizer of h on $[0, j]$ is $z^* \triangleq js/(s+t)$. Thus,

$$\max_{z \in [0, j]} h(z) = h(z^*) = \frac{(s+t)((s+t)x+j)}{s+t+j}.$$

□

5. Application to the Combinatorial Thompson Sampling policy. Employing multiple independent DPs as described in our Corollary 3.7 has practical applications in various contexts. One notable example is the stochastic semi-bandit problem [37], an extension of the standard multi-armed bandits (MAB) problem. An example of a semi-bandit problem is formulated as follows. We consider a set of n independent Bernoulli distributions (referred to as base arms), each with an unknown mean $\theta_k \in [0, 1]$, $k \in [n] \triangleq \{1, \dots, n\}$. At each round $t \in \mathbb{N}^*$, an agent selects an action $A_t \in \mathcal{A}$, where $\mathcal{A} \subset \mathcal{P}([n])$ is a fixed action space. For each base arm k in action A_t , an outcome $X_{k,t} \sim \text{Ber}(\theta_k)$ is drawn independently from the environment and observed as feedback. The agent gains a reward $\sum_{k \in A_t} X_{k,t}$ before moving to the next round. The agent's goal is to minimize the expected regret over $T \in \mathbb{N}^*$ rounds, $R_T \triangleq T \sum_{k \in A^*} \theta_k - \sum_{t \in [T]} \mathbb{E}[\sum_{k \in A_t} \theta_k]$, where $A^* \in \arg \max_{A \in \mathcal{A}} \sum_{k \in A} \theta_k$.

A commonly used policy for this problem is Combinatorial Thompson Sampling (CTS) [48, 60]. In CTS, each base arm $k \in [n]$ is associated with a maintained prior distribution $\text{Beta}(1 + N_{k,t} \bar{\theta}_{k,t}, 1 + N_{k,t}(1 - \bar{\theta}_{k,t}))$, where, at the beginning of round t , $N_{k,t}$ (resp. $\bar{\theta}_{k,t}$) is the number of observations (resp. empirical mean) of base arm k . At round t , the agent draws, for each base arm k , an independent sample $\theta_{k,t}^+$ from the corresponding prior. Then, the action to be played is chosen as $A_t \in \arg \max_{A \in \mathcal{A}} \sum_{k \in A} \theta_{k,t}^+$ (we assume that linear optimization is computationally efficient over \mathcal{A}).

CTS can be compared with two well-known alternative policies: CUCB [13, 49] and ESCB [14]. Unlike CTS, which is based on sampling, these policies directly build a confidence region C_t , for the vector of outcomes, and then play an action $A_t \in \arg \max_{A \in \mathcal{A}} \max_{(\theta_{1,t}^+, \dots, \theta_{K,t}^+) \in C_t} \sum_{k \in A} \theta_{k,t}^+$. CUCB is conservative but efficient, using the Cartesian product of the individual outcome confidence intervals, whereas ESCB leverages stochastic independence between the base arms but is generally inefficient. CTS strikes a good balance by leveraging independence while remaining efficient. These three policies are often qualified as optimistic, which essentially means that the estimates $\theta_{k,t}^+$ are such that the event $\left\{ \sum_{k \in A^*} \theta_k \leq \sum_{k \in A^*} \theta_{k,t}^+ \right\}$ occurs with high probability.

To compare these policies in terms of expected regret, consider a simple semi-bandit instance where $\mathcal{A} \triangleq \{\{1, \dots, m\}, \{m+1, \dots, 2m\}, \dots, \{n-m+1, \dots, n\}\}$, with $n, m \in \mathbb{N}$ so that $|\mathcal{A}| = n/m \in \mathbb{N}$. We assume that each base arm in an action $j \in [n/m]$ follows an independent Bernoulli distribution of parameter p_j (so $\theta_k = p_{\lceil k/m \rceil}$ for $k \in [n]$), and that $p_1 = \max_{j \in [n/m]} p_j$. This reduces to a MAB problem with n/m actions and with a binomial reward $\text{Bin}(m, p_j)$ for each action $j \in [n/m]$. First, using a result from [38], we have the following lower bound on the asymptotic expected regret of any policy:

$$\liminf_{T \rightarrow \infty} \frac{R_T}{\log T} \geq \sum_{j=2}^{n/m} \frac{m(p_1 - p_j)}{\text{KL}(\text{Bin}(m, p_j) \parallel \text{Bin}(m, p_1))} = \sum_{j=2}^{n/m} \frac{p_1 - p_j}{\text{kl}(p_j, p_1)}.$$

Now, let us examine the upper bounds on $\limsup_{T \rightarrow \infty} R_T / \log T$, focusing on the specific semi-bandit instance described earlier for the sake of simplicity. When considering versions

of the policies based on a KL confidence region, CUCB has an upper bound that is m times larger than the lower bound mentioned earlier [13], whereas ESCB's upper bound matches this lower bound [14]. Since CTS aims to match the statistical performance of ESCB, a natural question arises: Can CTS achieve the same upper bound as ESCB? This question can be addressed using Corollary 3.7, as we will demonstrate next.

