JMLR: Workshop and Conference Proceedings 1–31

Minimax Number of Strata for Online Stratified Sampling given Noisy Samples

Alexandra Carpentier ALEXANDRA.CARPENTIER@INRIA.FR
INRIA Lille Nord-Europe, 40, avenue Halley, 59650 Villeneuve d'Ascq, France

Rémi Munos REMI.MUNOS@INRIA.FR
INRIA Lille Nord-Europe, 40, avenue Halley, 59650 Villeneuve d'Ascq, France

Abstract

We consider the problem of online stratified sampling for Monte Carlo integration of a function given a finite budget of n noisy evaluations to the function. More precisely we focus on the problem of choosing the number of strata K as a function of the budget n. We provide asymptotic and finite-time results on how an oracle that has access to the function would choose the partition optimally. In addition we prove a *lower bound* on the learning rate for the problem of stratified Monte-Carlo. As a result, we are able to state, by improving the bound on its performance, that algorithm MC-UCB, defined in (Carpentier and Munos, 2011a), is minimax optimal both in terms of the number of samples n and the number of strata K, up to a $\sqrt{\log(nK)}$. This enables to deduce a minimax optimal bound on the difference between the performance of the estimate outputted by MC-UCB, and the performance of the estimate outputted by the best oracle static strategy, on the class of Hölder continuous functions, and upt to a $\sqrt{\log(n)}$.

Keywords: Online learning, stratified sampling, Monte Carlo integration, regret bounds.

1. Introduction

The objective of this paper is to provide an efficient strategy for Monte-Carlo integration of a function f over a domain $[0,1]^d$. We assume that we can query the function n times. Querying the function at a time t and at a point $x_t \in [0,1]^d$ provides a noisy sample

$$\boldsymbol{f(x_t) + s(x_t)\epsilon_t}, \tag{1}$$

where ϵ_t is an independent sample drawn from ν_{x_t}. Here ν_x is a distribution with mean 0, variance 1 and whose shape may depend on x[1]. This model is actually very general (see Section 2).

Stratified sampling is a well-known strategy to reduce the variance of the estimate of the integral of f, when compared to the variance of the estimate provided by crude Monte-Carlo. The principle is to partition the domain in K subsets called *strata* and then to sample in each stratum (see (Rubinstein and Kroese, 2008)[Subsection 5.5] or (Glasserman, 2004)). If the variances of the strata are known, there exists an optimal static allocation strategy which allocates the samples proportionally to the measure of the stratum times their standard deviation (see Equation 3 in this paper for a reminder). We refer to this

1. It is the usual model for functions in heterocedastic noise. We isolate the standard deviation on a point x, $s(x)$, in the expression of the noise, since this quantity is very relevant.

© A. Carpentier & R. Munos.

allocation as optimal oracle strategy for a given partition. In the case that the variations of f and the standard deviation of the noise s are unknown, it is not possible to adopt this strategy.

Consider first that the partition of the space is fixed. A way around this problem is to estimate the variations of the function and the amount of noise on the function in the strata *online* (exploration) while allocating the samples according to the estimated optimal oracle proportions (exploitation). This setting is considered in (Etoré and Jourdain, 2010; Grover, 2009; Carpentier and Munos, 2011a). In the long version (Carpentier and Munos, 2011b) of the last paper, the authors propose the so-called MC-UCB algorithm which is based on Upper-Confidence-Bounds (UCB) on the standard deviation. They provide upper bounds for the difference between the mean-squared error[2] of the estimate provided by MC-UCB and the mean-squared error of the estimate provided by the optimal oracle strategy (optimal oracle variance). The algorithm performs almost as well as the optimal oracle strategy. However, the authors of (Carpentier and Munos, 2011b) do not infirm nor assess the optimality of their algorithm with a lower bound as benchmark. As a matter of fact, no lower bound on the rate of convergence (to the oracle optimal strategy) for the problem of stratified Monte-Carlo exists, to the best of our knowledge. Still in the same paper (Carpentier and Munos, 2011b), the authors do not at all discuss on how to *stratify* the space. In particular, they do not pose the problem of what an *optimal oracle* partition of the space is, and do not try to answer on whether it is possible or not to attain it.

The next step is thus to efficiently design the partition. There are some interesting papers on that topic such that (Glasserman et al., 1999; Kawai, 2010; Etoré et al., 2011). The recent, state of the art, work of Etoré et al. (2011) describes a strategy that samples *asymptotically* almost as efficiently as the optimal oracle strategy, and at the same time adapts the direction and number of the strata online. This is a very difficult problem. The authors do not provide proofs of convergence of their algorithm. However for static allocation of the samples, they present some properties of the stratified estimate when the number of strata goes to infinity and provide convergence results under the optimal oracle strategy. As a corollary, they prove that the more strata there are, the smallest the optimal oracle variance.

Contributions: The more strata there are, the smaller the variance of the estimate computed when following the optimal oracle strategy. However, the more strata there are, the more difficult it is to estimate the variance within each of these strata, and thus the more difficult it is to perform almost as well as the optimal oracle strategy. Choosing the number of strata is thus crucial and this is the problem we address in this paper. This defines a trade-off similar to the one in model selection (and in all its variants, e.g. density estimation, regression...): The wider the class of models considered, i.e. the larger the number of strata, the smaller the distance between the true model and the best model of the class, i.e. the approximation error. But the larger the estimation error.
Paper (Etoré et al., 2011), although proposing no finite time bounds, develops very interesting ideas for bounding the first term, i.e. the approximation error. As pointed out in paper e.g. (Carpentier and Munos, 2011a), it is possible to build algorithms that have a small estimation error. By constructing tight and finite-time bounds for the approximation

2. The mean squared error is measured with respect to the quantity of interest, i.e. the integral of f.

error, it is thus possible to propose a number of strata that minimizes an upper bound on the performance. It is however not clear how consistent this choice is, i.e. how much it can be improved. The essential ingredients for choosing efficiently a partition are thus lower bounds *on the estimation error, and on the approximation error*.

The objective of this paper is to propose a method for choosing the minimax-optimal number of strata. Our contributions are the following.

- We first present results on what we call the *quality* $Q_{n,\mathcal{N}}$ of a given partition in K strata $\mathcal{N}$ (i.e., using the previous analogy to model selection, this would represent the approximation error). Using very mild assumptions we compute a lower bound on the variance of the estimate given by the optimal oracle strategy on the optimal oracle partition. Then if the function and the standard deviation of the noise are $\alpha-$Hölder, and also if the strata satsfy some assumptions, we prove that $Q_{n,\mathcal{N}} = O(\frac{K^{\alpha/d}}{n})$. This bound is also minimax optimal on the class of $\alpha-$Hölder functions.

- We then present results on the estimation error for the estimate outputted by algorithm MC-UCB of (Carpentier and Munos, 2011a) (pseudo-regret in the terminology of (Carpentier and Munos, 2011a)). In this paper, we improve the analysis of the MC-UCB algorithm when compared to paper (Carpentier and Munos, 2011a) in terms of the dependence on K. The problem independent bound on the pseudo-regret in (Carpentier and Munos, 2011a) is of order[3] $\tilde{O}(Kn^{-4/3})$, and we tighten this bound in this paper so that it is of order $\tilde{O}(K^{1/3}n^{-4/3})$.

- We provide the first *lower bound* (on the pseudo-regret) for the problem of online Stratified Sampling. The bound $\Omega(K^{1/3}n^{-4/3})$ is tight and *matches the upper-bound of MC-UCB both in terms of the number of strata and the number of samples*. This is the main contribution of the paper, and we believe that the proof technique for this bound is original.

- Finally, we combine the results on the quality and on the pseudo-regret of MC-UCB to provide a value on the number of strata leading to a minimax-optimal trade-off (up to a $\sqrt{\log(n)}$) on the class of $\alpha-$Hölder functions.

The rest of the paper is organized as follows. In Section 2 we formalize the problem and introduce the notations used throughout the paper. Section 3 states the results on the quality of a partition. Section 4 improves the analysis of the MC-UCB algorithm, and establishes the lower bound on the pseudo-regret. Section 5 reports the best trade-off to choose the number of strata. And in Section 6, we illustrate how important it is to choose carefully the number of strata. We finally conclude the paper and suggest future works.

2. Setting

We consider the problem of numerical integration of a function $f : [0,1]^d \to \mathbb{R}$ with respect to the uniform (Lebesgue) measure. We dispose of a budget of n queries (samples) to the function, and we can allocate this budget *sequentially*. When querying the function at a time t and at a point x_t, we receive a noisy sample $X(t)$ of the form described in Equation 1.

3. Here $\tilde{O}$ is a O up to a polynomial $\log(n)$.

We now assume that the space is stratified in K Lebesgue measurable strata that form a partition $\mathcal{N}$. We index these strata, called Ω_k, with indexes $k \in \{1, \dots, K\}$, and write w_k their measure, according to the Lebesgue measure. We write $\mu_k = \frac{1}{w_k} \int_{\Omega_k} \mathbb{E}_{\epsilon \sim \nu_x}[f(x) + s(x)\epsilon] dx = \frac{1}{w_k} \int_{\Omega_k} f(x) dx$ their mean and $\sigma_k^2 = \frac{1}{w_k} \int_{\Omega_k} \mathbb{E}_{\epsilon \sim \nu_x}[(f(x) + s(x)\epsilon - \mu_k)^2] dx$ their variance. These mean and variance correspond to the mean and variance of the random variable $X(t)$ when the coordinate x at which the noisy evaluation of f is observed is chosen uniformly at random on the stratum Ω_k.

We denote by $\mathcal{A}$ an algorithm that allocates online the budget by selecting at each time step $1 \leq t \leq n$ the index $k_t \in \{1, \dots, K\}$ of a stratum and then sampling uniformly the corresponding stratum Ω_{k_t}. The objective is to return the best possible estimate $\hat{\mu}_n$ of the integral of the function f. We write $T_{k,n} = \sum_{t \leq n} \mathbb{I}\{k_t = k\}$ the number of samples in stratum Ω_k up to time n. We denote by $\left(X_{k,t}\right)_{1 \leq k \leq K, 1 \leq t \leq T_{k,n}}$ the samples in stratum Ω_k, and we define $\hat{\mu}_{k,n} = \frac{1}{T_{k,n}} \sum_{t=1}^{T_{k,n}} X_{k,t}$ the empirical means. We estimate the integral of f by $\hat{\mu}_n = \sum_{k=1}^{K} w_k \hat{\mu}_{k,n}$.

If we allocate a deterministic number of samples T_k to each stratum Ω_k and if the samples are independent and chosen uniformly on each stratum Ω_k, we have

$$\mathbb{E}(\hat{\mu}_n) = \sum_{k \leq K} w_k \mu_k = \sum_{k \leq K} \int_{\Omega_k} f(u) du = \int_{[0,1]^d} f(u) du = \mu,$$

and also

$$\mathbb{V}(\hat{\mu}_n) = \sum_{k \leq K} \frac{w_k^2 \sigma_k^2}{T_k},$$

where the expectation and the variance are computed according to all the samples that the algorithm collected.

For a given algorithm $\mathcal{A}$ allocating $T_{k,n}$ samples drawn *uniformly* within stratum Ω_k, we denote by *pseudo-risk* the quantity

$$L_{n,\mathcal{N}}(\mathcal{A}) = \sum_{k \leq K} \frac{w_k^2 \sigma_k^2}{T_{k,n}}. \tag{2}$$

Note that if an algorithm $\mathcal{A}^*$ has access the variances σ_k^2 of the strata, it can choose to allocate the budget in order to minimize the pseudo-risk, i.e., sample each stratum $T_k^* = \frac{w_k \sigma_k}{\sum_{i \leq K} w_i \sigma_i} n$ times (this is the so-called oracle allocation). These optimal numbers of samples can be non-integer values, in which case the proposed optimal allocation is not realizable. But we still use it as a benchmark. The pseudo-risk for this algorithm (which is also the variance of the estimate here since the sampling strategy is deterministic) is then

$$L_{n,\mathcal{N}}(\mathcal{A}^*) = \frac{\left(\sum_{k \leq K} w_k \sigma_k\right)^2}{n} = \frac{\Sigma_{\mathcal{N}}^2}{n}, \tag{3}$$

where $\Sigma_{\mathcal{N}} = \sum_{k \leq K} w_k \sigma_k$. We also refer in the sequel as optimal proportion to $\lambda_k = \frac{w_k \sigma_k}{\sum_{i \leq K} w_i \sigma_i}$, and to optimal oracle strategy to this allocation strategy. Although, as already

mentioned, the optimal allocations (and thus the optimal pseudo-risk) might not be realizable, it is still very useful in providing a lower-bound. No static (even oracle) algorithm has a pseudo-regret lower than $L_{n,\mathcal{N}}(\mathcal{A}^*)$ on partition $\mathcal{N}$.

It is straightforward to see that the more refined the partition $\mathcal{N}$ the smaller $L_{n,\mathcal{N}}(\mathcal{A}^*)$. We thus define the *quality of a partition* $Q_{n,\mathcal{N}}$ as the difference between the variance $L_{n,\mathcal{N}}(\mathcal{A}^*)$ of the estimate provided by the optimal oracle strategy on partition $\mathcal{N}$, and the infimum of the variance of the optimal oracle strategy on *any* partition (optimal oracle partition) (with an arbitrary number of strata):

$$Q_{n,\mathcal{N}} = L_{n,\mathcal{N}}(\mathcal{A}^*) - \inf_{\mathcal{N}' measurable} L_{n,\mathcal{N}'}(\mathcal{A}^*). \tag{4}$$

We also define the *pseudo-regret* of an algorithm $\mathcal{A}$ on a given partition $\mathcal{N}$, the difference between its pseudo-risk and the variance of the optimal oracle strategy:

$$R_{n,\mathcal{N}}(\mathcal{A}) = L_{n,\mathcal{N}}(\mathcal{A}) - L_{n,\mathcal{N}}(\mathcal{A}^*). \tag{5}$$

We will assess the performance of an algorithm $\mathcal{A}$ by comparing its pseudo risk to the minimum possible variance of an optimal oracle strategy on the optimal oracle partition:

$$L_{n,\mathcal{N}}(\mathcal{A}) - \inf_{\mathcal{N}' measurable} L_{n,\mathcal{N}'}(\mathcal{A}^*) = R_{n,\mathcal{N}}(\mathcal{A}) + Q_{n,\mathcal{N}}. \tag{6}$$

Using the analogy of model selection mentioned in the Introduction, the quality $Q_{n,\mathcal{N}}$ is similar to the approximation error and the pseudo-risk $R_{n,\mathcal{N}}(\mathcal{A})$ to the estimation error.

Motivation for the model $f(x) + s(x)\epsilon_t$. Assume that a learner can, at each time t, choose a point x and collect an observation $F(x, W_t)$, where W_t is an independent noise, that can however depend on x. It is the general model for representing evaluations of a noisy function. There are many settings where one needs to integrate accurately a noisy function without wasting too much budget, like for instance pollution survey. Set $f(x) = \mathbb{E}_{W_t}[F(x, W_t)]$, and $s(x)\epsilon_t = F(x, W_t) - f(x)$. Since by definition ϵ_t is of mean 0 and variance 1, we have in fact $s(x) = \sqrt{\mathbb{E}_{\nu_x}[(F(x, W_t) - f(x))^2]}$ and $\epsilon_t = \frac{F(x,W_t)-f(x)}{s(x)}$. Observing $F(x, W_t)$ is equivalent to observing $f(x) + s(x)\epsilon_t$, and this implies that the model that we choose is also very general.

There is also a important setting where this model is relevant, and this is for the integration of a function F in high dimension d^*. Stratifying in dimension d^* seems hopeless, since the budget n has to be exponential with d^* if one wants to stratify in every direction of the domain: this is the curse of dimensionality. It is necessary to reduce the dimension by choosing *a small amount* of directions $(1, \dots, d)$ that are particularly relevant, and control/stratify only in these d directions[4]. Then the control/stratification is only on the first d coordinates, so when sampling at at a time t, one chooses $x = (x_1, \dots, x_d)$, and the other $d^* - d$ coordinates $U(t) = (U_{d+1}(t), \dots, U_{d^*}(t))$ are uniform random variables on $[0, 1]^{d^*-d}$ (without any control). When sampling in x at a time t, we observe $F(x, U(t))$. By writing $f(x) = \mathbb{E}_{U(t)\sim\mathcal{U}([0,1]^{d^*-d})}[F(x, U(t))]$, and $s(x)\epsilon_t = F(x, U(t)) - f(x)$, we obtain that the model we propose is also valid in this case.

4. This is actually a very common technique for computing the price of options, see (Glasserman, 2004).

3. The quality of a partition: Analysis of the term $Q_{n,\mathcal{N}}$.

In this Section, we focus on the *quality* of a partition defined in Section 2.

Convergence under very mild assumptions As mentioned out in Section 2, the more refined the partition $\mathcal{N}$ of the space, the smaller $L_{n,\mathcal{N}}(\mathcal{A}^*)$, and thus $\Sigma_{\mathcal{N}}$. Through this monotony property, we know that $\inf_{\mathcal{N}} \Sigma_{\mathcal{N}}$ is also the limit of the $(\Sigma_{\mathcal{N}_p})_p$ of a sequence of partitions $(\mathcal{N}_p)_p$ such that the diameter of each stratum goes to 0. We state in the following Proposition that for *any* such sequence, $\lim_{p\to+\infty} \Sigma_{\mathcal{N}_p} = \int_{[0,1]^d} s(x)dx$. Consequently $\inf_{\mathcal{N}} \Sigma_{\mathcal{N}} = \int_{[0,1]^d} s(x)dx$.

Proposition 1 *Let $(\mathcal{N}_p)_p = (\Omega_{k,p})_{k\in\{1,\dots,K_p\},p\in\{1,\dots,+\infty\}}$ be a sequence of measurable partitions (where K_p is the number of strata of partition $\mathcal{N}_p$) such that*

- AS1: *$0 < w_{k,p} \leq \upsilon_p$, for some sequence $(\upsilon_p)_p$, where $\upsilon_p \to 0$ for $p \to +\infty$.*
- AS2: *The diameters according to the $||.||_2$ norm on $\mathbb{R}^d$ of the strata are such that $\max_k Diam(\Omega_{k,p}) \leq D(w_{k,p})$, for some real valued function $D(\cdot)$, such that $D(w) \to 0$ for $w \to 0$.*

If the functions m and s are in $\mathbb{L}_2([0,1]^d)$, then

$$\lim_{p\to+\infty} \Sigma_{\mathcal{N}_p} = \inf_{\mathcal{N} measurable} \Sigma_{\mathcal{N}} = \int_{[0,1]^d} s(x)dx,$$

which implies that $n \times Q_{n,\mathcal{N}_p} \to 0$ for $p \to +\infty$.

Proof [Sketch of Proof. The full proof is in the Supplementary material (Appendix B)] The form of the model and the definition of σ_k imply that

$$\sigma_k^2 = \frac{1}{w_k}\int_{\Omega_k} \big(f(x) - \frac{1}{w_k}\int_{\Omega_k} f(u)du\big)^2 dx + \frac{1}{w_k}\int_{\Omega_k} s(x)^2 dx. \tag{7}$$

We first prove that the result hold for uniformly continuous functions, and then generalize to $\mathbb{L}_2$ functions based on a density argument.

Step 1: Convergence when m and s are uniformly continuous: Assume that m and s are uniformly continuous with respect to the $||.||_2$ norm. For any $\upsilon > 0$, there exists η *s.t.* $\forall x, |s(x+u) - s(x)| \leq \upsilon$ and $|f(x+u) - f(x)| \leq \upsilon$ where $u \in \mathcal{B}_{2,d}(\eta)$. We choose K large enough so that the size of the strata is smaller than υ, and their diameter is smaller than η (it is possible to do so since the diameter of the strata shrinks to 0 as $K \to \infty$). From Equation 7 we deduce that

$$\sigma_k^2 - (\frac{1}{w_k}\int_{\Omega_k} s)^2 \leq 2\upsilon^2,$$

and using the concavity of the square-root function, we have $\sum_k w_k\sigma_k - \int_{[0,1]^d} s \leq \sqrt{2}\upsilon$, which concludes the proof for uniformly continuous functions.

Step 2: Generalization to the case where f and s are in $\mathbb{L}_2([0,1)^d)$: From the density property of the uniformly continuous functions in $\mathbb{L}_2([0,1]^d)$ (with respect to the $||.||_2$ norm),

we deduce that for any K and υ, there exists two uniformly continuous function f_υ and s_υ such that:

$$\left| \sum_{k=1}^{K} w_k \sigma_k - \sum_{k=1}^{K} \sqrt{w_k} \sqrt{\int_{\Omega_k} \left(f_\upsilon(x) + \int_{\Omega_k} f_\upsilon(u)du \right)^2 dx - \frac{1}{w_k} \int_{\Omega_k} s_\upsilon^2(x)dx} \right| \leq \upsilon,$$

and also that $\int_\Omega |s(x) - s_\upsilon(x)|dx \leq \sqrt{\frac{\upsilon}{2}}$. One concludes by combining those two inequalities with Step 1. ■

In Proposition 1, even though the optimal oracle allocation might not be realizable (in particular if the number of strata is larger than the budget), we can still compute the quality of a partition, as defined in 4. It does not correspond to any reachable pseudo-risk, but rather to a lower bound on any (even oracle) static allocation.

When f and s are in $\mathbb{L}_2([0,1]^d)$, for any appropriate sequence of partitions $(\mathcal{N}_p)_p$, $\Sigma_{\mathcal{N}_p}$ (which is the principal ingredient of the variance of the optimal oracle allocation) converges to the smallest possible $\Sigma_{\mathcal{N}}$ for given f and s. Note however that this condition is not sufficient to obtain a *rate*.

Finite-Time analysis under Hölder assumption: We make the following assumption on the functions f and s.

Assumption 1 *The functions f and s are $(M,\alpha)-$Hölder continuous, i.e., for $g \in \{m, s\}$, for any x and $y \in [0,1]^d$, $|g(x) - g(y)| \leq M||x-y||_2^\alpha$.*

The Hölder assumption enables to consider arbitrarily non-smooth functions (for small α, the function can vary arbitrarily fast), and is thus a fairly general assumption.

We also consider the following partitions in K squared strata.

Definition 2 *We write $\mathcal{N}_K$ the partition of $[0,1]^d$ in K hyper-cubic strata of measure $w_k = w = \frac{1}{K}$ and side length $(\frac{1}{K})^{1/d}$: we assume for simplicity that there exists an integer l such that $K = l^d$.*

The following Proposition holds.

Proposition 3 *Under Assumption 1 we have for any partition $\mathcal{N}_K$ as defined in Definition 2 that*

$$\Sigma_{\mathcal{N}_K} - \int_{[0,1]^d} s(x)dx \leq \sqrt{2d}M(\frac{1}{K})^{\alpha/d}, \tag{8}$$

which implies

$$Q_{n,\mathcal{N}_K} \leq \frac{2\sqrt{2d}M\Sigma_{\mathcal{N}_1}}{n}(\frac{1}{K})^{\alpha/d},$$

where $\mathcal{N}_1$ stands for the "partition" with one stratum.

Proof [Sketch of Proof for Proposition 3] We deduce from Assumption 1 that

$$\frac{1}{w_k}\int_{\Omega_k}\left(f(x) - \frac{1}{w_k}\int_{\Omega_k} f(u)du\right)^2 dx + \frac{1}{w_k}\int_{\Omega_k} s^2(x)dx - \left(\frac{1}{w_k}\int_{\Omega_k} s(u)du\right)^2 \leq 2M^2 d(\frac{1}{K})^{2\alpha/d}.$$

Then, by using Equation 7 and by summing over all strata, we deduce Equation 8. Now the result on the quality follows from the fact that $\Sigma_{\mathcal{N}_K}^2 - \left(\int_{[0,1]^d} s(x)dx\right)^2 = (\Sigma_{\mathcal{N}_K}^2 - \left(\int_{[0,1]^d} s(x)dx\right)^2)(\Sigma_{\mathcal{N}_K}^2 + \left(\int_{[0,1]^d} s(x)dx\right)^2) \leq 2\Sigma_{\mathcal{N}_1}(\Sigma_{\mathcal{N}_K} - \int_{[0,1]^d} s(x)dx)$. ■

The full proof for this Proposition is in the Supplementary Material (Appendix C).

3.1. General comments

The impact of α and d: The quantity $Q_{n,\mathcal{N}_K}$ increases with the dimension d, because the Hölder assumption becomes less constraining when d increases. This can easily be seen since a squared strata of measure w has a diameter of order $w^{1/d}$. $Q_{n,\mathcal{N}_K}$ decreases with the smoothness α of the function, which is a logic effect of the Hölder assumption. Note also that when defining the partitions $\mathcal{N}_K$ in Definition 2, we made the crucial assumption that $K^{1/d}$ is an integer. This fact is of little importance in small dimension, but will matter in high dimension, as we will enlighten in the last remark of Section 5.

Minimax optimality of this rate: The rate $n^{-1}K^{-\alpha/d}$ is minimax optimal on the class of $\alpha-$Hölder functions since for any n and K one can easily build a function with Hölder exponent α such that the corresponding $\Sigma_{\mathcal{N}_K}$ is at least $\int_{[0,1]^d} s(x)dx + cK^{-\alpha/d}$ for some constant c.

Discussion on the shape of the strata: Whatever the shape of the strata, as long as their diameter goes to 0[5], $\Sigma_{\mathcal{N}_K}$ converges to $\int_{[0,1]^d} s(x)dx$. The shape of the strata have an influence only on the negligible term, i.e. the speed of convergence to this quantity. This result was already made explicit, in a different setting and under different assumptions, in (Etoré et al., 2011). Choosing small strata of same shape and size is also minimax optimal on the class of Hölder functions. Working on the shape of the strata could, however, improve the speed of convergence in some specific cases, e.g. when the noise is very localized. It could also be interesting to consider strata of varying size, and make this size depend on the specific problem.

The decomposition of the variance: Note that the variance σ_k^2 within each stratum Ω_k comes from two sources. First, σ_k^2 comes from the noise, that contributes to it by $\frac{1}{w_k}\int_{\Omega_k} s(x)^2dx$. Second, the mean f is not a constant function, thus its contribution to σ_k^2 is $\frac{1}{w_k}\int_{\Omega_k}\left(f(x) - \frac{1}{w_k}\int_{\Omega_k} f(u)du\right)^2dx$. Note that when the size of Ω_k goes to 0, this later contribution vanishes, and the optimal allocation is thus proportional to $\sqrt{w_k\int_{\Omega_k} s(x)^2dx + o(1)} = \int_{\Omega_k} s(x)dx + o(1)$. This means that for small strata, the variation in the mean are negligible when compared to the variation due to the noise.

4. Algorithm MC-UCB and a matching lower bound

4.1. Algorithm $MC-UCB$

In this Subsection, we describe a slight modification of the algorithm $MC-UCB$ introduced in (Carpentier and Munos, 2011a). The only difference is that we change the form of the

5. And note that in this *noisy* setting, if the diameter of the strata does not go to 0 on non homogeneous part of m and s, then the standard deviation corresponding to the allocation is larger than $\int_{[0,1]^d} s(u)du$.

high-probability upper confidence bound on the standard deviations, in order to improve the elegance of the proofs, and we refine their analysis. The algorithm takes as input two parameters b and $f_{\max}$ which are linked to the distribution of the arms, δ which is a (small) probability, and the partition $\mathcal{N}_K$. We remind in Figure 1 the algorithm $MC-UCB$.