Let $\varepsilon > 0$. We have $R_T = \sum_{j=2}^{n/m} \sum_{t=1}^T \mathbb{P}(A_t = \text{action } j) m(p_1 - p_j)$, so it is sufficient to get $\sum_{t=1}^T \mathbb{P}(A_t = \text{action } j) \leq \frac{(1+\varepsilon)\log(T)}{m\text{kl}(p_j, p_1)} + o(\log(T))$. We can thus set aside the rounds where $N_{jm,t} \leq \frac{(1+\varepsilon)\log(T)}{m\text{kl}(p_j, p_1)}$, constituting the leading bound (notice that all the counters $N_{(j-1)m+1,t}, \dots, N_{jm,t}$ are equal). In addition, we can focus on the intersection of several high-probability events, which are listed as follows.

- The optimism event: $mp_1 \leq \sum_{k \in A^*} \theta_{k,t}^+$.
- For all $k \in A_t$, $\bar{\theta}_{k,t} \leq p_1$.
- $(1+\varepsilon) \sum_{k \in A_t} \text{kl}(\bar{\theta}_{k,t}, p_1) \geq \sum_{k \in A_t} \text{kl}(\theta_k, p_1) = m\text{kl}(p_j, p_1)$ (see [41]).

In summary, it is sufficient to show that these events are mutually exclusive for the remaining rounds where $N_{jm,t} > \frac{(1+\varepsilon)\log(T)}{m\text{kl}(p_j, p_1)}$. From the optimism event and the policy's definition we have $mp_1 \leq \sum_{k \in A^*} \theta_{k,t}^+ \leq \sum_{k \in A_t} \theta_{k,t}^+$. From Corollary 3.7, we get that this event holds with a conditional probability bounded by $\exp(-N_{jm,t} \sum_{k \in A_t} \text{kl}(\bar{\theta}_{k,t}, p_1))$. Thus, with high probability, $N_{jm} \sum_{k \in A_t} \text{kl}(\bar{\theta}_{k,t}, p_1) \leq \log(T)$. This, with $N_{jm,t} > \frac{(1+\varepsilon)\log(T)}{m\text{kl}(p_j, p_1)}$, contradicts the last event listed above.

We have demonstrated, using a toy example, how our results can establish the statistical optimality of the CTS policy. These findings could be valuable in more general and practical contexts. Specifically, it would be interesting to show that the expected regret rate of CTS aligns with that of ESCB in problems where ESCB is computationally inefficient.

6. Conclusion. In this paper, we presented a new method for bounding the cumulant generating function (CGF) of Dirichlet Processes (DPs). The proposed non-asymptotic bound achieves asymptotic optimality as $\alpha \rightarrow \infty$. It is expressed as the convex conjugate of α times the large deviation principle rate function for the DP, represented by the reversed Kullback-Leibler divergence. This approach enables the construction of confidence regions for sums of independent DPs, making it useful for various applications.

REFERENCES

- [1] Raghu R Bahadur and SL Zabell, *Large deviations of the sample mean in general vector spaces*, The Annals of probability (1979), 587–621.
- [2] Sergei Bernstein, *On a modification of chebyshev's inequality and of the error formula of laplace*, Ann. Sci. Inst. Sav. Ukraine, Sect. Math **1** (1924), no. 4, 38–49.
- [3] Christopher M Bishop and Nasser M Nasrabadi, *Pattern recognition and machine learning*, vol. 4, Springer, 2006.
- [4] David M Blei and Michael I Jordan, *Variational inference for Dirichlet process mixtures*, Bayesian Analysis **1** (2005), 2006.
- [5] David M Blei, Andrew Y Ng, and Michael I Jordan, *Latent dirichlet allocation*, Journal of machine Learning research **3** (2003), no. Jan, 993–1022.
- [6] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart, *Concentration inequalities: A nonasymptotic theory of independence*, Oxford university press, 2013.
- [7] V Buldygin and K Moskvichova, *The sub-gaussian norm of a binary random variable*, Theory of probability and mathematical statistics **86** (2013), 33–49.
- [8] Valerii V Buldygin and Yu V Kozachenko, *Sub-gaussian random variables*, Ukrainian Mathematical Journal **32** (1980), 483–489.
- [9] Valerii Vladimirovich Buldygin and IU V Kozachenko, *Metric characterization of random variables and random processes*, vol. 188, American Mathematical Soc., 2000.