Input: b, $f_{\max}$, δ, $\mathcal{N}_K$, set $A = 2\sqrt{(1+3b+4f_{\max}^2)\log(2nK/\delta)}$
Initialize: Sample 2 states in each strata.
for $t = 2K+1, \dots, n$ **do**
  Compute $B_{k,t} = \frac{w_k}{T_{k,t-1}}\left(\hat{\sigma}_{k,t-1} + A\sqrt{\frac{1}{T_{k,t-1}}}\right)$ for each stratum $k \leq K$
  Sample a point in stratum $k_t \in \arg\max_{1\leq k\leq K} B_{k,t}$
end for
Output: $\hat{\mu}_n = \sum_{k=1}^K w_k \hat{\mu}_{k,n}$

Figure 1: The pseudo-code of the MC-UCB algorithm. The empirical standard deviations and means $\hat{\sigma}_{k,t}^2$ and $\hat{\mu}_{k,t}$ are computed using Equations 9 and 10.

The estimates of $\hat{\sigma}_{k,t-1}^2$ and $\hat{\mu}_{k,t-1}$ are computed according to

$$\hat{\sigma}_{k,t-1}^2 = \frac{1}{T_{k,t-1}} \sum_{i=1}^{T_{k,t-1}} (X_{k,i} - \hat{\mu}_{k,t-1})^2 \,, \tag{9}$$

and

$$\hat{\mu}_{k,t-1} = \frac{1}{T_{k,t-1}} \sum_{i=1}^{T_{k,t-1}} X_{k,i} \,. \tag{10}$$

4.2. Upper bound on the pseudo-regret of algorithm MC-UCB.

We first state the following Assumption on the noise ϵ_t:

Assumption 2 *There exist $b > 0$ such that $\forall x \in [0,1]^d$, $\forall t$, and $\forall \lambda < \frac{1}{b}$,*

$$\mathbb{E}_{\nu_x}\Big[\exp(\lambda\epsilon_t)\Big] \leq \exp\Big(\frac{\lambda^2}{2(1-\lambda b)}\Big), \quad \textit{and} \quad \mathbb{E}_{\nu_x}\Big[\exp(\lambda\epsilon_t^2 - \lambda)\Big] \leq \exp\Big(\frac{\lambda^2}{2(1-\lambda b)}\Big).$$

This is a kind of sub-Gaussian assumption, satisfied for e.g., Gaussian as well as bounded distributions. We also state an assumption on f and s.

Assumption 3 *The functions f and s are bounded by $f_{\max}$.*

Note that since the functions f and s are defined on $[0,1]^d$, if Assumption 1 is satisfied, then Assumption 3 holds with $f_{\max} = \max(f(0), s(0)) + Md^{\alpha/2}$. We now prove the following bound on the pseudo-regret. Note that we state it on partitions $\mathcal{N}_K$, but that it in fact holds for any partition in K strata.

Proposition 4 *Under Assumptions 2 and 3, on partition $\mathcal{N}_K$, when $n \geq 4K$, we have*

$$\mathbb{E}[R_{n,\mathcal{N}_K}(\mathcal{A}_{MC-UCB})] \leq 24\sqrt{2}\Sigma_{\mathcal{N}_K}\sqrt{(1+3b+4f_{\max}^2)}\Big(\frac{f_{\max}+4}{4}\Big)^{1/3}\frac{K^{1/3}}{n^{4/3}}\sqrt{\log(nK)}+\frac{14K\Sigma_{\mathcal{N}_K}^2}{n^2}.$$

The proof, given in the Supplementary Material (Appendix A), is close to the one of MC-UCB in (Carpentier and Munos, 2011a). But an improved analysis leads to a better dependency in terms of number of strata K. We remind that in paper (Carpentier and Munos, 2011a), the bound is of order $\tilde{O}(Kn^{-4/3})$. This improvement is crucial here since the larger K is, the closer $\Sigma_{\mathcal{N}_K}$ is from $\int_{[0,1]^d} s(x)dx$. The next Subsection states that the rate $K^{1/3}\tilde{O}(n^{-4/3})$ of MC-UCB is optimal both in terms of K and n.

4.3. Lower Bound

We now study the minimax rate for the pseudo-regret of any algorithm on a given partition $\mathcal{N}_K$. Note that we state it for partitions $\mathcal{N}_K$, but that it holds for any partition in K strata of equal measure.

Theorem 5 *Let $K \in \mathbb{N}$. Let* inf *be the infimum taken over all online stratified sampling algorithms on $\mathcal{N}_K$ and* sup *represent the supremum taken over all environments, then:*

$$\inf \sup \mathbb{E}[R_{n,\mathcal{N}_K}] \geq C\frac{K^{1/3}}{n^{4/3}},$$

where C is a numerical constant.

Proof [Sketch of proof (The full proof is reported in Appendix D)] We consider a partition with $2K$ strata. On the K first strata, the samples are drawn from Bernoulli distributions of parameter μ_k where $\mu_k \in \{\frac{\mu}{2}, \mu, 3\frac{\mu}{2}\}$, and on the K last strata, the samples are drawn from a Bernoulli of parameter $1/2$. We write $\sigma = \sqrt{\mu(1-\mu)}$ the standard deviation of a Bernoulli of parameter μ. We index by υ a set of 2^K possible environments, where $\upsilon = (\upsilon_1, \ldots, \upsilon_K) \in \{-1, +1\}^K$, and the K first strata are defined by $\mu_k = \mu + \upsilon_k\frac{\mu}{2}$. Write $\mathbb{P}_\sigma$ the probability under such an environment, also consider $\mathbb{P}_\sigma$ the probability under which all the K first strata are Bernoulli with mean μ.

We define Ω_υ the event on which there are less than $\frac{K}{3}$ arms not pulled correctly for environment υ (i.e. for which $T_{k,n}$ is larger than the optimal allocation corresponding to μ when actually $\mu_k = \frac{\mu}{2}$, or smaller than the optimal allocation corresponding to μ when $\mu_k = 3\frac{\mu}{2}$). See the Appendix D for a precise definition of these events. Then, the idea is that there are so many such environments that any algorithm will be such that for at least one of them we have $\mathbb{P}_\sigma(\Omega_\upsilon) \leq \exp(-K/72)$. Then we derive by a variant of Pinsker's inequality applied to an event of small probability that $\mathbb{P}_\upsilon(\Omega_\upsilon) \leq \frac{KL(\mathbb{P}_\sigma,\mathbb{P}_\upsilon)}{K} = O(\frac{\sigma^{3/2}n}{K})$. Finally, by choosing σ of order $(\frac{K}{n})^{1/3}$, we have that $\mathbb{P}_\upsilon(\Omega_\upsilon^c)$ is bigger than a constant, and on Ω_υ^c we know that there are more than $\frac{K}{3}$ arms not pulled correctly. This leads to an expected pseudo-regret in environment υ of order $\Omega(\frac{K^{1/3}}{n^{4/3}})$. ■

This is the first lower-bound for the problem of online stratified sampling for Monte-Carlo. Note that this bound is of same order as the upper bound for the pseudo-regret of

algorithm MC-UCB. It means that this algorithm is, up to a constant, minimax optimal, both in terms of the number of samples and in terms of the number of strata. It however holds only on the partitions $\mathcal{N}_K$ (we conjecture that a similar result holds for *any* measurable partition $\mathcal{N}$, but with a bound of order $\Omega\Big(\sum_{x\in\mathcal{N}} \frac{w_x^{2/3}}{n^{4/3}}\Big)$).

5. Best trade-off between $Q_{n,\mathcal{N}_K}$ and $R_{n,\mathcal{N}_K}(\mathcal{A}_{MC-UCB})$

5.1. Best trade-off

We consider in this Section the hyper-cubic partitions $\mathcal{N}_K$ as defined in Definition 2, and we want to find the best number of strata K_n as a function of n. Using the results in Section 3 and Subsection 4.1, it is possible to deduce an optimal number of strata K to give as parameter to algorithm $MC-UCB$. Note that since the performance of the algorithm is defined as the sum of the quality of partition $\mathcal{N}_K$, i.e. $Q_{n,\mathcal{N}_K}$ and of the pseudo-regret of the algorithm MC-UCB, namely $R_{n,\mathcal{N}_K}(\mathcal{A}_{MC-UCB})$, one wants to (i) on the one hand take many strata so that $Q_{n,\mathcal{N}_K}$ is small but (ii) on the other hand, pay attention to the impact this number of strata has on the pseudo-regret $R_{n,\mathcal{N}_K}(\mathcal{A}_{MC-UCB})$. A good way to do that is to choose K_n in function of n such that $Q_{n,\mathcal{N}_{K_n}}$ and $R_{n,\mathcal{N}_{K_n}}(\mathcal{A}_{MC-UCB})$ are of the same order.

Theorem 6 *Under Assumptions 1 and 2 (since on $[0,1]^d$, Assumption 1 implies Assumption 3, by setting $f_{\max} = X(1) + Md^{\alpha/2}$), choosing $K_n = \Big(\lfloor (n^{\frac{d}{d+3\alpha}})^{1/d} \rfloor\Big)^d (\leq n^{\frac{d}{d+3\alpha}} \leq n)$, we have*

$$\mathbb{E}[L_n(\mathcal{A}_{MC-UCB})] - \frac{1}{n}\Big(\int_{[0,1]^d} s(x)dx\Big)^2 \leq Cd^{\frac{2\alpha}{3d}+\frac{1}{2}}\sqrt{\log(n)}n^{-\frac{d+4\alpha}{d+3\alpha}}(1+d^\alpha n^{-\frac{\alpha}{d+3\alpha}}),$$

where $c = 70(1+M)\Sigma_{\mathcal{N}_K}\sqrt{(1+3b+4(f(0)+s(0)+M)^2)}\Big(\frac{(f(0)+s(0)+M)+4}{4}\Big)^{1/3}$.
This leads to, if $d \ll n$, the simplified bound is

$$\mathbb{E}[L_n(\mathcal{A}_{MC-UCB})] - \frac{1}{n}\Big(\int_{[0,1]^d} s(x)dx\Big)^2 = \tilde{O}(n^{-\frac{d+4\alpha}{d+3\alpha}}).$$

Proof [Proof of Theorem 6] The definition of K_n implies that $K_n \geq \Big((n^{\frac{d}{d+3\alpha}} - 1)^{1/d}\Big)^d \geq n^{\frac{d}{d+3\alpha}}\Big(1 - \frac{d}{n^{\frac{1}{d}(\frac{d}{d+3\alpha})}}\Big)$. Also, trivially, $K_n \leq n^{\frac{d}{d+3\alpha}}$. By plugging these lower and upper bounds, in respectively $Q_{n,\mathcal{N}_{K_n}}$ and $R_{n,\mathcal{N}_{K_n}}$, we obtain the the final bound. ■

We can also prove a matching minimax lower bound using the results in Theorem 5.

Theorem 7 *Let* sup *represent the supremum taken over all $\alpha-$Hölder functions and* inf *be the infimum taken over all algorithms that partition the space in convex strata of same shape, then the following holds true:*

$$\inf\sup \mathbb{E}L_n(\mathcal{A}) - \frac{1}{n}\Big(\int_{[0,1]^d} s(x)dx\Big)^2 = \Omega(n^{-\frac{d+4\alpha}{d+3\alpha}}).$$

Proof [Proof of Theorem 7] This is a direct consequence of Theorem 5 and the second comment of Subsection 3.1. ■

5.2. Discussion

Optimal pseudo-risk. The dominant term in the pseudo-risk of MC-UCB with proper number of strata is $\frac{(\inf_{\mathcal{N}} \Sigma_{\mathcal{N}})^2}{n} = \frac{1}{n}\big(\int_{[0,1]^d} s(x)dx\big)^2$ (the other term is negligible). This means that algorithm MC-UCB is almost as efficient as the optimal oracle strategy on the optimal oracle partition. In comparison, the variance of the estimate given by crude Monte-Carlo is $\int_{[0,1]^d}\big(f(x) - \int_{[0,1]^d} f(u)du\big)^2 dx + \int_{[0,1]^d} s(x)^2 dx$. Thus MC-UCB enables to have the term coming from the variations in the mean vanish, and the noise term decreases (since by Cauchy-Schwarz, $\big(\int_{[0,1]^d} s(x)dx\big)^2 \leq \int_{[0,1]^d} s(x)^2 dx$).

minimax-optimal trade-off for algorithm MC-UCB. The optimal trade-off on the number of strata K_n of order $n^{\frac{d}{d+3\alpha}}$ depends on the dimension and the smoothness of the function. The higher the dimension, the more strata are needed in order to have a decent speed of convergence for $\Sigma_{\mathcal{N}_K}$. The smoother the function, the less strata are needed.
It is yet important to remark that this trade-off is not exact. We provide an almost minimax-optimal order of magnitude for K_n, in terms of n, so that the rate of convergence of the algorithm is minimax-optimal up to a $\sqrt{\log(n)}$.

Link between risk and pseudo-risk. It is important to compare the pseudo-risk $L_n(\mathcal{A}) = \sum_{k=1}^K \frac{w_k^2\sigma_k^2}{T_{k,n}}$ and the true risk $\mathbb{E}[(\hat{\mu}_n - \mu)^2]$. Note that those quantities are in general not equal for an algorithm $\mathcal{A}$ that allocates the samples in a dynamic way: indeed, the quantities $T_{k,n}$ are in that case stopping times and the variance of estimate $\hat{\mu}_n$ is not equal to the pseudo-risk. However, in the paper (Carpentier and Munos, 2011b), the authors highlighted for $MC-UCB$ some links between the risk and the pseudo-risk. More precisely, they established links between $L_n(\mathcal{A})$ and $\sum_{k=1}^K w_k^2 \mathbb{E}[(\hat{\mu}_{k,n} - \mu_k)^2]$. This step is possible since $\mathbb{E}[(\hat{\mu}_{k,n} - \mu_k)^2] \leq \frac{w_k^2\sigma_k^2}{\underline{T}_{k,n}^2}\mathbb{E}[T_{k,n}]$, where $\underline{T}_{k,n}$ is a lower-bound on the number of pulls $T_{k,n}$ on a high probability event. Then they bounded the cross products $\mathbb{E}[(\mu_{k,n} - \mu_k)(\hat{\mu}_{p,n} - \mu_p)]$ and provided some upper bounds on those terms. A tight analysis of these terms as a function of the number of strata K remains to be investigated.

Knowledge of the Hölder exponent. In order to be able to choose properly the number of strata to achieve the rate in Theorem 6, it is needed to possess a proper lower bound on the Hölder exponent of the function: indeed, the rougher the function is, the more strata are required. On the other hand, such a knowledge on the function is not always available and an interesting question is whether it is possible to estimate this exponent fast enough. There are interesting papers on that subject like (Hoffmann and Lepski, 2002) where the authors tackle the problem of regression and prove that it is possible, up to a certain extent, to adapt to the unknown smoothness of the function. The authors in (Giné and Nickl, 2010) add to that (in the case of density estimation) and prove that it is even possible under the assumption that the function attain its Hölder exponent to have a proper estimation of

this exponent and thus adaptive confidence bands. An idea would be to try to adapt those results in the case of finite sample.

MC-UCB On a noiseless function. Consider the case where $s = 0$ almost surely, i.e. the samples collected are noiseless. Proposition 1 ensures that $\inf_{\mathcal{N}} \Sigma_{\mathcal{N}} = 0$: it is thus possible in this case to achieve a pseudo-risk that has a faster rate than $O(\frac{1}{n})$. If the function m is smooth, e.g. Hölder with a not too low exponent α, it is efficient to use low discrepancy methods to integrate the functions. An idea is to stratify the domain in n hyper-rectangular strata of minimal diameter, and to pick at random one sample per stratum. The variance of the resulting estimate is of order $O(\frac{1}{n^{1+2\alpha/d}})$. Algorithm MC-UCB is not as efficient as a low discrepancy schemes: it needs a number of strata $K < n$ in order to be able to estimate the variance of each stratum. Its pseudo-risk is then of order $O(\frac{1}{nK^{2\alpha/d}})$.
It is however only true when the observations are noiseless. Otherwise, the order for the variance of the estimate is in $1/n$, no matter what strategy the learner chooses.

In high dimension. The first bound in Theorem 6 expresses precisely how the performance of the estimate outputted by MC-UCB depends on d. The first bound states that the quantity $L_n(\mathcal{A}) - \frac{1}{n}\Big(\int_{[0,1]^d} s(x)dx\Big)^2$ is negligible when compared to $1/n$ when n is exponential in d. This is not surprising since our technique aims at stratifying equally in every direction. It is not possible to stratify in every directions of the domain if the function lies in a very high dimensional domain.
This is however *not* a reason for not using our algorithm in high dimension. Indeed, stratifying even in a small number of strata already reduces the variance, and in high dimension, any variance reduction techniques are welcome. As mentioned in the end of Section 2, the model that we propose for the function is suitable for modeling d^* dimensional functions that we only stratify in $d < d^*$ directions (and $d \ll n$). A reasonable trade-off for d can also be inferred from the bound, but we believe that what a good choice of d is depends a lot of the problem. We then believe that it is a good idea to select the number of strata in the minimax way that we propose. Again, having a very high dimensional function that one stratifies in only a few directions is a very common technique in financial mathematics, for pricing options (practitioners stratify an infinite dimensional process in only 1 to 5 carefully chosen dimensions).

6. Numerical experiment: influence of the number of strata in the Pricing of an Asian option

We consider the pricing problem of an Asian option introduced in (Glasserman et al., 1999) and later considered in (Kawai, 2010; Etoré and Jourdain, 2010). This uses a Black-Scholes model with strike C and maturity T. Let $(W(t))_{0\le t\le T}$ be a Brownian motion. The discounted payoff of the Asian option is defined as a function of W, by:

$$F((W)_{0\le t\le T}) = \exp(-rT)\max\Big[\int_0^T S_0 \exp\Big((r-\tfrac{1}{2}s_0^2)t + s_0 W_t\Big)dt - C, 0\Big], \qquad (11)$$

where S_0, r, and s_0 are constants, and the price is defined by the expectation $p = \mathbb{E}_W F(W)$.

We want to estimate the price p by Monte-Carlo simulations (by sampling on W). In order to reduce the variance of the estimated price, we can stratify the space of W.

Glasserman et al. (1999) suggest to stratify according to a one dimensional projection of W, i.e., by choosing a time t and stratifying according to the quantiles of W_t (and simulating the rest of the Brownian according to a Brownian Bridge, see (Kawai, 2010)). They further argue that the best direction for stratification is to choose $t = T$, i.e., to stratify according to the last time of T. This choice of stratification is also intuitive since W_T has the highest variance, the biggest exponent in the payoff (11), and thus the highest volatility. Kawai (2010) and Etoré and Jourdain (2010) also use the same direction of stratification. We stratify according to the quantiles of W_T, that is to say the quantiles of a normal distribution $\mathcal{N}(0, T)$. When stratifying in K strata, we stratify according to the $1/K$-th quantiles (so that the strata are hyper-cubes of same measure).

We choose the same numerical values as Kawai (2010): $S_0 = 100$, $r = 0.05$, $s_0 = 0.30$, $T = 1$ and $d = 16$. We discretize also, as in Kawai (2010), the Brownian motion in 16 equidistant times, so that we are able to simulate it. We choose $C = 120$.

In this paper, we only do experiments for MC-UCB, and exhibit the influence of the number of strata. For a comparison between MC-UCB and other algorithms, see (Carpentier and Munos, 2011a). By studying the range of the $F(W)$, we set the parameter of the algorithm MC-UCB to $A = 150 \log(n)$.

For $n = 200$ and $n = 2000$, we observe the influence of the number of strata in Figure 2. We observe the trade-off that we mentioned between pseudo-regret and quality, in the sense that the mean squared error of the estimate outputted by MC-UCB (when compared to the true integral of f) first decreases with K and then increases. Note that, without surprise, for a large n the minimum of mean squared error is reached with more strata. Finally, note that our technique is never outperformed by uniform stratified Monte-Carlo: it is a good idea to try to adapt.

7. Conclusion

In this paper we studied the problem of online stratified sampling for the numerical integration of a function given noisy evaluations, and more precisely we discussed the problem of choosing the *minimax-optimal number* of strata.

We explained why, to our minds, this is a crucial problem when one wants to design an efficient algorithm. We enlightened the fact that there is a trade-off between having many strata (and a good approximation error, called the quality of a partition), and not too many, in order to perform almost as well as the optimal oracle allocation on a given partition (small estimation error, called pseudo-regret).

When the function is noisy, the noise is the dominant quantity in the optimal oracle variance on the optimal oracle partition. Indeed, decreasing the size of the strata does not diminish the (local) variance of the noise. In this case, the pseudo-risk of algorithm MC-UCB is equal, up to negligible terms, to the mean squared error of the estimate outputted by the optimal oracle strategy on the best (oracle) partition, at a rate of $O(n^{-\frac{d+4\alpha}{d+3\alpha}})$ where α is the Hölder exponent of s and m. This rate is minimax optimal on the class of α-Hölder functions: it is not possible, up to a constant factor, to do better on simultaneously all α-Hölder functions.

We believe that there are (at least) three very interesting remaining open questions:

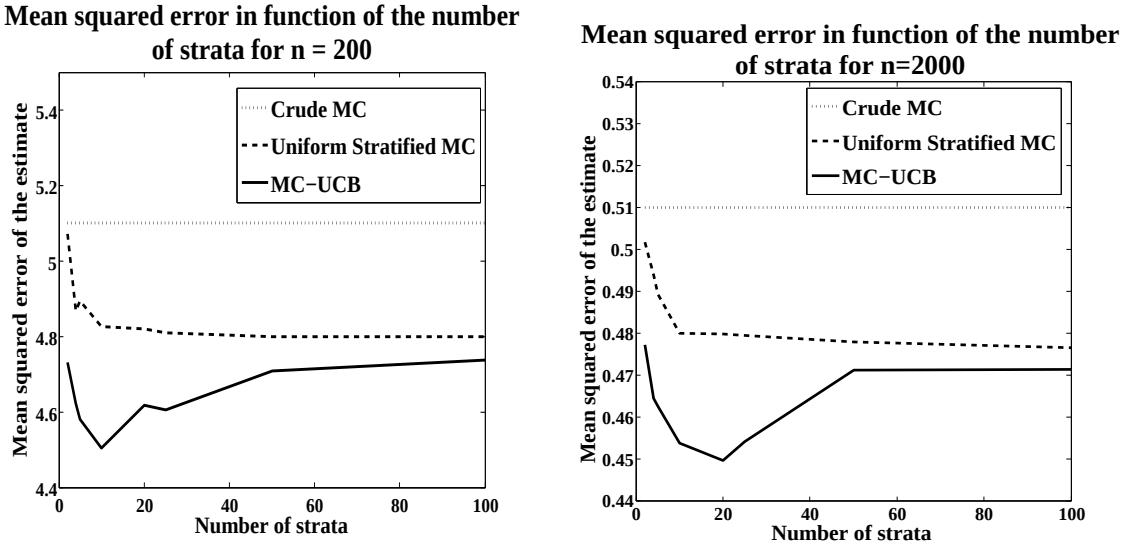


Figure 2: Mean squared error for uniform stratified sampling for different number of strata, for (Left:) n=200 and (Right:) n=2000.

- The first one is to investigate whether it is possible to estimate online the Hölder exponent *fast enough*. Indeed, one needs it in order to compute the proper number of strata for MC-UCB, and the lower bound on the Hölder exponent appears in the bound. It is thus a crucial parameter.

- The second direction is to build a more efficient algorithm in the noiseless case. We remarked that MC-UCB is not as efficient in this case as a simple non-adaptive method. The problem comes from the fact that in the case of a noiseless function, it is important to sample the space in a way that ensures that the points are as spread as possible. An interesting problem is thus to build an algorithm that mixes ideas from quasi Monte-Carlo and ideas from online stratified Monte-Carlo.

- Another question is the relevance of fixing the strata in advance. Although it is minimax-optimal on the class of $\alpha-$Hölder functions to have hyper-cubic strata of same measure, it might in some cases be more interesting to focus and stratify more finely at places where the function is rough. On that perspective, it could be more clever to have an adaptive procedure that also decides *where* to refine the strata.

References

A. Carpentier and R. Munos. Finite-time analysis of stratified sampling for monte carlo. In *In Neural Information Processing Systems (NIPS)*, 2011a.

A. Carpentier and R. Munos. Finite-time analysis of stratified sampling for monte carlo. Technical report, INRIA-00636924, 2011b.

Pierre Etoré and Benjamin Jourdain. Adaptive optimal allocation in stratified sampling methods. *Methodol. Comput. Appl. Probab.*, 12(3):335–360, September 2010.

Pierre Etoré, Gersende Fort, Benjamin Jourdain, and Éric Moulines. On adaptive stratification. *Ann. Oper. Res.*, 2011. to appear.

E. Giné and R. Nickl. Confidence bands in density estimation. *The Annals of Statistics*, 38 (2):1122–1170, 2010.

P. Glasserman. *Monte Carlo methods in financial engineering*. Springer Verlag, 2004. ISBN 0387004513.

P. Glasserman, P. Heidelberger, and P. Shahabuddin. Asymptotically optimal importance sampling and stratification for pricing path-dependent options. *Mathematical Finance*, 9 (2):117–152, 1999.

V. Grover. Active learning and its application to heteroscedastic problems. *Department of Computing Science, Univ. of Alberta, MSc thesis*, 2009.

M. Hoffmann and O. Lepski. Random rates in anisotropic regression. *Annals of statistics*, pages 325–358, 2002.

R. Kawai. Asymptotically optimal allocation of stratified sampling with adaptive variance reduction by strata. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 20(2):1–17, 2010. ISSN 1049-3301.

R.Y. Rubinstein and D.P. Kroese. *Simulation and the Monte Carlo method*. Wiley-interscience, 2008. ISBN 0470177942.

Appendix A. Proof of Theorem 10

A.1. The main tool: a high probability bound on the standard deviations

Upper bound on the standard deviation:

Lemma 8 *Let Assumption 2 hold and $n \geq 2$. Define the following event*

$$\xi = \xi_{K,n}(\delta) = \bigcap_{1\leq k\leq K,\, 2\leq t\leq n} \left\{ \left| \sqrt{\frac{1}{t-1}\sum_{i=1}^{t}\left(X_{k,i} - \frac{1}{t}\sum_{j=1}^{t} X_{k,j}\right)^2} - \sigma_k \right| \leq A\sqrt{\frac{1}{t}} \right\}, \tag{12}$$

where $A = 2\sqrt{(1+3b+4\bar{V})\log(2nK/\delta)}$. Then $\Pr(\xi) \geq 1-\delta$.

Note that the first term in the absolute value in Equation 12 is the empirical standard deviation of arm k computed as in Equation 9 for t samples. The event ξ plays an important role in the proofs of this section and a number of statements will be proved on this event.

Proof Under Assumption 2 we have for $f_{\max}^2 \geq \max_k \sigma_k^2$ with probability $1-\delta$ because of the results of Lemma 15

$$\left| \sqrt{\frac{1}{t-1}\sum_{i=1}^{t}\left(X_{k,i} - \frac{1}{t}\sum_{j=1}^{t} X_{k,j}\right)^2} - \sigma_k \right| \leq 2\sqrt{\frac{(1+3b+4f_{\max}^2)\log(2/\delta)}{t}}. \tag{13}$$

Then by doing a simple union bound on (k,t), we obtain the result. ∎

We deduce the following corollary when the number of samples $T_{k,t}$ are random.