- [10] Apostolos N Burnetas and Michael N Katehakis, *Optimal adaptive policies for sequential allocation problems*, Advances in Applied Mathematics **17** (1996), no. 2, 122–142.
- [11] George Casella, *Christian robert*, Monte Carlo Statistical Methods (2005).
- [12] Narasinga R Chaganty, *Large deviations for joint distributions and statistical applications*, Sankhyā: The Indian Journal of Statistics, Series A (1997), 147–166.
- [13] Wei Chen, Yajun Wang, Yang Yuan, and Qinshi Wang, *Combinatorial multi-armed bandit and its extension to probabilistically triggered arms*, Journal of Machine Learning Research **17** (2016), no. 50, 1–33.
- [14] Richard Combes, Mohammad Sadegh Talebi Mazraeh Shahi, Alexandre Proutiere, and marc lelarge, *Combinatorial bandits revisited*, Advances in Neural Information Processing Systems 28 (C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, eds.), Curran Associates, Inc., 2015, pp. 2116–2124.
- [15] Peter Congdon, *Applied bayesian modelling*, John Wiley & Sons, 2014.
- [16] H Cramer, *Sur un nouveau theoreme limite de la theorie des probabilités, colloquium on theory of probability*, Paris Hermann Cramw (1937).
- [17] I. Csizar and F. Matus, *Information projections revisited*, IEEE Transactions on Information Theory **49** (2003), no. 6, 1474–1490.
- [18] I. Csizsár and P.C. Shields, *Information theory and statistics: A tutorial*, Foundations and Trends® in Communications and Information Theory **1** (2004), no. 4, 417–528.
- [19] Amir Dembo, *Large deviations techniques and applications*, Springer, 2009.
- [20] Monroe D Donsker and SR Srinivasa Varadhan, *Asymptotic evaluation of certain markov process expectations for large time. iv*, Communications on pure and applied mathematics **36** (1983), no. 2, 183–212.
- [21] Hani Doss and Thomas Sellke, *The tails of probabilities chosen from a dirichlet prior*, The Annals of Statistics **10** (1982), no. 4, 1302–1305.
- [22] Lutz Dumbgen, *New goodness-of-fit tests and their application to nonparametric confidence sets*, Annals of statistics (1998), 288–314.
- [23] Richard S Ellis, *Large deviations for a general class of random vectors*, The Annals of Probability **12** (1984), no. 1, 1–12.
- [24] Michael Fekete, *Über die verteilung der wurzeln bei gewissen algebraischen gleichungen mit ganzzahligen koeffizienten*, Mathematische Zeitschrift **17** (1923), no. 1, 228–249.
- [25] Shui Feng, *Large deviations for dirichlet processes and poisson-dirichlet distribution with two parameters*, Electronic Journal of Probability **12** (2007), no. none, 787 – 807.
- [26] ———, *Hierarchical dirichlet process and relative entropy*, Electronic Communications in Probability **28** (2023), 1–12.
- [27] Thomas S Ferguson, *A bayesian analysis of some nonparametric problems*, The annals of statistics (1973), 209–230.
- [28] Ayalvadi Ganesh and Neil O’Connell, *An inverse of sanov’s theorem*, Statistics & Probability Letters **42** (1999), no. 2, 201–206.
- [29] Ayalvadi J Ganesh and Neil O’connell, *A large-deviation principle for dirichlet posteriors*, Bernoulli (2000), 1021–1034.
- [30] Aurélien Garivier, Hédi Hadiji, Pierre Menard, and Gilles Stoltz, *Kl-ucb-switch: optimal regret bounds for stochastic bandits from both a distribution-dependent and a distribution-free viewpoints*, 2022.
- [31] Jürgen Gärtner, *On large deviations from the invariant measure*, Theory of Probability & Its Applications **22** (1977), no. 1, 24–39.
- [32] Subhashis Ghosal and Aad Van der Vaart, *Fundamentals of nonparametric bayesian inference*, vol. 44, Cambridge University Press, 2017.
- [33] W Hoeffding, *Probability inequalities for sums of bounded random variables*, Journal of the American Statistical Association **58** (1963), 13–30.
- [34] Junya Honda and Akimichi Takemura, *Non-asymptotic analysis of a new bandit algorithm for semi-bounded rewards*, Journal of Machine Learning Research **16** (2015), no. 113, 3721–3756.
- [35] Ian R James and James E Mosimann, *A new characterization of the dirichlet distribution through neutrality*, The Annals of Statistics **8** (1980), no. 1, 183–189.
- [36] Michael Kearns and Lawrence Saul, *Large deviation methods for approximate probabilistic inference*, arXiv preprint arXiv:1301.7392 (2013).
- [37] Branislav Kveton, Zheng Wen, Azin Ashkan, and Csaba Szepesvari, *Tight regret bounds for stochastic combinatorial semi-bandits*, International Conference on Artificial Intelligence and Statistics, 2015.
- [38] Tze L Lai and Herbert Robbins, *Asymptotically efficient adaptive allocation rules*, Advances in Applied Mathematics **6** (1985), no. 1, 4–22.
- [39] Eugene Lukacs, *A characterization of the gamma distribution*, The Annals of Mathematical Statistics **26** (1955), no. 2, 319–324.