Corollary 9 *For any $k = 1,\dots,K$ and $t = 2K,\dots,n$, let $\{X_{k,i}\}_i$ be n i.i.d. random variables drawn from ν_k, satisfying Assumption 2. Let $T_{k,t}$ be any random variable taking values in $\{2,\dots,n\}$. Let $\hat{\sigma}_{k,t}^2$ be the empirical variance computed from Equation 9. Then, on the event ξ, we have:*

$$|\hat{\sigma}_{k,t} - \sigma_k| \leq A\sqrt{\frac{1}{T_{k,t}}}\,, \tag{14}$$

where $A = 2\sqrt{(1+3b+4\bar{V})\log(2nK/\delta)}$.

A.2. Main Demonstration

We first state and prove the following Lemma and then use this result to prove Theorem 10.

Theorem 10 *Let Assumption 2 hold. For any $0 < \delta \leq 1$ and for $n \geq 4K$, the algorithm MC-UCB launched on a partition $\mathcal{N}_K$ satisfies*

$$\mathbb{E}L_n \leq \frac{\Sigma_{\mathcal{N}_K}^2}{n} + 24\sqrt{2}\Sigma_{\mathcal{N}_K}\sqrt{(1+3b+4f_{\max}^2)}\left(\frac{f_{\max}+4}{4}\right)^{1/3}\frac{K^{1/3}}{n^{4/3}}\sqrt{\log(nK)} + \frac{14K\Sigma_{\mathcal{N}_K}^2}{n^2}.$$

Proof

Step 1. Lower bound of order $\widetilde{O}(n^{2/3})$. Let k be the index of an arm such that $T_{k,n} \geq \frac{n}{K}$ (this implies $T_{k,n} \geq 3$ as $n \geq 4K$, and arm k is thus pulled after the initialization) and let $t+1 \leq n$ be the last time at which it was pulled [6], i.e., $T_{k,t} = T_{k,n} - 1$ and $T_{k,t+1} = T_{k,n}$. From Equation 14 and the fact that $T_{k,n} \geq \frac{n}{K}$, we obtain on ξ

$$B_{k,t} \leq \frac{w_k}{T_{k,t}} \left(\sigma_k + 2A \sqrt{\frac{1}{T_{k,t}}} \right) \leq \frac{K w_k \Big(\sigma_k + 2A \Big)}{n}, \tag{15}$$

where the second inequality follows from the facts that $T_{k,t} \geq 1$, $w_k \sigma_k \leq \Sigma_{\mathcal{N}_K}$, and $w_k \leq \sum_k w_k = 1$. Since at time $t+1$ the arm k has been pulled, then for any arm q, we have

$$B_{q,t} \leq B_{k,t}. \tag{16}$$

From the definition of $B_{q,t}$, and also using the fact that $T_{q,t} \leq T_{q,n}$, we deduce on ξ that

$$B_{q,t} \geq \frac{2A w_q}{T_{q,t}^{3/2}} \geq \frac{2A w_q}{T_{q,n}^{3/2}} \ . \tag{17}$$

Combining Equations 15–17, we obtain on ξ

$$\frac{2A w_q}{T_{q,n}^{3/2}} \leq \frac{K w_k \Big(\sigma_k + 2A \Big)}{n}.$$

Finally, this implies on ξ that for any q because $w_k = w_q$,

$$T_{q,n} \geq \Big(\frac{2A}{\sigma_k + 2A} \frac{n}{K} \Big)^{2/3}. \tag{18}$$

This implies that $\forall q, T_{q,n} \geq C \Big(\frac{n}{K} \Big)^{2/3}$ where $C = \Big(\frac{2A}{\max_k \sigma_k + 2A} \Big)^{2/3}$.

Step 2. Properties of the algorithm. We first remind the definition of $B_{q,t+1}$ used in the MC-UCB algorithm

$$B_{q,t+1} = \frac{w_q}{T_{q,t}} \left(\hat{\sigma}_{q,t} + A \sqrt{\frac{1}{T_{q,t}}} \right).$$

Using Corollary 9 it follows that, on ξ

$$\frac{w_q \sigma_q}{T_{q,t}} \leq B_{q,t+1} \leq \frac{w_q}{T_{q,t}} \left(\sigma_q + 2A \sqrt{\frac{1}{T_{q,t}}} \right). \tag{19}$$

Let $t+1 \geq 2K+1$ be the time at which an arm q is pulled for the last time, that is $T_{q,t} = T_{q,n} - 1$. Note that there is at least one arm such that this happens as $n \geq 4K$. Since at $t+1$ arm q is chosen, then for any other arm p, we have

$$B_{p,t+1} \leq B_{q,t+1} \ . \tag{20}$$

6. Note that such an arm always exists for any possible allocation strategy given the constraint $n = \sum_q T_{q,n}$.

From Equation 19 and $T_{q,t} = T_{q,n} - 1$, we obtain on ξ

$$B_{q,t+1} \leq \frac{w_q}{T_{q,t}}\left(\sigma_q + 2A\sqrt{\frac{1}{T_{q,t}}}\right) = \frac{w_q}{T_{q,n}-1}\left(\sigma_q + 2A\sqrt{\frac{1}{T_{q,n}-1}}\right). \tag{21}$$

Furthermore, since $T_{p,t} \leq T_{p,n}$, then on ξ

$$B_{p,t+1} \geq \frac{w_p\sigma_p}{T_{p,t}} \geq \frac{w_p\sigma_p}{T_{p,n}}. \tag{22}$$

Combining Equations 20–22, we obtain on ξ

$$\frac{w_p\sigma_p}{T_{p,n}}(T_{q,n}-1) \leq w_q\left(\sigma_q + 2A\sqrt{\frac{1}{T_{q,n}-1}}\right).$$

Summing over all q such that the previous Equation is verified, i.e. such that $T_{q,n} \geq 3$, on both sides, we obtain on ξ

$$\frac{w_p\sigma_p}{T_{p,n}}\sum_{q|T_{q,n}\geq 3}(T_{q,n}-1) \leq \sum_{q|T_{q,n}\geq 3} w_q\left(\sigma_q + 2A\sqrt{\frac{1}{T_{q,n}-1}}\right).$$

This implies

$$\frac{w_p\sigma_p}{T_{p,n}}(n-3K) \leq \sum_{q=1}^{K} w_q\left(\sigma_q + 2A\sqrt{\frac{1}{T_{q,n}-1}}\right). \tag{23}$$

Step 3. Lower bound. Plugging Equation 18 in Equation 23,

$$\begin{aligned}\frac{w_p\sigma_p}{T_{p,n}}(n-3K) &\leq \sum_q w_q\left(\sigma_q + 2A\sqrt{\frac{1}{T_{q,n}-1}}\right)\\ &\leq \sum_q w_q\left(\sigma_q + 2A\sqrt{\frac{2K^{2/3}}{Cn^{2/3}}}\right)\\ &\leq \Sigma_{\mathcal{N}_K} + \frac{2\sqrt{2}A}{\sqrt{C}}\frac{K^{1/3}}{n^{1/3}},\end{aligned}$$

on ξ, since $T_{q,n} - 1 \geq \frac{T_{q,n}}{2}$ (as $T_{q,n} \geq 2$). Finally as $n \geq 4K$, we obtain on ξ the following bound

$$\frac{w_p\sigma_p}{T_{p,n}} \leq \frac{\Sigma_{\mathcal{N}_K}}{n} + \frac{4\sqrt{2}A}{\sqrt{C}}\frac{K^{1/3}}{n^{4/3}} + \frac{12K\Sigma_{\mathcal{N}_K}}{n^2}. \tag{24}$$

Step 4. Regret. By summing and using Equation 24 which holds for all p, we obtain on ξ (with probability $1-\delta$)

$$L_n = \sum_p \frac{w_p^2\sigma_p^2}{T_{p,n}} \leq \frac{\Sigma^2_{\mathcal{N}_K}}{n} + \frac{4\Sigma_{\mathcal{N}_K}\sqrt{2}A}{\sqrt{C}}\frac{K^{1/3}}{n^{4/3}} + \frac{12K\Sigma^2_{\mathcal{N}_K}}{n^2}.$$

This implies since $\mathbb{E}L_n = \mathbb{E}[L_n \mathbb{I}\{\xi\}] + \mathbb{E}[L_n \mathbb{I}\{\xi^c\}]$ and since $\delta = n^{-2}$

$$\begin{aligned}\mathbb{E}L_n &\leq \frac{\Sigma^2_{\mathcal{N}_K}}{n} + \frac{4\Sigma_{\mathcal{N}_K}\sqrt{2}A}{\sqrt{C}}\frac{K^{1/3}}{n^{4/3}} + \frac{12K\Sigma^2_{\mathcal{N}_K}}{n^2} + (\sum_p w_p^2\sigma_p^2)n^{-2} \\ &\leq \frac{\Sigma^2_{\mathcal{N}_K}}{n} + \frac{4\Sigma_{\mathcal{N}_K}\sqrt{2}A}{\sqrt{C}}\frac{K^{1/3}}{n^{4/3}} + \frac{14K\Sigma^2_{\mathcal{N}_K}}{n^2}.\end{aligned}$$

Since $\delta = n^{-2}$, we have $A \leq 6\sqrt{(1+3b+4\bar{V})\log(nK)}$ and $C \geq \left(\frac{4}{f_{\max}+4}\right)^{2/3}$, this leads to

$$\mathbb{E}L_n \leq \frac{\Sigma^2_{\mathcal{N}_K}}{n} + 24\sqrt{2}\Sigma_{\mathcal{N}_K}\sqrt{(1+3b+4f^2_{\max})}\left(\frac{f_{\max}+4}{4}\right)^{1/3}\frac{K^{1/3}}{n^{4/3}}\sqrt{\log(nK_n)} + \frac{14K\Sigma^2_{\mathcal{N}_K}}{n^2}.$$

■

Appendix B. Proof of Proposition 1

Step 1: Expression of the variance of the stratified estimate. Note that the samples $f(x) + s(x)\epsilon_t$ where $\epsilon_t \sim \nu_x$ and $\mathbb{E}_{\nu_x}[\epsilon_t] = 0$, $\mathbb{V}_{\nu_x}[\epsilon_t] = 1$ the ϵ_t are independent. We have

$$\begin{aligned}\sigma_k^2 &= \frac{1}{w_k}\int_{\Omega_k} \mathbb{E}_{\nu_x}[(X_x(t) - \mu_k)^2]dx \\ &= \frac{1}{w_k}\int_{\Omega_k} \mathbb{E}_{\nu_x}\Big[(f(x) + s(x)\epsilon_t - \frac{1}{w_k}\int_{\Omega_k} f(u)du)^2\Big]dx \\ &= \frac{1}{w_k}\int_{\Omega_k} \mathbb{E}_{\nu_x}\Big[(f(x) - \frac{1}{w_k}\int_{\Omega_k} f(u)du)^2\Big]dx + \frac{1}{w_k}\int_{\Omega_k} \mathbb{E}_{\nu_x}\Big[s(x)^2\epsilon_t^2\Big]dx \\ &= \frac{1}{w_k}\int_{\Omega_k} \big(f(x) - \frac{1}{w_k}\int_{\Omega_k} f(u)du\big)^2 dx + \frac{1}{w_k}\int_{\Omega_k} s(x)^2 dx\end{aligned}$$

Step 2: Proof for the uniformly continuous functions. We first prove the result for a subset of $L_2([0,1]^d)$, namely the set of functions m and s that are uniformly continuous.

Proposition 11 *If the functions f and s are uniformly continuous and if the strata satisfy the Assumptions of Proposition 1, we have*

$$\sum_k w_{k,n}\sigma_{k,n} - \int_{[0,1]^d} s(x)dx \to 0$$

Proof

Let $\upsilon > 0$. As s and f are uniformly continuous, we know that $\forall x$, $\exists \eta$ such that $|s(x+u) - s(x)| \leq \upsilon$ and $|f(x+u) - f(x)| \leq \upsilon$ where $u \in \mathcal{B}_{2,d}(\eta)$[7].
By Assumption AS1, we know that $w_{k,n} \leq \upsilon_n$. Note that the diameter of strata $\Omega_{k,n}$ is

7. We denote by $B_{2,d}(\eta)$ the ball of center 0 and radius η according to the $||.||_2$ norm.

smaller than $D(w_{k,n}) \leq D(\upsilon_n)$. Let us choose n big enough, i.e. such that $D(\upsilon_n) \leq \eta$ and $\upsilon_n \leq \upsilon$.
We have

$$\begin{aligned}\sigma_{k,n}^2 - (\frac{1}{w_{k,n}}\int_{\Omega_{k,n}} s)^2 &= \frac{1}{w_{k,n}}\int_{\Omega_{k,n}} s^2 - \Big(\frac{1}{w_{k,n}}\int_{\Omega_{k,n}} s\Big)^2 + \frac{1}{w_{k,n}}\int_{\Omega_{k,n}} \Big(f - \frac{1}{w_{k,n}}\int_{\Omega_{k,n}} f\Big)^2 \\ &= \frac{1}{w_{k,n}}\int_{\Omega_{k,n}} \Big(s - \frac{1}{w_{k,n}}\int_{\Omega_{k,n}} s\Big)^2 + \frac{1}{w_{k,n}}\int_{\Omega_{k,n}} \Big(f - \frac{1}{w_{k,n}}\int_{\Omega_{k,n}} f\Big)^2 \\ &\leq \upsilon^2 + \upsilon^2 \leq 2\upsilon^2.\end{aligned}$$

Because of concavity of the square-root function, we get

$$\sigma_{k,n} - (\frac{1}{w_{k,n}}\int_{\Omega_{k,n}} s) \leq \sqrt{2}\upsilon.$$

By summing we get

$$\sum_k w_{k,n}\sigma_{k,n} - \int_{[0,1]^d} s \leq \sqrt{2}\upsilon.$$

■

Step 3: Density of uniformly continuous functions in $L_2([0,1]^d)$. We first remind a property of the functions in $L_2([0,1]^d)$.