- [40] James Lynch and Jayaram Sethuraman, *Large deviations for processes with independent increments*, The annals of probability **15** (1987), no. 2, 610–627.
- [41] Odalric-Ambrym Maillard, Rémi Munos, and Gilles Stoltz, *Finite-Time Analysis of Multi-armed Bandits Problems with Kullback-Leibler Divergences*, To appear in Proceedings of the 24th annual Conference On Learning Theory, COLT '11, 2011.
- [42] Olivier Marchal and Julyan Arbel, *On the sub-gaussianity of the beta and dirichlet distributions*, Electronic Communications in Probability **22** (2017), no. none, 1 – 14.
- [43] Peter Müller, Fernando A Quintana, and Garritt Page, *Nonparametric bayesian inference in applications*, Statistical Methods & Applications **27** (2018), 175–206.
- [44] Kevin P. Murphy, *Probabilistic machine learning: An introduction*, MIT Press, 2022.
- [45] Ian Osband, Daniel Russo, and Benjamin Van Roy, *(more) efficient reinforcement learning via posterior sampling*, Advances in Neural Information Processing Systems (C.J. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K.Q. Weinberger, eds.), vol. 26, Curran Associates, Inc., 2013.
- [46] Ian Osband and Benjamin Van Roy, *Why is posterior sampling better than optimism for reinforcement learning?*, Proceedings of the 34th International Conference on Machine Learning (Doina Precup and Yee Whye Teh, eds.), Proceedings of Machine Learning Research, vol. 70, PMLR, 06–11 Aug 2017, pp. 2701–2710.
- [47] John Paisley, *A simple proof of the stick-breaking construction of the dirichlet process*, Princeton University: Princeton, NJ, USA (2010).
- [48] Pierre Perrault, Etienne Boursier, Vianney Perchet, and Michal Valko, *Statistical Efficiency of Thompson Sampling for Combinatorial Semi-Bandits*, arXiv preprint arXiv:2006.06613 (2020).
- [49] Pierre Perrault, Vianney Perchet, and Michal Valko, *Finding the bandit in a graph: Sequential search-and-stop*, Proceedings of Machine Learning Research (Kamalika Chaudhuri and Masashi Sugiyama, eds.), Proceedings of Machine Learning Research, vol. 89, PMLR, 2019, pp. 1668–1677.
- [50] Gilles Pisier, *Subgaussian sequences in probability and fourier analysis*, arXiv preprint arXiv:1607.01053 (2016).
- [51] Yu V Prokhorov, *Convergence of random processes and limit theorems in probability theory*, Theory of Probability & Its Applications **1** (1956), no. 2, 157–214.
- [52] Maxim Raginsky, Igal Sason, et al., *Concentration of measure inequalities in information theory, communications, and coding*, Foundations and Trends® in Communications and Information Theory **10** (2013), no. 1-2, 1–246.
- [53] Ivan N Sanov, *On the probability of large deviations of random variables*, United States Air Force, Office of Scientific Research, 1958.
- [54] Ivan Nicolaevich Sanov, *On the probability of large deviations of random variables*, Selected Translations in Mathematical Statistics and Probability **1** (1961), 213–244.
- [55] Yee Whye Teh, Michael I Jordan, Matthew J Beal, and David M Blei, *Hierarchical dirichlet processes*, Journal of the American Statistical Association **101** (2006), no. 476, 1566–1581.
- [56] Hans van der Weide, *Gamma processes*, Max MENDEL (1997), 77.
- [57] SR Srinivasa Varadhan, *Asymptotic probabilities and differential equations*, Communications on Pure and Applied Mathematics **19** (1966), no. 3, 261–286.
- [58] ———, *Large deviations and applications*, SIAM, 1984.
- [59] Anatolii Moiseevich Vershik, Marc Yor, and Natalia Vladimirovna Tsilevich, *Remarks on the markov–krein identity and quasi-invariance of the gamma process*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov.(POMI) **283** (2001), 21–36.
- [60] Siwei Wang and Wei Chen, *Thompson sampling for combinatorial semi-bandits*, International Conference on Machine Learning, PMLR, 2018, pp. 5114–5122.