Proposition 12 *The uniformly continuous functions according to the $||.||_2$ norm are dense in $L_2([0,1]^d)$.*

Proof The result follows directly from the facts that

- The continuous functions are dense in $L_2(\Omega)$ (Stone-Weierstrass Theorem).
- The uniformly continuous functions on a compact space Ω according to the $||.||_2$ norm are dense in the space of continuous functions.
- $[0,1]^d$ is a compact.

■

This means that we can approximate with arbitrary precision according to the $||.||_2$ measure on $L_2([0,1]^d)$ any function in $L_2([0,1]^d)$ by an uniformly continuous function.
Using this proposition, we can prove the following Lemma.

Lemma 13 *For a given n and a given υ, there exist two uniformly continuous function m_υ and s_υ such that:*

$$\Big|\sum_{k=1}^{K_n} w_{k,n}\sigma_{k,n} - \sum_{k=1}^{K_n} \sqrt{w_{k,n}}\sqrt{\int_{\Omega_{k,n}} \Big(f_\upsilon(x) + \int_{\Omega_{k,n}} f_\upsilon(u)du\Big)^2 dx - \frac{1}{w_{k,n}}\int_{\Omega_{k,n}} s_\upsilon^2(x)dx}\Big| \leq \upsilon.$$

Proof Let us fix n and υ.
Let m_υ be an uniformly continuous function such that

$$\int_\Omega (f(x) - f_\upsilon(x))^2 dx \leq \min_k(w_{k,n})\frac{\upsilon}{2},$$

and s_υ be an uniformly continuous function such that

$$\int_\Omega (s(x) - s_\upsilon(x))^2 dx \leq \min_k(w_{k,n})\frac{\upsilon}{2}.$$

It is possible because of $w_{k,n} > 0$ and because the uniformly continuous functions are dense in $L_2([0,1]^d)$ by Proposition 12.
Note that we thus have

$$\frac{1}{w_{k,n}}\int_{\Omega_{k,n}} (f(x) - f_\upsilon(x))^2 dx \leq \frac{\upsilon}{2},$$

and

$$\frac{1}{w_{k,n}}\int_{\Omega_{k,n}} (s(x) - s_\upsilon(x))^2 dx \leq \frac{\upsilon}{2}.$$

Note also that $\frac{1}{w_{k,n}}\int_{\Omega_{k,n}}(s(x) - s_\upsilon(x))^2 dx \geq \left|\frac{1}{w_{k,n}}\int_{\Omega_{k,n}} s(x)^2 dx - \frac{1}{w_{k,n}}\int_{\Omega_{k,n}} s_\upsilon(x)^2 dx\right|$.
Simple triangle inequality leads to

$$\left|\frac{1}{w_{k,n}}\int_{\Omega_{k,n}}(f(x) - \frac{1}{w_{k,n}}\int_{\Omega_{k,n}} f(u)du)^2 dx - \frac{1}{w_{k,n}}\int_{\Omega_{k,n}}(f_\upsilon(x) - \frac{1}{w_{k,n}}\int_{\Omega_{k,n}} f_\upsilon(u)du)^2 dx\right| \leq \frac{\upsilon}{2}.$$

Now note that as $\sigma_{k,n}^2 = \frac{1}{w_{k,n}}\int_{\Omega_{k,n}}(f(x) - \frac{1}{w_{k,n}}\int_{\Omega_{k,n}} f(u)du)^2 dx + \frac{1}{w_{k,n}}\int_{\Omega_{k,n}} s(x)^2 dx$, we know that the variance of the function on strata $\Omega_{k,n}$ is arbitrarily close to the variance of its approximation.
By convexity, one gets

$$\left|\sigma_{k,n} - \sqrt{\frac{1}{w_{k,n}}\int_{\Omega_{k,n}}\left(f_\upsilon(x) - \frac{1}{w_{k,n}}\int_{\Omega_{k,n}} f_\upsilon(u)du\right)^2 dx + \frac{1}{w_{k,n}}\int_{\Omega_{k,n}} s_\upsilon^2(x)dx}\right| \leq \upsilon.$$

And finally, by summing

$$\left|\sum_{k=1}^{K_n} w_{k,n}\sigma_{k,n} - \sum_{k=1}^{K_n}\sqrt{w_{k,n}}\sqrt{\int_{\Omega_{k,n}}\left(f_\upsilon(x) + \int_{\Omega_{k,n}} f_\upsilon(u)du\right)^2 dx - \frac{1}{w_{k,n}}\int_{\Omega_{k,n}} s_\upsilon^2(x)dx}\right| \leq \upsilon.$$

■

Step 4: Combination of all the preliminary results to finish the proof. Finally, we finish the demonstration of Proposition 1.

Let $\upsilon > 0$ and f_υ and s_υ be as in Lemma 13.
We know that

$$\left|\sum_{k=1}^{K_n} w_{k,n}\sigma_{k,n} - \sum_{k=1}^{K_n}\sqrt{w_{k,n}}\sqrt{\int_{\Omega_{k,n}}\left(f_\upsilon(x) + \int_{\Omega_{k,n}} f_\upsilon(u)du\right)^2 dx - \frac{1}{w_{k,n}}\int_{\Omega_{k,n}} s_\upsilon^2(x)dx}\right| \leq \upsilon,$$

and also that

$$\int_\Omega (s(x) - s_\upsilon(x))^2 dx \leq \min_k(w_{k,n})\frac{\upsilon}{2} \leq \frac{\upsilon}{2}.$$

Note that by Cauchy-Schwartz:

$$\int_\Omega |s(x) - s_\upsilon(x)| dx \leq \sqrt{\int_\Omega (s(x) - s_\upsilon(x))^2 dx} \leq \sqrt{\frac{\upsilon}{2}}.$$

Note also that Proposition 11 tells us that $\exists n$ such that

$$\sum_{k=1}^{K_n} \sqrt{w_{k,n}} \sqrt{\int_{\Omega_{k,n}} \Big(f_\upsilon(x) - \frac{1}{w_{k,n}} \int_{\Omega_{k,n}} f_\upsilon(u)du\Big)^2 dx + \int_{\Omega_{k,n}} s_\upsilon^2(x)dx} - \int_{[0,1]^d} s_\upsilon(x)dx \leq \upsilon.$$

When combining all those results, one gets the desired result.

Note finally that if we choose the strata as being small boxes of size $\frac{1}{K}$ and side $(\frac{1}{K})^{1/d}$, then the assumptions of Proposition 1 is verified.

Appendix C. Proof of Proposition 3

Note first that

$$\sigma_k^2 = \frac{1}{w_k}\int_{\Omega_k} \Big(f(x) - \frac{1}{w_k}\int_{\Omega_k} f(u)du\Big)^2 dx + \frac{1}{w_k}\int_{\Omega_k} s^2(x)dx.$$

The term in f As the function f is $(\alpha, M)-$ Hölder, we know that $\forall (x,y) \in \Omega, |f(x) - f(y)| \leq M||x-y||_2^\alpha$.
Using that we get

$$\begin{aligned}\frac{1}{w_k}\int_{\Omega_k} \Big(f(x) - \frac{1}{w_k}\int_{\Omega_k} f(u)du\Big)^2 dx &\leq M^2 D(\Omega_k)^{2\alpha} \\ &\leq M^2 d(\frac{1}{K})^{2\alpha/d}.\end{aligned}$$

The term in s As the function s is $(\alpha, M)-$ Hölder, we know that $\forall (x,y) \in \Omega, |s(x) - s(y)| \leq M||x-y||_2^\alpha$.

$$\begin{aligned}\frac{1}{w_k}\int_{\Omega_k} s^2(x)dx - (\frac{1}{w_k}\int_{\Omega_k} s(u)du)^2 = \frac{1}{w_k}\int_{\Omega_k} \big(s(x) - \frac{1}{w_k}\int_{\Omega_k} s(u)du\big)^2 dx &\leq M^2 D(\Omega_k)^{2\alpha} \\ &\leq M^2 d(\frac{1}{K})^{2\alpha/d}.\end{aligned}$$

Finally... By combining those two results

$$\begin{aligned}w_k\sigma_k - \int_{\Omega_k} s(x)dx &\leq w_k\sqrt{\sigma_k^2 - (\frac{1}{w_k}\int_{\Omega_k} s(x)dx)^2} \\ &\leq w_k\sqrt{M^2 d(\frac{1}{K})^{2\alpha/d} + M^2 d(\frac{1}{K})^{2\alpha/d}}.\end{aligned}$$

By summing over all the strata, one obtains

$$\Sigma_{\mathcal{N}_K} - \int_{[0,1]^d} s(x)dx \leq \sqrt{2d}M(\frac{1}{K})^{\alpha/d}.$$

Appendix D. Lower bound

Let us write the proof of the lower bound using the terminology of multi-armed bandits. Each arm k represents a stratum and the distribution associated to this arm is defined as the distribution of the noisy samples of the function collected when sampling uniformly on the strata.

Let us choose $\mu < 1/2$ and $\alpha = \frac{\mu}{2}$. Consider $2K$ Bernoulli bandits (i.e., $2K$ strata where the samples follow Bernoulli distributions) where the K first bandits have parameter $(\mu_k)_{1\leq k\leq K}$ and the K last ones have parameter $1/2$. The μ_k take values in $\{\mu-\alpha, \mu, \mu+\alpha\}$.

Define $\sigma^2 = \mu(1-\mu)$ the variance of a Bernoulli of parameter μ, and is such that $\sqrt{\frac{1}{2}\mu} \leq \sigma \leq \sqrt{\mu}$. We wite $\sigma_{-\alpha}$ and $\sigma_{+\alpha}$ the two other standard deviations, and notice that $\frac{1}{2}\sqrt{\mu} \leq \sigma_{-\alpha} \leq \sqrt{\mu}$, and $\sqrt{\frac{1}{2}\mu} \leq \sigma_{+\alpha} \leq \sqrt{\mu}$.

We consider the 2^K bandit environments $M(\upsilon)$ (characterized by $\upsilon = (\upsilon_k)_{1\leq k\leq K} \in \{-1,+1\}^K$) defined by $(\mu_k = \mu + \upsilon_k\alpha)_{1\leq k\leq K}$. We write $\mathbb{P}_\upsilon$ the probability with respect to the environment $M(\upsilon)$ at time n. We also write $M(\sigma)$ the environment defined by all K first arms having a parameter σ, and write $\mathbb{P}_\sigma$ the associated probability at time n.

The optimal oracle allocation for environment $M(\upsilon)$ is to play arm $k \leq K$, $t_k(\upsilon) = \frac{\sigma_{\upsilon_k\alpha}}{\sum_{i=1}^K \sigma_{\upsilon_i\alpha}+K/2} n$ times and arm $k > K$, $t_k(\upsilon) = \frac{1/2}{\sum_{i=1}^K \sigma_{\upsilon_i\alpha}+K/2} n$ times. The corresponding quadratic error of the resulting estimate is $l(\upsilon) = \frac{(\sum_{i=1}^K \sigma_{\upsilon_i\alpha}+K/2)^2}{(2K)^2 n}$. For the environment $M(\sigma)$, the optimal oracle allocation is to play arm $k \leq K$, $t(\sigma) = \frac{\sigma}{K\sigma+K/2} n$ times (and arm $k > K$, $t_2(\sigma) = \frac{1/2}{K\sigma+K/2} n$ times).

Consider deterministic algorithms first (extension to randomized algorithms will be discussed later). An algorithm is a set (for all $t = 1$ to $n-1$) of mappings from any sequence $(r_1, \ldots, r_t) \in \{0,1\}$ of t observed samples (where $r_s \in \{0,1\}$ is the sample observed at the s-th round) to the choice of an arm $I_{t+1} \in \{1, \ldots, 2K\}$. Write $T_k(r_1, \ldots, r_n)$ the (random variable) corresponding to the number of pulls of arm k up to time n. We thus have $n = \sum_{k=1}^{2K} T_k$.

Now, consider the set of algorithms that know that the K first arms have parameter $\mu_k \in \{\mu-\alpha, \mu, \mu+\alpha\}$, and that also know that the K last arms have their parameters in $\{1/4, 3/4\}$. Given this knowledge, an optimal algorithm will not pull any arm $k \leq K$ more than $\left(\frac{\sigma_{+\alpha}}{K\sigma_{-\alpha}+\sqrt{3}K/4}\right) n$ times. Indeed, the optimal oracle allocation in *all* such environments allocates less than $\left(\frac{\sigma_{+\alpha}}{K\sigma_{-\alpha}+\sqrt{3}K/4}\right) n$ samples to each arm $k \leq K$. In addition, since the samples of all arms are independent, a sample collected from arm k does not provide any information about the relative allocations among the other arms. Thus, once an arm has been pulled as many times as recommended by the optimal oracle strategy, there is no need to allocate more samples to that arm. Writing $\mathbb{A}$ the class of all algorithms that do not know the set of possible environments, $\mathbb{A}_\upsilon$ the class of algorithms that know the set of possible environments $M(\upsilon)$ and $\mathbb{A}_{opt}$ the subclass of $\mathbb{A}_\upsilon$ that pull all arms $k \leq K$ less than $\left(\frac{\sigma_{+\alpha}}{K\sigma_{-\alpha}+\sqrt{3}K/4}\right) n$ times, we have

$$\inf_{\mathbb{A}} \sup_{M(\upsilon)} \mathbb{E}R_n \geq \inf_{\mathbb{A}_\upsilon} \sup_{M(\upsilon)} \mathbb{E}R_n = \inf_{\mathbb{A}_{opt}} \sup_{M(\upsilon)} \mathbb{E}R_n,$$

where the first inequality comes from the fact that algorithms in $\mathbb{A}_\upsilon$ possess more information than those in $\mathbb{A}$, which they can use or not. Thus $\mathbb{A} \subset \mathbb{A}_\upsilon$.

Now for any $\upsilon = (\upsilon_1, \dots, \upsilon_K)$, define the events

$$\Omega_\upsilon = \{\omega : \forall \mathcal{U} \subset \{1, \dots, K\} : |\mathcal{U}| \leq \frac{K}{3} \text{ and } \forall k \in \mathcal{U}^c, \upsilon_k T_k \geq \upsilon_k t(\sigma)\}.$$

Note that by definition

$$\Omega_\upsilon = \bigcup_{p=1}^{\frac{K}{3}} \bigcup_{\mathcal{U} \subset \{1,\dots,K\}:|\mathcal{U}|=p} \left\{ \Big\{ \bigcap_{k \in \mathcal{U}} \{\upsilon_k T_k < \upsilon_k t(\sigma)\} \Big\} \bigcap \Big\{ \bigcap_{k \in \mathcal{U}^C} \{\upsilon_k T_k \geq \upsilon_k t(\sigma)\} \Big\} \right\}.$$

By the sub-additivity of the probabilities, we have

$$\mathbb{P}_\sigma(\Omega_\upsilon) \leq \sum_{p=1}^{\frac{K}{3}} \sum_{\mathcal{U} \subset \{1,\dots,K\}:|\mathcal{U}|=p} \mathbb{P}\left[\left\{ \Big\{ \bigcap_{k \in \mathcal{U}} \{\upsilon_k T_k < \upsilon_k t(\sigma)\} \Big\} \bigcap \Big\{ \bigcap_{k \in \mathcal{U}^C} \{\upsilon_k T_k \geq \upsilon_k t(\sigma)\} \Big\} \right\} \right].$$

The events $\left\{ \Big\{ \bigcap_{k \in \mathcal{U}} \{\upsilon_k T_k < \upsilon_k t(\sigma)\} \Big\} \bigcap \Big\{ \bigcap_{k \in \mathcal{U}^C} \{\upsilon_k T_k \geq \upsilon t(\sigma)\} \Big\} \right\}$ are disjoint for different υ, and form a partition of the space, thus $\sum_\upsilon \mathbb{P}_\sigma \left[\left\{ \Big\{ \bigcap_{k \in \mathcal{U}} \{\upsilon_k T_k < \upsilon_k t(\sigma)\} \Big\} \bigcap \Big\{ \bigcap_{k \in \mathcal{U}^C} \{\upsilon T_k \geq \upsilon_k t(\sigma)\} \Big\} \right\} \right] = 1$.

We deduce that

$$\begin{aligned}
\sum_\upsilon \mathbb{P}_\sigma(\Omega_\upsilon) &\leq \sum_\upsilon \sum_{p=1}^{\frac{K}{3}} \sum_{\mathcal{U} \subset \{1,\dots,K\}:|\mathcal{U}|=p} \mathbb{P}_\sigma \left[\left\{ \Big\{ \bigcap_{k \in \mathcal{U}} \{\upsilon T_k < \upsilon_k t(\sigma)\} \Big\} \bigcap \Big\{ \bigcap_{k \in \mathcal{U}^C} \{\upsilon_k T_k \geq \upsilon_k t(\sigma)\} \Big\} \right\} \right] \\
&= \sum_{p=1}^{\frac{K}{3}} \sum_{\mathcal{U} \subset \{1,\dots,K\}:|\mathcal{U}|=p} \sum_\upsilon \left[\left\{ \Big\{ \bigcap_{k \in \mathcal{U}} \{\upsilon_k T_k < \upsilon_k t(\sigma)\} \Big\} \bigcap \Big\{ \bigcap_{k \in \mathcal{U}^C} \{\upsilon T_k \geq \upsilon_k t(\sigma)\} \Big\} \right\} \right] \\
&= \sum_{p=1}^{\frac{K}{3}} \sum_{\mathcal{U} \subset \{1,\dots,K\}:|\mathcal{U}|=p} 1 \\
&= \sum_{p=1}^{\frac{K}{3}} \binom{K}{p}.
\end{aligned}$$

Since there are 2^K environments υ, we have

$$\min_\upsilon \mathbb{P}_\sigma(\Omega_\upsilon) \leq \frac{1}{2^K} \sum_\upsilon \mathbb{P}_\sigma(\Omega_\upsilon) \leq \frac{1}{2^K} \sum_{p=1}^{\frac{K}{3}} \binom{K}{p}.$$

Note that $\frac{1}{2^K}\sum_{p=1}^{\frac{K}{3}} \left(\begin{array}{c} K \\ p \end{array} \right) = \mathbb{P}(\sum_{k=1}^K X_k \leq \frac{K}{3})$ where $(X_1, \ldots, X_K)$ are K independent Bernoulli random variables of parameter $1/2$. By Chernoff-Hoeffding's inequality, we have $\mathbb{P}(\sum_{k=1}^K X_k \leq \frac{K}{3}) = \mathbb{P}(\frac{1}{K}\sum_{k=1}^K X_k - \frac{1}{2} \leq \frac{K}{6}) \leq \exp(-K/72)$. Thus there exists $\upsilon_{\min}$ such that $\mathbb{P}_\sigma(\Omega_{\upsilon_{\min}}) \leq \exp(-K/72)$.

Let us write $p = \mathbb{P}_{\upsilon_{\min}}(\Omega_{\upsilon_{\min}})$ and $p_\sigma = \mathbb{P}_\sigma(\Omega_{\upsilon_{\min}})$. Let $kl(a,b) = a\log(\frac{a}{b}) + (1-a)\log(\frac{1-a}{1-b})$ denote the KL for Bernoulli distributions with parameters a and b. Note that because $\forall \Omega$, $KL(\mathbb{P}_{\upsilon_{\min}}(.|\Omega), \mathbb{P}_\sigma(.|\Omega)) \geq 0$, we have

$$kl(p, p_\sigma) \leq KL(\mathbb{P}_{\upsilon_{\min}}, \mathbb{P}_\sigma).$$

From that we deduce that $p(\log(p) - \log(p_\sigma)) + (1-p)(\log(1-p) - \log(1-p_\sigma)) \leq KL(\mathbb{P}_{\upsilon_{\min}}, \mathbb{P}_\sigma)$, which leads to

$$p \leq \max(\frac{36}{K}\Big(KL(\mathbb{P}_{\upsilon_{\min}}, \mathbb{P}_\sigma)\Big), \exp(-K/72)). \tag{25}$$

Let us now consider any environment (υ). Let $R_t = (r_1, \ldots, r_t)$ be the sequence of observations, and let $\mathbb{P}_\upsilon^t$ be the law of R_t for environment $M(\upsilon)$. Note first that $\mathbb{P}_\upsilon = \mathbb{P}_\upsilon^n$. Adapting the chain rule for Kullback-Leibler divergence, we get

$$\begin{aligned}
&KL(\mathbb{P}_\upsilon^n, \mathbb{P}_\sigma^n) \\
&= KL(\mathbb{P}_\upsilon^1, \mathbb{P}_\sigma^1) + \sum_{t=2}^n \sum_{R_{t-1}} \mathbb{P}_\upsilon^{t-1}(R_{t-1}) KL(\mathbb{P}_\upsilon^t(.|R_{t-1}), \mathbb{P}_\sigma^t(.|R_t)) \\
&= KL(\mathbb{P}_\sigma^1, \mathbb{P}_\upsilon^1) + \sum_{t=2}^n \Big[\sum_{R_{t-1}|\upsilon_{I_t}=+1} \mathbb{P}_\sigma^{t-1}(R_{t-1}) kl(\mu+\alpha, \mu) + \sum_{R_{t-1}|\upsilon_{I_t}=-1} \mathbb{P}_\sigma^{t-1}(R_{t-1}) kl(\mu-\alpha, \mu) \Big] \\
&= kl(\mu-\alpha, \mu)\mathbb{E}_\upsilon[\sum_{k:\upsilon_k=-1} T_k] + kl(\mu+\alpha, \mu)\mathbb{E}_\upsilon[\sum_{k:\upsilon_k=+1} T_k].
\end{aligned}$$

We thus have, using the property that $kl(a,b) \leq \frac{(a-b)^2}{b(1-b)}$,

$$\begin{aligned}
KL(\mathbb{P}_\upsilon, \mathbb{P}_\sigma) &= kl(\mu-\alpha, \mu)\mathbb{E}_\upsilon[\sum_{k:\upsilon_k=-1} T_k] + kl(\mu+\alpha, \mu)\mathbb{E}_\upsilon[\sum_{k:\upsilon_k=+1} T_k] \\
&\leq \mathbb{E}_\sigma[\sum_{k \leq K} T_k] \frac{\alpha^2}{\mu(1-\mu)} \\
&= E_\sigma[\sum_{k \leq K} T_k] \frac{\alpha^2}{\sigma^2}.
\end{aligned}$$

Note that for an algorithm in $\mathbb{A}_{opt}$, we have $\sum_{k=1}^{K} T_k \leq T_k \leq K\Big(\frac{\sigma_{+\alpha}}{K\sigma_{-\alpha}+\sqrt{3}K/4}\Big)n$. Since $\alpha = \frac{\mu}{2}$ and $0 < \mu \leq \frac{1}{2}$ we have

$$\begin{aligned} KL(\mathbb{P}_\upsilon, \mathbb{P}_\sigma) &\leq \Big(K\frac{\sigma_{+\alpha}}{K\sigma_{-\alpha} + \sqrt{3}K/4}\Big)\frac{\alpha^2}{\sigma^2}n \\ &\leq 4\sigma_{+\alpha}\frac{\alpha^2}{\sigma^2}n \\ &\leq 8\frac{\alpha^2}{\sigma}n, \end{aligned}$$

We thus deduce using Equation 25

$$\begin{aligned} \mathbb{P}_{\upsilon_{\min}}(\Omega_{\upsilon_{\min}}) = p &\leq \max(\frac{18}{K}\Big(KL(\mathbb{P}_{\upsilon_{\min}}, \mathbb{P}_\sigma)\Big), \exp(-K/72)) \\ &\leq \frac{144}{K}\frac{\alpha^2}{\sigma}n. \end{aligned}$$

Now choose $\sigma \leq \frac{1}{7}(\frac{K}{n})^{1/3}$ (as $\alpha = \frac{\mu}{2} = \frac{\sigma^2}{2}$). Note that this implies that $\mathbb{P}_{\upsilon_{\min}}(\Omega_{\upsilon_{\min}}) \leq \frac{1}{2}$.

Let $\omega \in \Omega^c_{\upsilon_{\min}}$. We know that for ω, there are at least $\frac{K}{3}$ arms among the K first which are not pulled correctly: either $\frac{K}{6}$ arms among the arms with parameter $\mu - \alpha$ or among the arms with parameter $\mu + \alpha$ are not pulled correctly. Assume that for this fixed ω, there are $\frac{K}{6}$ arms among the arms with parameter $\mu - \alpha$ which are not pulled correctly. Let $\mathcal{U}(\omega)$ be this subset of arms.

We write $\Delta T = \sum_{k\in\mathcal{U}} T_k - \frac{K}{6}t(\sigma_{-\alpha})$ the number of times those arms are over pulled. Note that on ω we have $\Delta T \geq \frac{K}{6}t(\sigma) - t(\sigma_{-\alpha})$. We have

$$\begin{aligned} \Delta T = \frac{K}{6}t(\sigma) - \frac{K}{6}t(\sigma_{-\alpha}) &= \frac{1}{6}\frac{K\sigma}{K\sigma + K/2}n - \frac{1}{6}\frac{K\sigma_{-\alpha}}{\sum_{i=1}^{K}\sigma_{\upsilon_i\alpha} + K/2}n \\ &\geq \frac{1}{6}\frac{K\sigma}{K\sigma + K/2}n - \frac{1}{6}\frac{K\sigma/\sqrt{2}}{\sqrt{3}K\sigma/\sqrt{2} + K/2}n \\ &\geq \frac{1}{6}\frac{1}{K\sigma + K/2}\frac{1}{\sqrt{3}K\sigma/\sqrt{2} + K/2}\Big(K^2\sigma/2 - K^2\sigma/2\sqrt{2}\Big)n \\ &\geq \frac{1}{2}(1 - 1/\sqrt{2})\sigma n \\ &\geq \frac{1}{35}K^{1/3}n^{2/3} \end{aligned}$$

Thus on ω, the regret is such that

$$
\begin{aligned}
R_{n,v_{\min}}(\omega) &\geq \sum_{k=1}^{3K} \frac{w_k^2\sigma_k^2}{T_k(\omega)} - \frac{1}{(2K)^2}\frac{\left(\sum_{i=1}^K \sigma_{v_i\alpha} + K/2\right)^2}{n} \\
&\geq \sum_{k\in\mathcal{U}(\omega)} \frac{w_k^2\sigma_k^2}{T_k(\omega)} + \sum_{k\in\mathcal{U}(\omega)^C} \frac{w_k^2\sigma_k^2}{T_k(\omega)} - \frac{1}{(2K)^2}\frac{\left(\sum_{i=1}^K \sigma_{v_i\alpha} + K/2\right)^2}{n} \\
&\geq \frac{1}{K^2}\frac{K}{6}\frac{\sigma_{-\alpha}^2}{t_k(\sigma_{-\alpha}) + 6\Delta T/K} + \frac{\left(\sum_{i=1}^K \sigma_{v_i\alpha} - K\sigma_{-\alpha}/6 + K/2\right)^2}{(2K - K/6)^2(n-\Delta T)} - \frac{1}{(2K)^2}\frac{\left(\sum_{i=1}^K \sigma_{v_i\alpha} + K/2\right)^2}{n} \\
&\geq \frac{1}{(2K)^2}\frac{\left(\sum_{i=1}^K \sigma_{v_i\alpha} + K/2\right)^2}{n}\frac{1 + \left(\frac{\left(\sum_{i=1}^K \sigma_{v_i\alpha}+K/2\right)\Delta T}{\left(K\sigma_{-\alpha}/6\right)n} - \frac{\left(\sum_{i=1}^K \sigma_{v_i\alpha}+K/2\right)\Delta T}{\left(\sum_{i=1}^K \sigma_{v_i\alpha}-K\sigma_{-\alpha}/6+K/2\right)n}\right)}{\left(1 + \frac{6\Delta T\left(\sum_{i=1}^K \sigma_{v_i\alpha}+K/2\right)}{K\sigma_{-\alpha}n}\right)\left(1 - \frac{\left(\sum_{i=1}^K \sigma_{v_i\alpha}+K/2\right)\Delta T}{\left(\sum_{i=1}^K \sigma_{v_i\alpha}-K\sigma_{-\alpha}/6+K/2\right)n}\right)} \\
&\quad - \frac{1}{(2K)^2}\frac{\left(\sum_{i=1}^K \sigma_{v_i\alpha} + K/2\right)^2}{n} \\
&\geq \frac{1}{(2K)^2}\frac{\left(\sum_{i=1}^K \sigma_{v_i\alpha} + K/2\right)^2}{n}\frac{\left(\frac{\left(\sum_{i=1}^K \sigma_{v_i\alpha}+K/2\right)\Delta T}{\left(\sum_{i=1}^K \sigma_{v_i\alpha}-K\sigma_{-\alpha}/6+K/2\right)n}\right)\left(\frac{\left(\sum_{i=1}^K \sigma_{v_i\alpha}+K/2\right)\Delta T}{\left(K\sigma_{-\alpha}/6\right)n}\right)}{\left(1 + \frac{6\Delta T\left(\sum_{i=1}^K \sigma_{v_i\alpha}+K/2\right)}{K\sigma_{-\alpha}n}\right)\left(1 - \frac{\left(\sum_{i=1}^K \sigma_{v_i\alpha}+K/2\right)\Delta T}{\left(\sum_{i=1}^K \sigma_{v_i\alpha}-K\sigma_{-\alpha}/6+K/2\right)n}\right)} \\
&\geq C\frac{(\Delta T)^2}{n^3\sigma} \\
&\geq C\frac{K^{1/3}}{n^{4/3}},
\end{aligned}
$$

where C is a numerical constant. Note that for events ω where there are $\frac{K}{6}$ arms among the arms with parameter $\mu+\alpha$ which are not pulled correctly, the same result holds.

Note finally that $\mathbb{P}(\Omega_{v_{\min}}^c) \geq 1/2$. We thus have that the regret is bigger than

$$
\begin{aligned}
\mathbb{E}R_{n,v_{\min}} &\geq \sum_{\omega\in\Omega_{v_{\min}}^c} R_{n,v_{\min}}(\omega)\mathbb{P}_{v_{\min}}(\omega) \\
&\geq \sum_{\omega\in\Omega_{v_{\min}}^c} C\frac{K^{1/3}}{n^{4/3}}\mathbb{P}_{v_{\min}}(\omega) \\
&\geq \frac{1}{2}C\frac{K^{1/3}}{n^{4/3}},
\end{aligned}
$$

which proves the lower bound for deterministic algorithms. Now the extension to randomized algorithms is straightforward: any randomized algorithm can be seen as a static (i.e., does not depend on samples) mixture of deterministic algorithms (which can be defined before the game starts). Each deterministic algorithm satisfies the lower bound above in expectation, thus any static mixture does so too.

Appendix E. Large deviation inequalities for independent sub-Gaussian random variables

We first state Bernstein inequality for large deviations of independent random variables around their mean.

Lemma 14 *Let $(X_1, \ldots, X_n)$ be n independent random variables of mean $(\mu_1, \ldots, \mu_n)$ and of variance $(\sigma_1^2, \ldots, \sigma_n^2)$. Assume that there exists $b > 0$ such that for any $\lambda < \frac{1}{b}$, for any $i \leq n$, it holds that $\mathbb{E}\Big[\exp(\lambda(X_i - \mu_i))\Big] \leq \exp\left(\frac{\lambda^2 \sigma_i^2}{2(1-\lambda b)}\right)$. Then with probability $1-\delta$*

$$|\frac{1}{n}\sum_{i=1}^n X_i - \frac{1}{n}\sum_{i=1}^n \mu_i| \leq \sqrt{\frac{2(\frac{1}{n}\sum_{i=1}^n \sigma_i^2)\log(2/\delta)}{n}} + \frac{b\log(2/\delta)}{n}.$$

Proof If the assumptions of Lemma 14 are verified, then

$$\begin{aligned}
\mathbb{P}\Big(\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i \geq n\upsilon\Big) &= \mathbb{P}\left[\exp\left(\lambda(\textstyle\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i)\right) \geq \exp(n\lambda\upsilon)\right] \\
&\leq \mathbb{E}\left[\frac{\exp\left(\lambda(\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i)\right)}{\exp(n\lambda\upsilon)}\right] \\
&\leq \textstyle\prod_{i=1}^n \mathbb{E}\left[\frac{\exp\left(\lambda(X_i-\mu_i)\right)}{\exp(\lambda\upsilon)}\right] \\
&\leq \exp(\tfrac{\lambda^2}{2}\textstyle\sum_{i=1}^n \frac{\sigma_i^2}{2(1-\lambda b)} - n\lambda\upsilon).
\end{aligned}$$

By setting $\lambda = \frac{n\upsilon}{\sum_{i=1}^n \sigma_i^2 + bn\upsilon}$ we obtain

$$\mathbb{P}\Big(\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i \geq n\upsilon\Big) \leq \exp(-\frac{n^2\upsilon^2}{2(\sum_{i=1}^n \sigma_i^2 + bn\upsilon)}).$$

By an union bound we obtain

$$\mathbb{P}\Big(|\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i| \geq n\upsilon\Big) \leq 2\exp(-\frac{n^2\upsilon^2}{2(\sum_{i=1}^n \sigma_i^2 + bn\upsilon)}).$$

This means that with probability $1-\delta$,

$$|\frac{1}{n}\sum_{i=1}^n X_i - \frac{1}{n}\sum_{i=1}^n \mu_i| \leq \sqrt{\frac{2(\frac{1}{n}\sum_{i=1}^n \sigma_i^2)\log(2/\delta)}{n}} + \frac{b\log(2/\delta)}{n}.$$

■

We also state the following Lemma on large deviations for the variance of independent random variables.

Lemma 15 *Let $(X_1, \ldots, X_n)$ be n independent random variables of mean $(\mu_1, \ldots, \mu_n)$ and of variance $(\sigma_1^2, \ldots, \sigma_n^2)$. Assume that there exists $b > 0$ such that for any $\lambda < \frac{1}{b}$, for any $i \leq n$, it holds that $\mathbb{E}\Big[\exp(\lambda(X_i - \mu_i))\Big] \leq \exp\left(\frac{\lambda^2\sigma_i^2}{2(1-\lambda b)}\right)$ and also $\mathbb{E}\Big[\exp(\lambda(X_i-\mu_i)^2 - \lambda\sigma_i^2)\Big] \leq \exp\left(\frac{\lambda^2\sigma_i^2}{2(1-\lambda b)}\right)$.*

Let $V = \frac{1}{n}\sum_i(\mu_i - \frac{1}{n}\sum_i \mu_i)^2 + \frac{1}{n}\sum_n \sigma_i^2$ be the variance of a sample chosen uniformly at random among the n distributions, and $\hat{V} = \frac{1}{n}\sum_{i=1}^n \left(X_i - \frac{1}{n}\sum_{j=1}^n X_j\right)^2$ the corresponding empirical variance. Then with probability $1-\delta$,

$$|\sqrt{\hat{V}} - \sqrt{V}| \leq 2\sqrt{\frac{(1+3b+4V)\log(2/\delta)}{n}}.$$

Proof By decomposing the estimate of the empirical variance in bias and variance, we obtain with probability $1-\delta$

$$\begin{aligned}
\hat{V} =& \frac{1}{n}\sum_i (X_i - \frac{1}{n}\sum_j \mu_j)^2 - (\frac{1}{n}\sum_i X_i - \frac{1}{n}\sum_i \mu_i)^2 \\
=& \frac{1}{n}\sum_i (X_i - \mu_i)^2 + 2\frac{1}{n}\sum_i (X_i - \mu_i)\frac{1}{n}\sum_i (\mu_i - \frac{1}{n}\sum_j \mu_j) \\
&+ \frac{1}{n}\sum_i (\mu_i - \frac{1}{n}\sum_j \mu_j)^2 - (\frac{1}{n}\sum_i X_i - \frac{1}{n}\sum_i \mu_i)^2 \\
=& \frac{1}{n}\sum_i (X_i - \mu_i)^2 + \frac{1}{n}\sum_i (\mu_i - \frac{1}{n}\sum_j \mu_j)^2 - (\frac{1}{n}\sum_i X_i - \frac{1}{n}\sum_i \mu_i)^2.
\end{aligned}$$

We then have by the definition of V that with probability $1-\delta$

$$\hat{V} - V = \frac{1}{n}\sum_{i=1}^n (X_i - \mu_i)^2 - \frac{1}{n}\sum_{i=1}^n \sigma_i^2 - (\frac{1}{n}\sum_i X_i - \frac{1}{n}\sum_i \mu_i)^2. \tag{26}$$

If the assumptions of Lemma 15 are verified, we have with probability $1-\delta$

$$\begin{aligned}
\mathbb{P}\Big(\sum_{i=1}^n (X_i-\mu_i)^2 - \sum_{i=1}^n \sigma_i^2 \geq n\upsilon\Big) &= \mathbb{P}\left[\exp\left(\lambda(\sum_{i=1}^n |X_i-\mu_i|^2 - \sum_{i=1}^n \sigma_i^2)\right) \geq \exp(n\lambda\upsilon)\right] \\
&\leq \mathbb{E}\left[\frac{\exp\left(\lambda(\sum_{i=1}^n |X_i-\mu_i|^2 - \sum_{i=1}^n \sigma_i^2)\right)}{\exp(n\lambda\upsilon)}\right] \\
&\leq \prod_{i=1}^n \mathbb{E}\left[\frac{\exp\left(\lambda(|X_i-\mu_i|^2 - \sigma_i^2)\right)}{\exp(\lambda\upsilon)}\right] \\
&\leq 2\exp(\frac{\lambda^2}{2}\sum_{i=1}^n \frac{\sigma_i^2}{2(1-\lambda b)} - n\lambda\upsilon).
\end{aligned}$$

If we take $\lambda = \frac{n\upsilon}{\sum_{i=1}^n \sigma_i^2 + nb\upsilon}$ we obtain with probability $1-\delta$

$$\mathbb{P}\Big(\sum_{i=1}^n (X_i-\mu_i)^2 - \sum_{i=1}^n \sigma_i^2 \geq n\upsilon^2\Big) \leq \exp(-\frac{n^2\upsilon^2}{2(\sum_{i=1}^n \sigma_i^2 + bn\upsilon)}). \tag{27}$$

By a union bound we get with probability $1-\delta$ that

$$\mathbb{P}\Big(|\sum_{i=1}^{n}(X_i-\mu_i)^2-\sum_{i=1}^{n}\sigma_i^2|\geq n\upsilon\Big)\leq 2\exp(-\frac{n^2\upsilon^2}{2(\sum_{i=1}^{n}\sigma_i^2+bn\upsilon)}).$$

This means that with probability $1-\delta$,

$$|\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu_i)^2-\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2|\leq\sqrt{\frac{2(\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2)\log(2/\delta)}{n}}+\frac{b\log(2/\delta)}{n}. \tag{28}$$

Finally, by combining Equations 26 and 28 with Lemma 14, we obtain with probability $1-\delta$

$$\begin{aligned}|\hat{V}-V|&\leq\frac{4(\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2)\log(2/\delta)}{n}+\frac{2b^2\log(2/\delta)^2}{n^2}+\sqrt{\frac{2(\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2)\log(2/\delta)}{n}}+\frac{b\log(2/\delta)}{n}\\&\leq\sqrt{\frac{2(\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2)\log(2/\delta)}{n}}+\frac{(3b+4\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2)\log(2/\delta)}{n}\\&\leq\sqrt{\frac{2V\log(2/\delta)}{n}}+\frac{(3b+4V)\log(2/\delta)}{n},\end{aligned}$$

when $n\geq b\log(2/\delta)$ and because $V\geq\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2$.

This implies with probability $1-\delta$ that

$$\begin{aligned}&V-\sqrt{\frac{2V\log(2/\delta)}{n}}+\frac{\log(2/\delta)}{2n}\leq\hat{V}+\frac{(3b+4V)\log(2/\delta)}{n}+\frac{\log(2/\delta)}{2n}\\&\Leftrightarrow\sqrt{V}-\sqrt{\frac{\log(2/\delta)}{2n}}\leq\sqrt{\hat{V}+\frac{(1+3b+4V)\log(2/\delta)}{n}}\\&\Rightarrow\sqrt{V}-\sqrt{\frac{\log(2/\delta)}{2n}}\leq\sqrt{\hat{V}}+\sqrt{\frac{(1+3b+4V)\log(2/\delta)}{n}}\\&\Rightarrow\sqrt{V}\leq\sqrt{\hat{V}}+2\sqrt{\frac{(1+3b+4V)\log(2/\delta)}{n}}.\end{aligned}$$

On the other hand, we have also with probability $1-\delta$

$$\begin{aligned}&\hat{V}\leq V+\sqrt{\frac{2V\log(2/\delta)}{n}}+\frac{(3b+4V)\log(2/\delta)}{n}\\&\Rightarrow\sqrt{\hat{V}}\leq\sqrt{V}+2\sqrt{\frac{(1+3b+4V)\log(2/\delta)}{n}}.\end{aligned}$$

Finally, we have with probability $1-\delta$

$$|\sqrt{\hat{V}}-\sqrt{V}|\leq 2\sqrt{\frac{(1+3b+4V)\log(2/\delta)}{n}}. \tag{29}$$

■